



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN

(A Project of Sri Marudhar Kesari Jain Trust)

Permanently affiliated to Thiruvalluvar University || Recognized Under Sec. 2(f) & 12(B)
of the UGC Act, 1956 || Re-Accredited with A Grade by NAAC (3rd Cycle)
|| An ISO 9001:2015 Certified institution
Vaniyambadi 635751, TamilNadu

**PG Department of Computer Applications
and
PG & Research Department of Computer Science**

Proceedings of AICTE Sponsored

TWO DAYS NATIONAL LEVEL VIRTUAL CONFERENCE

on

**RECENT TRENDS IN RESEARCH &
INFORMATION TECHNOLOGY**

22.09.2021 & 23.09.2021

ISBN Number : 978-93-91387-20-4

www.mkjc.in



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI

(A Project of Sri Marudhar Kesari Jain Trust)



PG Department of Computer Applications &

PG & Research Department of Computer Science

Proceedings

on

AICTE Sponsored Two Days National Level Virtual Conference



CHIEF EDITOR

Dr. M. Inbavalli

Principal & Convenor

EDITORS

Dr. K. Priya

Research Coordinator & Organizing Secretary

Ms. M. Prithi

Assistant Professor & Organizing Secretary

ISBN: 978-93-91387-20-4

Management Trustees



Sri. M. Vimmal Chand Jain
President



Sri. J. Rathanalal Jain
Vice President



Sri. C. Lickmichand Jain
Secretary



Sri. N. Sripal Kumar Jain
Treasurer



Sri. M. Sudarsan Kumar Jain
Trustee



Sri. K. Rajesh Kumar Jain
Trustee



Sri. N. Lalith Kumar Jain
Trustee



Sri. K. Anand Kumar Jain
Trustee



Sri. U. Rishab Jain
Trustee

ABOUT COLLEGE

Marudhar Kesari Jain College for Women, Vaniyambadi(26 Years of Excellence)MKJC, established in the year 1994, Recognized under Sec. 2(f) & 12(B) of the UGC Act1956, approved by Govt. of Tamilnadu, and is Permanently Affiliated to Thiruvalluvar University. It has been re-accredited with "A" grade by NAAC, Bengaluru(3rd Cycle) in the year 2019 and certified by BSI as an ISO9001:2015 Certified Institution. The focus of the College is RURAL DEVELOPMENT by enhancing Women Education, especially those from inclusive society with RIGHT KNOWLEDGE and makes them self - sufficient to develop holistic personality and excellence in their walks of life. The college is offering 16 UG Courses, 12 PG Courses, 7 M.Phil Courses and 6 Ph.D Courses with 2834 students, 181 Faculties and 159 Non Teaching staff. TamilNadu State Private Colleges Association-Skill Indian Award 2021 awarded Our College with the “Best College Award for Excellence in Education” for the endeavour and supremacy. The MKJC takes its proud history of achievements in promoting academic and professional excellence that has led to strong outcomes in securing above 90% of results, 5 Gold Medals and 39 University Ranks in the year 2019 and also 100% of results and 5 Gold Medals in the Academic year 2020. For over 26 years MKJC has achieved a long list of accolades with stories of success within the region.



ABOUT THE DEPARTMENTS

PG & RESEARCH DEPARTMENT OF COMPUTER SCIENCE

The Department of Computer Science was established in 1994. Department is currently offering Under Graduate and Post Graduate with Research Programmes with a vision to prepare young women to acquire the opportunities in the field of Information Technology.

The department is well equipped laboratories with state of the art computing facilities with internet facilities and 13 dedicated faculties to enhance the research and teaching activities.

Courses Offered: B.Sc., M.Sc., M.Phil., Ph.D

PG DEPARTMENT OF COMPUTER APPLICATIONS

The Department of Computer Applications laid its corner stone in the year 1998 with B.C.A Computer Applications and spread its wings in post graduate on M.C.A in the year 2002, approved by AICTE. The department has dedicated faculties, state – of –art infrastructure and well equipped laboratories which is being constantly updated with latest computers & software's.

The department prepares students to design, develop and implement software applications. Department of computer applications emphasis on the latest programming languages, cutting edge technologies and tools to make the students ready for the IT industry.

Courses Offered: BCA & MCA





MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri Marudhar Kesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.

Shri.M.Vimmal Chand Jain

President



MESSAGE

I am very happy to know that a National Conference on “Recent Innovations in Computer Science” organized by Department of Computer Science and Computer Applications of this institution.

I understand that many eminent Resource persons and elite scholars are taking part in the deliberations of the conference.

I am confident that the conference would provide a unique forum for the participants to exchange their expertise knowledge in the recent development in the field of Computer Science and Computer Applications.

I extend my warm greetings and best wishes to the convener, organizing committee, staff and students of the Department of Computer Science and Computer Applications. I wish this National Conference a great success and I also wish them to achieve everything possible to take the institution towards higher goals and broader horizons.

Shri.M.Vimmal Chand Jain

President

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN,VANIYAMBADI.

Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(TirupatturDist) **Phone: 04174-**

224300/225300 Email:principal@mkjc.in



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri Marudhar Kesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.

Shri.C.LickmichandJi Jain

Secretary



MESSAGE

I am very much delighted to know that the faculty of PG and Research Department of Computer Science and PG Department of Computer Applications is organizing Two days National Level Virtual Conference on **“Recent Trends in Research & Information Technology”** in our institution.

I am sure that it will provide an opportunity for eminent Scientists and Researchers to share their experiences and exchange their innovative ideas in **“Recent Trends in Research & Information Technology”**.

This National Conference would also enrich the knowledge of the participants in the vital areas of Computer Science which in turn would help the researchers and students to probe into the active research and contribute valuable research findings.

I hope all the distinguished invites, delegates and scholars would feel benefited out of the academic deliberations would enjoy their visit. My hearty congratulations to the convener, organizing committee, all staff and students.

I wish the conference a great success.

Shri.C.Lickmichand Ji Jain

Secretary

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI.
Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(Tirupattur Dist) **Phone: 04174-224300/225300 Email:principal@mkjc.in**



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri Marudhar Kesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.

Dr. M. Inbavalli

Principal & Convinor



MESSAGE

It is my privilege to express my profound happiness and blessings to the AICTE funded Two days National Level Virtual Conference on **“Recent Trends in Research & Information Technology”** organized by the PG and Research Department of Computer Science and PG Department of Computer Applications of our Institution. I have immense pleasure to say that, this conference will enlighten the evolutionary minds of staff, students and Research Scholars. I am sure that this conference will satisfy the needs and requirements of intellectual participants. I hope that this one day intellectual meet will bring new values and dimensions in the field of Computer Science and Applications.

Recently the importance of Computer Science is gaining its momentum. I hope this Virtual National Conference will provide a platform for the exchange of research knowledge in the field of ‘Computer Science’ and will enlighten the future citizens of India to realize their role in it. I take this opportunity to thank our most respected President, Secretary and Trust Members for given me the opportunity to join this Team of Spirit as Patron. I also congratulate the Head of the Departments of Computer Science and Applications, Research Coordinator, committed staff members and enthusiastic students for arranging such type of innovative, informative and eye opening Conference.

I pray and wish this Virtual National Conference a grand Success.

Dr. M. Inbavalli

Principal & Convinor

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI.

Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(Tirupattur Dist) **Phone: 04174-224300/225300 Email:principal@mkjc.in**



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri Marudhar Kesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.

Mrs. S. Abirami
Head, Department of Computer Applications



MESSAGE

Dear Participants,

Scientific seminars and technical conferences enable Engineers, Technologists and Researchers to keep abreast of the latest development in their area of work. This two days AICTE Sponsored National Virtual Conference on “**Recent Trends in Research & INFORMATION TECHNOLOGY**” held during 22nd September and 23rd September 2021 is yet another milestone for Department of Computer Applications.

I am sure the Keynote session of the Conference will pave a way for new contributions and themes for future issues in diverse topics. Paper presentations sessions will set trends for future work in new direction of innovative research.

This conference is a remarkable outcome of the excellent teamwork and I would like to record my sincere thanks to the President, Vice President, Secretary, Trust Members, Principal, PRO, Convener, Organizing Committee, Faculties and Student of the Department of Computer Applications and Computer Science for their contribution and untiring efforts to make up this technical festival a grand success.

With Best Wishes

Ms.S.Abirami

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI.
Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(Tirupattur Dist) **Phone: 04174-**
224300/225300 Email:principal@mkjc.in



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri Marudhar Kesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.

Ms. L. Hemalatha
Head, Department of Computer Science



MESSAGE

Dear Participants,

It is a great pleasure for me that our Computer Science and Computer Applications department is conducting a this two days AICTE Sponsored National Virtual Conference on “**Recent Trends in Research & Information Technology**” held during 22nd September and 23rd September 2021.

I extend my heartfelt thanks to the President, Vice President, Secretary, Trust Members, Principal, PRO, Convener, Organizing Committee, Faculties and Student of the department of Computer Science and Computer Applications for their endeavour and their efforts in making this conference a grand success.

I hope this conference will be enjoyable, memorable, and productive for participants and looking forward for the technological innovations that result from your networking and discussions.

I wish the conference a grand Success.

With Best Wishes

Ms. L. Hemalatha

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI.
Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(Tirupattur Dist) **Phone: 04174-224300/225300 Email:principal@mkjc.in**



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri Marudhar Kesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.



Dr. K. Priya

Research Coordinator



Ms. M. Prithi

Assistant Professor

EDITORS NOTE

Welcome to the Proceedings of the **AICTE Sponsored Two Days National Level Virtual Conference on “Recent Trends in Research & Information Technology”**. Ensuring a high quality conference with multidisciplinary knowledge sharing.

We would like to express our deepest appreciation to the authors whose technical contributions are presented in these proceedings. It is because of their excellent contributions and hard work that we have been able to prepare these proceedings. The significance of the research presented in this conference represents a step further towards maturity in all the areas.

We would like to thank all our Resource Persons who made all the efforts to synthesize the materials and their wide and rich experiences to deliver distinguished talks and for their valuable suggestions. We would also like to thank our President, Vice President, Secretary, Trust Members, Principal, National Advisory committee, Editorial/Reviewer Committee, track Juries for their great efforts in delivering interactive and excellent tutorials that address the learning needs of all levels, professionals, researchers, Postgraduates and graduates. Finally, we hope that the participants enjoy the outstanding conference program of the AICTE Sponsored Two Days National Level Virtual Conference on Recent Trends in Research & Information Technology organized on 22-Sep-2021 & 23-Sep-2021 in MKJC that we create for them.

The conference provides Paper presentation, selection of the best paper award and the best presenter award.

Dr. K. Priya

Ms. M. Prithi

Organizing Secretaries

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI.

Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(Tirupattur Dist) **Phone: 04174-**

224300/225300 Email:principal@mkjc.in



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri MarudharKesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.



Ms.A.Poornima
Assistant Professor



Ms.K.Suganya
Assistant Professor



Ms.G.Sasi rekha
Assistant Professor



Ms.N.Raghavi
Assistant Professor



Ms.M.Ponnarasi
Assistant Professor



Ms.M. Pragathi
Assistant Professor



Ms.K.Suganya
Assistant Professor



Ms.Syamala
Assistant Professor



Ms.M.Sarumathi
Assistant Professor

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI.

Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(Tirupattur Dist) **Phone: 04174-**

224300/225300 Email:principal@mkjc.in



MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI
(A Project of Sri MarudharKesari Jain Trust)



Permanent affiliated to Thiruvalluvar University|| Recognized Under Sec. 2(f) & 12(B) of the UGC Act, 1956||Re-Accredited with A Grade by NAAC (3rd Cycle)|| An ISO 9001:2015 Certified Institution Vaniyambadi- 635 751, TamilNadu.



Ms.J.Sasirekha
Assistant Professor



Ms.P.Prabavathy
Assistant Professor



Ms.T.Thenmozhi
Assistant Professor



Ms.L.Sumi
Assistant Professor



Ms.P.Monisha
Assistant Professor



Ms.S.Anitha
Assistant Professor



Ms.R.Padmalaatha
Assistant Professor



Ms.L.Shalini
Assistant Professor



Ms.S.Nishanthi
Assistant Professor

MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI.
Marudhar Nagar, Chinnakallupalli, Vaniyambadi-635 751.(Tirupattur Dist) **Phone: 04174-224300/225300 Email:principal@mkjc.in**

INDEX

SNO	TITLE OF THE PAPER	PAGE NO
1	IOT TO CONTROL THE TEMPERATURE OF THE SOLAR PANEL <i>Mr. M. Sanjheevi Raaman¹, Ms.J. Elavarasi ²,Ms.Sivasankari ³</i> <i>Asst.Professor, PG Dept of Computer Application, TN, India</i>	1
2	A STUDY ON NETWORK SECURITY AND CRYPTOGRAPHY <i>Ms.S.Anitha¹, Ms.R.Padmalatha²</i> <i>Asst.Professor, Department of Computer Applications TN ,India</i>	7
3	DETECTION OF MALARIA PARASITES USING DEEP LEARNING TECHNIQUES <i>Ms.G.Sasirekha¹ ,Ms.A.Poornima²</i> <i>Asst Professor, Department of Computer Science TN, India</i>	15
4	THE ROLE OF BIG DATA IN ECOMMERCE <i>A.Karthick Praveen Kumar¹, Sr. Data Architect, AWS,USA</i> <i>Ms.B.Sakthimala²</i> <i>Assistant Professor, Department of BBA, TN, India</i>	22
5	A STUDY ON SCOPE OF HYDERABAD AS THE IT STARTUP HUB OF INDIA <i>T. Tejaswi¹ , Dr. U. Devi Prasad²</i> <i>GITAM (Deemed to be University), Hyderabad Campus, Telangana</i>	27
6	SOCIAL DISTANCING DETECTION AND MONITORING THROUGH DENSITY DETECTION. <i>G. Bharathi¹ Dr.G.Anandharaj², R.Angelin Preethi³</i> <i>Dept of Computer Science, TN, India.</i>	32
7	IOT BASED UNDERGROUND CABLE FAULT DETECTION <i>Mr.Rajath V¹, Swetha Shekarappa .G²</i> <i>Department of Electrical & Electronics Engineering, Bangalore, India</i>	39
8	HYBRID APPROACH OF SET PARTITIONING IN HIERARCHICAL TREES WITH CDF WAVELET FOR MEDICAL IMAGE COMPRESSION <i>P.Jeyanthi¹ , Dr. K.Prabavathy²</i> <i>Assistant Professor,Department of Computer Science, Coimabtoe, India</i>	50
9	SECURED AND ENERGY EFFICIENCY ORIENTED CLUSTER ROUTING ALGORITHM FOR FANET BASED INTERNET OF THINGS <i>Ms.Revathi B¹,Mr.Arulanandam K²</i> <i>Department of Computer Applications, TN, India.</i>	59
10	MOTIVATING FACTORS OF WOMEN ENTREPRENEUR IN CHENNAI CITY <i>Dr.R.Dharmaragini¹ ,Dr.K.Rajamannar², R.Subitha Rani³</i> <i>Dept. of Commerce, TN, India</i>	67

11	<p>USING A MODIFIED MINIMUM SPANNING TREE TO IMPROVE CLUSTER BASED FEATURE SELECTION</p> <p><i>Dr.A.Suresh¹, Mr. A.Kaleemullah², Dr.P.Rizwan Ahmed³</i> <i>Department of Computer Science, TN, India</i></p>	74
12	<p>NUTRITIONAL STATUS AND RISK OF TYPE2 DIABETES AMONG PRE-MENOPAUSAL WOMEN IN KOLAGUR</p> <p><i>V.Shivani Abirami, Assistant Professor, Tamil Nadu, India.</i></p>	82
13	<p>IOT BASED SECURITY SYSTEM FOR HOME AUTOMATION</p> <p><i>Ms. R.Janani¹, Ms. S.Sneka², Ms. K.Nivetha³</i> <i>Dept. of Computer Application,Mkjc, Tamil Nadu, India.</i></p>	94
14	<p>REVIEW PAPER ON VARIOUS SOFTWARE TESTING TECHNIQUES & STRATEGIES</p> <p><i>Ms. K.Lavanya¹, Ms.S.Dhatchayani²</i> <i>Dept. of Computer Applications, Tamil Nadu, India.</i></p>	103
15	<p>FILE SYSTEMS FOR VARIOUS OPERATING SYSTEMS: A REVIEW</p> <p><i>Ms. L. Shalini, MKJC,Tamil Nadu, India</i></p>	114
16	<p>A STUDY ON CYBER SECURITY IN MODERN INTERNET OF THINGS</p> <p><i>Ms.N.Raghavi¹, Ms.M.Ponnarasi², Ms.K.Suganya³</i> <i>Assistant Professor Department of Computer science, TN, India</i></p>	123
17	<p>A STUDY ON ANALYSIS OF E-COMMERCE STRATEGIES FOR TRADITIONAL ORGANISATIONS MOVING ON TO THE INTERNET</p> <p><i>Ms.Shema¹, Dr. G. Deepalakshmi², A.Athiyaveni³</i> <i>Assistant Professor, TN, India</i></p>	131
18	<p>AN OVERVIEW OF CYBER SECURITY ON SOCIAL MEDIA: ISSUES, CHALLENGES AND SOLUTIONS</p> <p><i>Dr. K. Priya¹, P. Vishnu Priya²</i> <i>Department of Computer Science, TN, India.</i></p>	142
19	<p>DIGITAL MARKETING IN ECOMMERCE IN INDIA</p> <p><i>Ms. P.Priyanka, Assistant Professor, Dept. of Computer Applications, TN,India.</i></p>	149
20	<p>5G NETWORK SECURITY TECHNOLOGY AND IT'S CYBER ATTACKS</p> <p><i>Ms.T.Subhashini</i> <i>Assistant Professor, Dept. of Computer Applications.TN, India.</i></p>	160
21	<p>SMART HOME AUTOMATION SYSTEM USING IOT</p> <p><i>Dr.Priya¹, B.Vinisha², C.Kowdelya³, S.Banu Priya⁴</i> <i>Dept. of Computer Application, TN, India.</i></p>	178

22	A STUDY ON IMPACT OF COMPENSATION AND WORKING CONDITION TOWARDS JOB SATISFACTION OF COLLEGE TEACHERS <i>Ms. J. JEEVITHA¹, Dr. C. NITHYA², Department of Commerce, Tamil Nadu, India.</i>	183
23	AN ANALOGOUS STUDY BETWEEN WATERFALL AND INCREMENTAL SOFTWARE DEVELOPMENT LIFE CYCLE MODEL <i>Ms. S. Abirami, HOD, Department of Computer Application, Tamil Nadu, India</i>	194
24	ANALYSIS OF HOUSING DATA TO PREDICT HOUSE PRICE USING MULTIPLE LINEAR REGRESSION <i>Dr.G.Dona Rashmi¹, Dr. S.Poongodi² Assistant Professor, Department of Data Analytics (PG), TN, India</i>	203
25	SECURITY IN IOT: OVERVIEW <i>Ms. Uma Mageshwari¹, Dr.R.Santhi² Research & Evaluation Centre, Bharathiar University, Coimbatore</i>	209
26	AN ANALYSIS OF OPINION MINING USING DIFFERENT MACHINE LEARNING ALGORITHMS <i>Mrs. N. Sathyapriya Assistant Professor, Department of Computer Science, Udumalpet ,Tirupur, TN.</i>	213
27	NETWORK SECURITY AND CRYPTOGRAPHY <i>Ms.S.Brindha¹, Ms.V.Tamilselvi² Dept of Computer Science, Tamil Nadu, India.</i>	226
28	A REVIEW ON SECURITY AND PRIVACY ISSUES IN BIG DATA ANALYTICS <i>Ms.J.Shivani, Tamil Nadu, India</i>	231
29	A STUDY ON MACHINE LEARNING FOR WEB VULNERABILITY DETECTION:THE CASE OF CROSS-SITE REQUEST FORGERY <i>Ms.M PRAGATHI¹, Ms.K.Suganya², Assistant Professor Tamil Nadu, India.</i>	238
30	IMPRESSION OF COVID IN E-COMMERCE PLATFORM FOR THE DURATION OF EPIDEMIC <i>Mrs. Abarna Sri R¹, Mrs. B.Sasikala², Assistant Professor Department of Computer Science, TN, India.</i>	246
31	SECURITY ANALYSIS AND PRESERVING BLOCK-LEVEL DATA PREVENTION IN CLOUD STORAGE SERVICES <i>Ms. R.Sandhiya¹, Ms.P.Prabavathy², Assistant Professor, Tamil Nadu, India.</i>	252

32	A REVIEW ON IMPROVING SMART ATM SECURITY LEVELS <i>Ms. P. Indumathy¹, Ms. V. Kaviya², Ms. V. Nisha³ Dept. of Computer Science, Tamil Nadu, India.</i>	259
33	AN OVERVIEW OF BLOCKCHAIN TECHNOLOGY <i>Ms. R. Keerthana¹, Ms. D. Narmadha², Ms. S. Soundharya³ Dept. of Computer Applications, Tamil Nadu, India.</i>	264
34	REVIEW ON ARTICLE ABOUT CLOUD COMPUTING <i>G. Sasirekha¹, S. Deepthi², K. Sathya priya³ Dept. of Computer Science TN, India.</i>	274
35	ALGORITHM TO IMPLEMENT HIGHER ORDER MINING WITH LOGISTIC REGRESSION AND VARMA METHOD <i>Mrs. M. Deepanayaki,¹ Dr. Vidyaathulasiraman² Department of Computer Science, TN, India</i>	282
36	A REVIEW PAPER ON CLOUD COMPUTING IN INDUSTRIES <i>Ms. Kaviyanjali¹, Ms. P. Arishitha² Dept. of Computer Application TN, India.</i>	289
37	REVIEW ON ARTICLE ABOUT ARTIFICIAL INTELLIGENCE <i>Ms. S. Chandra Praba¹, Ms. P. Yogeshwari², Ms. R. S. Batma Jayani³ Dept. of Computer Science, Tamil Nadu, India</i>	295
38	BIO MEDICAL INFORMATICS <i>Ms. M. Haripriya¹, Ms. B. Swathi², Ms. B. Sneha³ Dept. Of Computer Application, TN, India</i>	304
39	DATA SCIENCE AND BIG DATA <i>R. Harshavardhini¹, . Lokeshwari², S. Sowmiya³ Dept. of computer Science TN, India.</i>	308
40	COMPUTER GRAPHICS, VISION, ANIMATION AND GAME SCIENCE <i>Ms. Nabeela Nousheen¹, Ms. P. Vaideeswari², Ms. A. Preethika³ Dept. of Computer Science, Tamil Nadu, India.</i>	319
41	DIGITAL TECHNOLOGY ON PSYCHOLOGY TREATMENTS <i>T. Shaktipriya¹, K. Sneha², R. Lekshana Priya³ Dept. of computer Science, TN, India.</i>	331
42	ARTIFICIAL INTELLIGENCE TRENDS IN DEEP LEARNING ANALYTICS <i>Mrs. S. Sowmiyasree¹, Mrs. V. Deepa² Assistant Professor, Tamil Nadu, India.</i>	343
43	INTERNET OF THINGS (IoT) BASED NETWORK SECURITY, CYBER SECURITY & DATA SECURITY <i>A. Indhumathi¹, R. Ayesha Riyaz², B. Pushpavalli³ Dept. of Computer Applications, TN, India.</i>	352

44	A STUDY ON DIGITAL MARKETING AND ITS IMPACTS ON CONSUMER BUYING BEHAVIOUR <i>Ms.S. Ranjitha¹, Ms.K.Divya²</i> <i>Assistant Professor, Tamil Nadu, India.</i>	358
45	SECURITY AND PRIVACY BY DESIGN FOR AHA-IOT APPLICATIONS AND SERVICES <i>Ms.P MONISHA,</i> <i>Assistant Professor of Computer Application, TN, India.</i>	367
46	NETWORK SECURITY AND CRYPTOGRAPHY <i>Ms.S.Bhavaniv¹, V.Kavi Bharathil², L.Sathiya³</i> <i>Dept. of Computer Science, TN, India.</i>	373
47	NETWORKING & SECURITY <i>Ms.J.Varshini¹, Ms.S.Soniya², MS.U.Kayalvizhi³</i> <i>Dept. of Computer Applications, TN, India</i>	379
48	BIG DATA ANALYTICS <i>Ms.M. Hemavathi¹, Ms.B. Vedhasri², Ms.B.Harini³</i> <i>Dept. of computer Application TN,India.</i>	385
49	CYBER SECURITY IN HEALTH CARE <i>Ms. V.Sasirekha</i> <i>Dept. of Computer Applications, TN, India.</i>	393
50	DESIGN AND IMPLEMENTATION OF MULTI-PATH ACCESS TRANSMISSION CONTROL PROTOCOL <i>Dr.S.Vijayarangam¹</i> <i>Associate Professor Hyderabad, Telangana, India.</i> <i>J.Sasirekha²</i> <i>Assistant Professor, Tamil Nadu, India.</i>	399
51	ECOMMERCE AND DIGITAL MARKETING <i>Devi Shree.N¹, Keerthana.A², Roja.D³</i> <i>Dept. of Computer Applications TN, India</i>	414
52	ARTIFICIAL SIGHT FOR THE VISUALLY IMPAIRED PEOPLE (WALKING STICK). <i>Ms. L. Sumi</i> <i>Assistant Professor, Department Of Computer Applications, TN, India</i>	421
53	ROBOTIC TRANSPOSE INTENT PARTITION FOR EXPLICITLY TRANSPOSE CAMERAS <i>Ms.Prithi.M¹, Dr. Tamizharasi. K²</i> <i>Dept. of Computer Science, TN, India</i>	427
54	THERAPEUTIC EFFECT OF APPLE CIDER VINEGAR ON TYPE II DIABETES MELLITUS <i>Dr.R.ShobanaDevi, Dept. of ND, TN, India</i>	435
55	COVID-19 DETECTION USING MACHINE LEARNING <i>Ms.Asma Farhana.M.N¹, Nehlath Harmain.H.E², Nandhini.V³</i> <i>Dept. of Computer Applications, TN, India</i>	443

56	CINEMA AND LITERATURE: TWO ART FORMS THAT BUILD AND GROW TOGETHER <i>Dr Hemamalini N, Assistant Professor, Dept. of English, TN, India</i>	452
57	PERFORMANCE ANALYSIS OF VARIOUS ENSEMBLE FEATURE SELECTION <i>Ms.Sandrilla R¹, Ms.Savitha Devi. M² Department of Computer Science, TN, India</i>	459
58	DATA MINING – TECHNIQUES, METHODS AND ALGORITHMS: A REVIEW ON TOOLS <i>Ms. L. Hemalatha HOD, Dept. of Computer Science, MKJC, TN, India</i>	479
59	ONLINE TEACHING TOOLS USED DURING THE PANDEMIC PERIOD WITH SPECIAL REFERENCE TO SCHOOLS AND COLLEGES IN VELLORE DISTRICT <i>M.Ashtalakshmi Assistant Professor & Head, Department of Commerce CA, MKJC, TN, India.</i>	488
60	IOT BASED SOLUTION FOR MONITORING OF POLLUTION FRUIT PESTICIDES IN FRESH FRUITS AND VEGITABLES IN MARKET <i>Ms.T.Thenmozhi¹, Ms.S.Nishanthi² Department of Computer Applications,Assistant Professor Marudhar Kesari Jain College For Women,Vaniyambadi, Tamilnadu, India.</i>	497
61	CRYPTOGRAPHY FROM OLDER TO MODERN ERA <i>K. Durgadevi¹ V. Harisha² K.R. Soundarya³ Marudhar Kesari jain college for women,vaniyambadi,TN,India</i>	506
62	BRAIN INTELLIGENCE MODEL: FUSING ARTIFICIAL INTELLIGENCE AND ARTIFICIAL LIFE <i>Dr.Vignesh Ramamoorthy H¹·Ms.Priyadharshini N²·Dr.LovelineZeema J³ Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India</i>	513
63	THE INFLUENCE OF TECHNOLOGY ON ENGLISH LANGUAGE AND LITERATURE <i>Dr.S.Varalakshmi, Vice-Principal Global College of Arts and Science for Women, Veppur,TN,India</i>	518

IOT TO CONTROL THE TEMPERATURE OF THE SOLAR PANEL

*Mr.Sanjheevi Raaman M¹ , Ms. Elavarasi J² , Ms. Sivasankari A³
Assistant Professor,PG Dept of Computer Applications
Shanmuga Industries Arts and Science College,Tamilnadu,India.*

ABSTRACT

Solar photovoltaic systems are a type of renewable energy that is widely used around the world. The efficiency, on the other hand, falls when the temperature of the solar panels rises. A cooling fan can be added on the rear side of the solar panel to prevent this problem and improve efficiency. The effectiveness of a solar system is also reduced by weather conditions and unanticipated events. To address this issue, an IoT (Internet of Things) system was utilised to monitor the solar system's status and manage the cooling fan. Arduino was the primary microprocessor in IoT systems. An IoT system can be easily and cheaply created using Arduino. After the complete system was designed and tested, the efficiency increased by about 4.7 percent. Although it is a tiny photovoltaic system with a capacity of 30W, its efficiency is projected to be boosted in the near future by applying it to a photovoltaic system with a capacity of more than 1kW.

Keywords: Arduino, Cooling fan, Internet of Things, Smart phone, Solar panel, and Wi-Fi

INTRODUCTION

IoT (Internet of Things) is becoming increasingly popular in the household and industry [1]. The Internet of Things (IoT) is a future network technology that allows people, things, things, and things to share information and control each other [2]. This enables remote control and monitoring from afar. Based on this IoT, it can be used to solar power systems that are now in use around the world.

Because a solar power generating system uses solar energy to generate electricity, the power generation efficiency decreases as the temperature of the solar panel rises [3]. At noon, the surface temperature of the solar panel ranges between 60 and 70 degrees. There are numerous approaches to avoid such a loss of efficiency. To begin, cooling water is sprayed on the surface of the solar panel to reduce the temperature. Second, apply a heat sink to lessen the temperature. Finally, utilise a cooling fan to reduce the temperature.

According to the findings of the experiments, the final of these three ways is the best for lowering the temperature of the solar panel.

This cooling fan, however, must be engaged when the temperature of the solar panel rises and deactivated when the temperature falls. Although the temperature can be controlled automatically, when the temperature sensor fails or is not perfect, it is impossible to control the entire system. As a result, the system is manually controlled via monitoring.

Arduino is the microcontroller utilised in this work. Arduino is popular among engineers due to its ease of use and low cost [4]. However, Arduino cannot communicate over long distances on its own. As a result, it operates the cooling fan via Wi-Fi connection between the Smartphone and the Arduino. The Arduino is paired with a Wi-Fi shield, and the Wi-Fi

shield contains a server. Various sensor and switch signals can be delivered to the smart phone and Arduino using this service. This allows you to turn on and off the switch immediately after verifying the temperature of the solar panel. We checked the normal operation of the Arduino and smart phone via Wi-Fi communication after building the actual model, and we also confirmed that the temperature of the solar panel has been reduced. The solar power generation system can generate more efficient electricity through the monitoring and control described above.

IMPORTANT SYSTEMCOMPONENTS

The system is made up of several crucial components. Important components include the Arduino, which serves as the brain, a Wi-Fi shield for Wi-Fi connectivity, an infrared temperature sensor for receiving infrared signals, and relays that activate the cooling fan. Voltage, current sensor, ambient temperature sensor, and LCD are all sub-components. The organic operation of these pieces can appropriately manage the temperature of the solar panel.

Arduino

The Arduino Mega utilised in this setup is depicted in Fig. 1. The primary chip is an atmega2560 with 54 digital input/output pins, 16 analogue input pins, 256kb flash memory, and 8KBSRAM.



Figure 1 shows an Arduino MEGA 2560. And, like C and Java, Arduino is coded in assembly language.

Wi-Fi protection

To communicate with Arduino and the smart phone, the Wi-Fi shield is utilised in conjunction with Arduino.

Infrared sensor

The infrared sensor is a device that detects infrared rays emitted by an item to determine its temperature. Remote measurement is possible, and no failures occur. In this system, a non-contact infrared sensor (MLX90614) is utilised to measure temperatures ranging from -70.01°C to $+382.19^{\circ}\text{C}$. It has a resolution of 0.01°C and communicates through I2C. Figure 2 depicts it.



Figure 2: MLX90614

Relay

The relay receives a signal from the Arduino and controls the cooling fan on and off. The maximum input voltage for AC is 250V and 30V for DC.

ADDITIONAL SENSORS AND DEVICES

Aside from the primary components, there are various sensors. Voltage and current sensors, ambient temperature sensors, cooling fans, and LCD displays are just a few examples. AC-DC converter, 30W solar panel, cradle, and power box are all included.

Fabrication Design

The entire system was designed utilising the above-mentioned components. Various sensors and relays are connected to the Arduino's core, and finally, one system is accomplished. A smart phone application for controlling the Arduino is also being developed. The software was created with user-friendliness in mind.

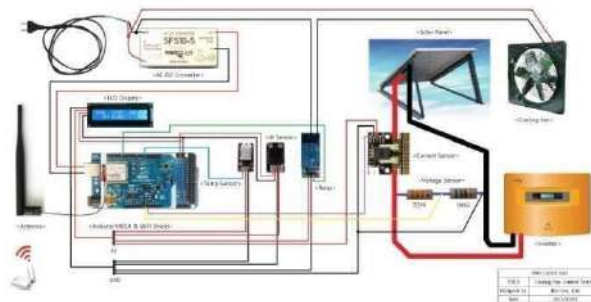


Fig. 3. The Schematic of the System

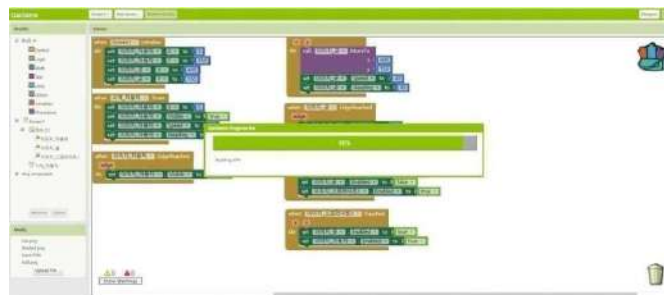


Fig. 4. App Inventor 2 interface

Programming Language

Finally, for the device to function, Arduino requires source code. Wi-Fi connection, LCD display, temperature, voltage, and current sensors, and switch operation are all covered in the source code. The code below only shows the most critical portions. The first is a header section in which you provide the header file to be used in your code.

The second section is for Wi-Fi connections, in which you can configure the SSID, Password, and encryption technique.

```
#define SSID "air" /  
#your wifi network SSID #define KEY "" / #your wifi network password  
#define AUTH "NONE" /  
#your wifi network security (NONE, WEP, WPA, WPA2)  
/*Server part*/  
// HTTP headers always start with a response code (e.g. HTTP/1.1 200 OK)  
// and a content-type so the clientem knows what's coming, then a blankline:  
JSN270.println("HTTP/1.1 200 OK");  
JSN270.println("Content-type:text/html");  
JSN270.println("Refresh: 30");JSN270.println();
```

Completed model

As shown in Fig. 5, a final model was built. Six cooling fans were connected to the solar panel, and a converter was installed to increase the input voltage to 12V. To directly verify the data, an LCD was attached to the main box. Finally, Figure 6 depicts the final application. There is a cooling fan switch button as well as an additional capability to check various data in real time.



Fig. 5. Final model

Relay

The relay receives a signal from the Arduino and controls the cooling fan on and off. The maximum input voltage for AC is 250V and 30V for DC.

SystemCheck

We verified that the created system operates normally. A smart phone was utilised to manage the cooling fan's on/off status, and a heater was employed to simulate sunlight. And, as indicated in Fig.7, use a heater to heat the solar panel.



Fig. 7. Testing set-up

The temperature of the solar panel was kept above 57°C, while the ambient temperature was kept at 28°C. The temperature reduced by 40°C after using the smart phone to control the cooling fan. It takes roughly 20 minutes at this time. Figure 8 depicts the temperature change of the solar panel. In general, the temperature of the solar panel is reduced.

The temperature of the solar panel was kept above 57°C, while the ambient temperature was kept at 28°C. The temperature reduced by 40°C after using the smart phone to control the cooling fan. It takes roughly 20 minutes at this time. Figure 8 depicts the temperature change of the solar panel. In general, the temperature of the solar panel is reduced.

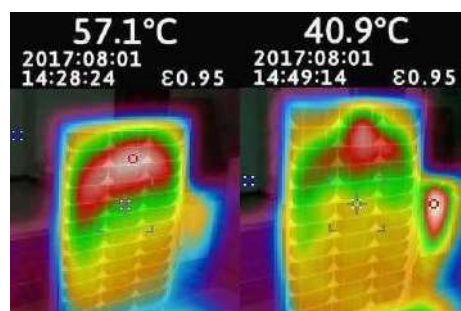


Fig. 8. Temperature change after cooling fan operation

CONCLUSION

We devised and built a solar panel temperature control system in this article. It is now feasible to adjust the temperature of the solar panel from anywhere and at any time using an Arduino and a Wi-Fi shield. The efficiency of electricity generation can be boosted by reducing the temperature of the solar panel. Furthermore, system control via smart phone is quite simple. In the near future, we must conduct further research into technologies that can considerably reduce the temperature of solar panels while also lowering costs. By implementing this technique to general household and large-scale photovoltaic systems, it is possible to maximise power generating efficiency.

REFERENCES

- 1.K. J. Park (2017), “Arduino turns your ideas into reality,” pp.198 of Edu-I
- O. Vermesan et al., “Internet of Things Strategic Research Roadmap,” Internet of Things-Global Technological and Societal Trends, vol. 1, 2011, pp.9-52.
- 2.S. Dubey, J. N. Sarvaiya, and B. Seshadri, “Temperature Dependent Photovoltaic(PV) Efficiency and Its Impact on PV Production in the World – A Review,” Energy Procedia, vol. 33, 2013,pp.311–321.
- 3.Adriansyah A, Dani, A.w., “Design of a Small Smart Home System Based on Arduino,” in Electrical Power, Electronics, Communications, Controls, and Informatics Seminar (EECCIS), 2014, pp.121-125, 27-28 August 2014.

A STUDY ON NETWORK SECURITY AND CRYPTOGRAPHY

Ms.S.Anitha¹ , Ms.R.Padmalatha²

*Assistant Professor, Department of Computer Applications
Marudhar Kesari Jain College for Women,Tamilnadu,India*

ABSTRACT

Network Security is a concept of securing data through wireless transmission with the help is used to ensure the contents of a message which are confidentiality means of cryptography. Data Security is the main aspects of secure data transmission over unreliable network. Network Security involves the authorization of access to data in a network, which is controlled by the network administrator. Network Security is used in various computer network sectors such as private and public. Networks used in the organizations, enterprises, institutions, etc..are in the form of private and public. The task of network security is not only ensuring the security of end systems but also to the entire network. Network Security is used in various applications like Government agencies, Organization, Enterprises, Bank, and Business etc. Cryptography nobody can understand the received message expect the one who has the decipher key, this is done when the sender includes a cryptographic operation called hash function on the original message. A hash function is a mathematical representation of the information, when any information arrives to receiver, the receiver calculates the value of this hash function. Security of data is done by a technique called cryptography. So one can say that Cryptography is an emerging technology, which is important for network security. In olden day's cryptography was used to keep the military information, diplomatic correspondence secure and in protecting the national security but the usage was limited .Now-a-days, the range of cryptography applications have been expanded a lot in this modern area after the development of communication. Cryptography is essentially required to ensure that data's are protected against penetrations and to prevent and it is also a powerful means in securing e-commerce.

INTRODUCTION

Network Security protects our network and data from breaches, intrusions and other threats. This is a vast and overarching term that describes hardware and software resolutions as well as processes or rules and configurations relating to network use, accessibility and overall threat protection. Network Security involves access control, virus and antivirus software, application security ,network analytics, types of network-related security [endpoint, web, wireless, firewalls, VPN encryption and many more. Network Security is the most vital component in information

security because it is responsible for securing all the information passed through networked computer. Network Security refers to hardware and software functions, characteristics, features, operational procedures, accountability, measures, access control, administrative and management policy required to provide an acceptable level of protection for hardware and software in a network. Internet has become more widespread, if an unauthorized person is able to get access to this network, he can not only spy on us but he can easily mess up our lives. Network Security and Cryptography is a concept of protecting the network and data transmission over a wireless network. A Network Security system typically relies on layers of protection and consists of multiple components including networking, monitoring and security software in addition to hardware's and appliances. All components work together to increase the overall security of the computer network. Security of data can be done by a technique called Cryptography.

Cryptography is the science of writing in secret code. Modern Cryptography exists at the intersection of the disciplines of mathematics, computer science, and electrical engineering. An application of cryptography includes ATM cards, computer password, and electronic commerce. The development of the World Wide Web resulted in broad use of cryptography for e-commerce and business applications. Cryptography is closely related to disciplines of cryptology and cryptanalysis. Techniques used for decrypting a message without any knowledge of the encryption details fall into the area of cryptanalysis. Cryptanalysis is what the layperson calls "breaking the code". The areas of cryptography and cryptanalysis together are called cryptology. Cryptography means "Hidden Secrets" is concerned with encryption.

Encryption is the process of converting ordinary information (called plaintext) into unintelligible text (called cipher text). Decryption is the reverse process of encryption, moving from the unintelligible cipher text back to plain text. Cryptosystem is the ordered list of elements of finite possible plaintext, cipher text, keys and the encryption and decryption algorithms which correspond to each key.



The various aspects in information security such as data confidentiality, data integrity, authentication, and non-repudiation are central to modern cryptography. The testing issue is the way

to successfully share scrambled information. Encode message with unequivocally secure key which is known just by sending and beneficiary end is a note worthy perspective to get strong security in sensor organize. The safe trade of key amongst sender and recipient is a lot of trouble some errand in as set imperative sensor arrange. information ought to be scrambled first by clients before it is outsourced to a remote distributed storage benefit and both information security and information get to security ought to be ensured to such an extent that distributed storage specialist organizations have no capacities to unscramble the information, and when the client needs to pursuit a few sections of the entire information, the distributed storage frame work will give the availability without recognizing what the segment of the encoded information to the client is about.

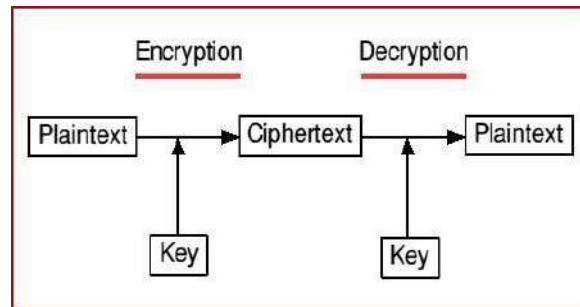
METHODOLOGY

Internet security is a tree branch of computer security specifically related to the Internet, often involving browser security but also network security on a more general level as it applies to other applications or operating systems on a whole. Its objective is to establish rules and measures to use against attacks over the Internet. The Internet represents an insecure channel for exchanging information leading to a high risk of intrusion or fraud, such as phishing. Different methods have been used to protect the transfer of data, including encryption. Network security involves the authorization of access to data in a network, which is controlled by the network administrator. Users choose or a reassigned an ID and password or other authenticating information that allows them access to information and programs within their authority.

CRYPTO GRAPHIC PRINCIPLES

Redundancy: All the encrypted message contain some redundancy, there is none ed of understanding the message by information.

Freshness: Time stamp is used in every message. For instance the time stamp is of 10sec for every message. Thereceiverkeepsthemessagearound10secto receive the message and filter the output within that10sec.The message exceeds the time stamp it is throw out.



CRYPTOGRAPHY GOALS

By using cryptography many goals can be achieved, These goals can be either all achieved at the same time in one application or only one of them, These goals are:

Confidentiality: It is the most important goal, that ensures that nobody can understand the received message except the one who has the decipher key.

Authentication: It is the process of proving the identity, that assures the communicating entity is the one that it claimed to be, This means that the user or the system can prove their own identities to other parties who don't have personal knowledge of their identities.

Data Integrity: It ensures that the received message has not been altered in any way from its original form, This can be achieved by using hashing at both sides the sender and the recipient in order to create a unique message digest and compare it with the one that received.

Non-Repudiation: It is a mechanism used to prove that the sender really sent this message, and the message was received by the specified party, so the recipient cannot claim that the message was not sent.

Access Control: It is the process of preventing an unauthorized use of resources. This goal controls who can have access to the resources, If one can access, under which restrictions and conditions the access can be occurred, and what is the permission level of a given access.

CRYPTO SYSTEM TYPES

Asymmetric cryptosystems

It uses two different keys to send and receive the messages. It uses public key for encryption and another key is used for decryption. Two users A and B need to communicate, A uses public key of B to encrypt the message. B uses private key to decipher the text. It is also called as public key cryptosystems. Diffie-Hellman key exchange generates both public and private key.

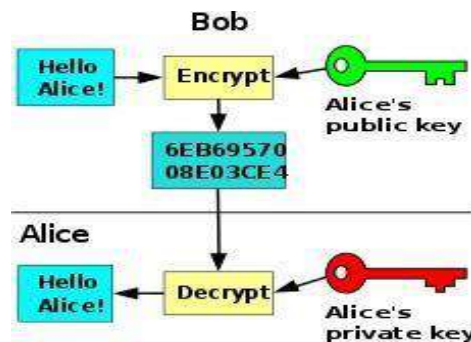


Fig. Asymmetric cryptosystems

Symmetric crypto systems

In Symmetric cryptosystems both the enciphering and deciphering keys are identical or sometimes bother related to each other. Both the key should be kept more secure otherwise in future secure communication will not be possible. Keys should be more secure and it should be exchanged in a secure channel between two users. Data Encryption Standard(DES) is example of Symmetric cryptosystems.

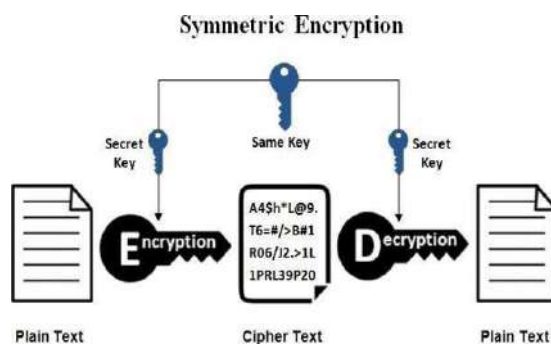
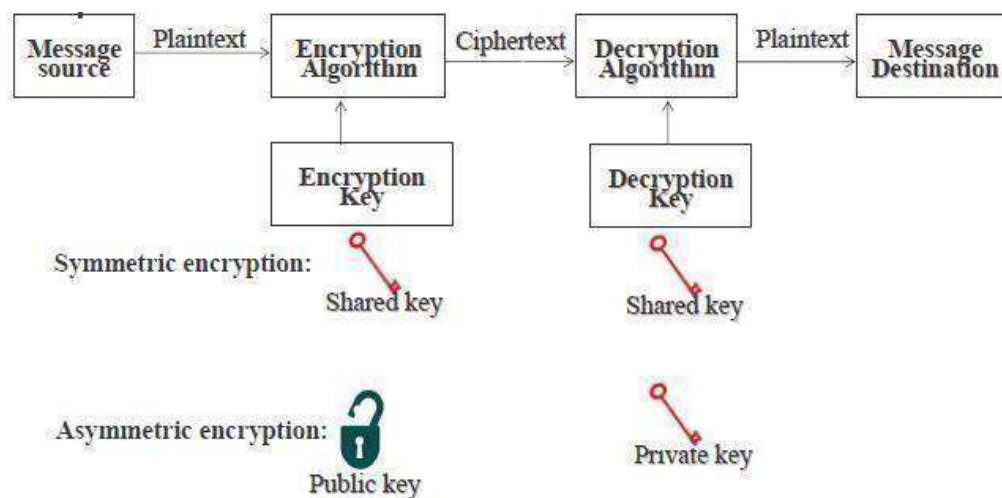


Fig. Symmetric cryptosystems

CRYPTOGRAPHIC MODEL & ALGORITHM

Encryption model

There are two encryption models namely they are as follows: Symmetric encryption and Asymmetric encryption. In Symmetric encryption, Encryption key Decryption key. In Asymmetric encryption, Encryption key Decryption key.



Algorithm

There are of course a wide range of cryptographic algorithms in use. The following are amongst the most well-known:

DES: This is the 'Data Encryption Standard'. This is a cipher that operates on 64-bit blocks of data, using a 56-bit key. It is a 'private key' system.

RSA: RSA is a public-key system designed by Rivest, Shamir, and Adleman.

HASH: A 'hash algorithm' is used for computing a condensed representation of a fixed length message/file. This is sometimes known as a 'message digest', or a 'fingerprint'.

MD5: MD5 is a 128-bit message digest function. It was developed by Ron Rivest.

AES: This is the Advanced Encryption Standard (using the Rijndael block cipher) approved by NIST.

SHA-1: SHA-1 is a hashing algorithm similar in structure to MD5, but producing a digest of 160 bits (20 bytes). Because of the large digest size, it is less likely that two different messages will have the same SHA-1 message digest. For this reason SHA-1 is recommended in preference to MD5.

HMAC: HMAC is a hashing method that uses a key in conjunction with an algorithm such as MD5 or SHA-1. Thus one can refer to HMAC-MD5 and HMAC-SHA1.

CONCLUSION:

With the explosive growth in the Internet, network and data security have become an inevitable concern for any organization whose internal private network is connected to the internet. The security for the data has become highly important. User's data privacy is a central question

over cloud. With more mathematical tools, cryptographic schemes are getting moreover satile and often involve multiple keys for a single application. The various schemes which are used in cryptography for Network security purpose. Encrypt message with strongly secure key which is known only by sending and recipient end, is a significant aspect to acquire robust security in cloud. The secure exchange of key between sender and receiver is an important task. The key management helps to maintain confidentiality of secret information from unauthorized users. It can also check the integrity of the exchanged message to verify the authenticity. Network security covers the use of cryptographic algorithms in network protocols and network applications. This paper briefly introduces the concept of computer security, focuses on the threats of computer network security. In the future, work can be done on key distribution and management as well as optimal cryptography algorithm for data security over clouds.

REFERENCE:

1. DaveDittrich, Network monitoring /Intrusion Detection Systems(IDS), University of Washington.
2. Algorithms:<http://www.cryptographyworld.com/algo.htm>
3. Data_Communication_and_Networking_by_Behrouz.A.Forouzan_4th.edition
4. Bellare, Mihir; Canetti, Ran ; Krawczyk, Hugo,"Hash Functions for Message Authentication",1996.
5. WilliamStallings,"CryptographyandNetworkSecurityPrincipleandPractice",FifthEdition,2011.
6. Diffie,W.,Hellman,M.E.:New directions in cryptography.IEEE Transactions On Information Theory 22,644–654(1976).
- 7.Gross,T.,M'odersheim,S.:Verticalprotocolcomposition.In:24thIEEEComputerSecurity FoundationsWorkshop(CSF2011). Publication197-Announcing the Advanced Encryption Standard (AES) . Federal Information Processing Standards,26Nov.2001.
8. Ralston, Anthony, Edwin D.Reilly, and David Hemmendinger .Encyclopedia of Computer Science.Fourthed.London, England: NaturePublishing Group,2000
- 9.Maurer,U.Secret key agreement by public discussion from common information. IEEE Transactions on Information Theory 39(3) , 733–742(1993)
- 10.ShyamNandanKumar,"TechniqueforSecurityofMultimediausingNeuralNetwork,"Paperid-IJRETM-2014-02-05-020,IJRETM,Vol:02,Issue:05,pp.1-7.Sep-2014
- Preneel.B (2010, September). Cryptography for network security: failures, successes and challenges.

11. International Conference on Mathematical Methods, Models, and Architectures for Computer Network Security (pp.36-54). Springer, Berlin, Heidelberg.
12. Panda, M. (2014). Security in wireless sensor networks using cryptographic techniques. American Journal of Engineering Research (AJER), 3(01), 50-56.
13. Dhamdhere Shubhangi .T., & Gumaste, S. V. Security in Wireless Sensor Network Using Cryptographic Techniques.
14. Simmonds, A; Sandilands, P; van Ekert, L (2004). "An Ontology for Network Security Attacks". Lecture Notes in Computer Science. Lecture Notes in Computer Science 3285:317-323.

DETECTION OF MALARIA PARASITES USING DEEP LEARNING TECHNIQUES*G.Sasirekha¹, A.Poornima²**Assistant Professor, Department of Computer Science
Marudhar Kesari Jain College for Women,Tamilnadu,India*

ABSTRACT

In this section an exemplary of malaria verdict was developed which is precise, fast and inexpensive, it is developed using the tainted tinny blood smear imageries. Concentration choices of plasmodium vermin and erythrocytes are used to create this methods. The pertinent topographies are extracted from the imageries of collected, re-handled infected and non-infected erythrocytes. Artificial Neural network (ANN) is the classifier which is used to evaluate and to create a database of the red blood cell samples based on the presentation of these topographies which are collected with the set of options with a planned reinforced strength. The selections are given with success and the consequences have shown using the protozoal infection discovery which are described as results.

Keywords: Artificial neural network and Convolutional neural networks.

INTRODUCTION

Peripheral blood pest of the species plasmodium is the reason of the contagious communicable disease called malaria. World Health organization has said that rendering to them the approximation was done that 262 million cases of malaria were there globally in the year 2000 also it leads to 839 deaths. But WHO also assessed that the number of malaria cases were reduced to 214 million and the total number of deaths decreased to 438 in the year 2015. From which majority of deaths were children from Sub-Saharan Africa. The reason behind this is the pitiable socio-economic circumstances and mosquitoes with good ecological conditions which makes the disease anticipation assets and health care conditions difficult. The old standard technique to diagnose the disease is manual microscopy among numerous techniques used. The above indicative method is leading to late verdict is the most protracted process and it is disposed to human error which is irrespective of the steps required in the manual impost leading to erroneous indicative even in knowledgeable hands. This approach which is manual of diagnosis is time overwhelming and may lead to discrepancy.

The request of training and knowledgeable technicians or pathologists is important. This will decrease the time occupied for broadcast the disease and this method should be digitized. The reliability is improved in the analysis.

The microscopic colour imageries is used to detect the malaria parasites and this detection inquiries into the use and request of the digital image dispensation. To describe the parasite recognition based on the strength and quality features an effectual method is proposed. Semi-automated analysis is the important purposes for the parasite discovery. Here this part is used to propose the other ordering and gratitude problems which are applied here. The associated two different types of grouping system increase its performance rate with the information given. Hence the features are extracted from the CNN and construct the classifier systems for the given purpose

Biological procedure of the property design among the neurons looks like the group of the animal cortical part there by galvanized and these are called as the convolutional networks. Receptive turf is the turf of regard which is the animal tissue neurons and separate answer stimuli exclusively in the restricted region of the field of above regard. This above receptive fields with the numerous neurons portion overlaps as like they cover with the comprehensive turf of regard. Comparing to dissimilar image arrangement algorithm this CNN has very little pre-processing algorithms. Antique procedures were hand-engineered and this gives the suggestions that the network studies the filters with this. The main advantage is that the independence from the preceding info and the human exertion in the feature style. Hence the need of application in image, video gratitude, recommender schemes and linguistic process is required.

2. PROPOSED SYSTEM

To detect the malarial parasite by using the microscopic colour imageries the machine learning methods are used in this proposed system. Convolution neural network is used to excerpt the structures and the parasite discovery is made in a well-organized technique and it is proposed. This parasite recognition is the important function of this analysis.

The general convolutional neural network topographies are used to appliance a novel method for the infected blood cell organization. The CNN can mechanically excerpt the options, study and categorize them which is square dimension. Comparing to the alternative image organization algorithms the CNN uses a equally very little pre-processing algorithms. So the network studies the filters based on the antique algorithms which is hand engineered. Image

organization has a good advantage of this self-governing from the previous knowledge and the human effort taken in the feature project.

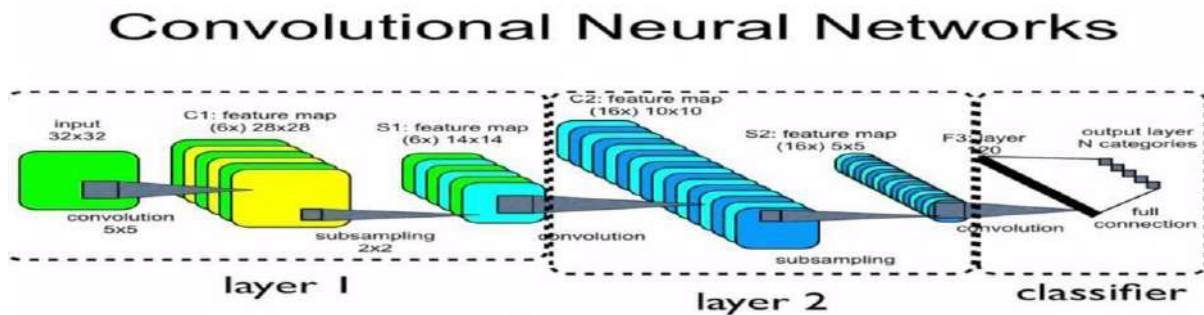


Fig 1. Convolutional Neural Networks

Computers unable to understand imageries in the way but are able to achieve calculations on numbers. Computers should comprehend the imageries only by changing it into numbers. Hence in computers the images are distorted into gray scale which is the variety of grey shades from white to black, the pixel is allocated by the value based on the darkness of the image. The computer computes the array by putting each numbers into an array

Convolutional neural network is used to build this model and this networks are proposed by Yann LeCun in 1988 which is the special architecture of the artificial neural networks. Visual cortex has features which is used by CNN (Convolutional neural network). It is possible to start implementing the neural network after done with pre-processing. In this we are having the complication layers with the max-pooling kernels that is 2×2 . By using the technique the maximum pixel value of a grid dimensions can be reduced. Also this helps to decrease the over suitable and the representations become more general and after this fully associated layers are added. The fully associated layers input should be two dimensional and the convolution layer output is four dimensional, for this the flattening layer is required among them. The fully connected layer which is connected is the soft max layer.

Model preparation is done with time after the model erection. An artificial convolutional neural network has been built with the recognized images. The dataset is divided into the train and the test dataset. Using this training dataset it is possible to build the classic and train the model at the end. At this phase the set of data is tested and loaded. The correctness of this data is verified and the dataset has not been seen by the classical. In the real world this classic is saved and used. Classic evaluation is the name of this phase and this means the model can be used to evaluate the new data.

The Python is the most influential encoding language and it is easy too. It is an effective method of an object oriented programming and has the methodical relaxed process of high-level data structures. Among most of the platforms this python makes itself a faultless language for quick application growth and scripting, in spite of python's sophisticated arrangement and dynamic typing with its inferred character.

In Python there are certain wide-ranging standard libraries which are freely available in the basis or in the binary format of all countless platforms from the internet websites of python, like www.python.org, and it can extent freely. This site also has the python classic, agendas and gears and large documentation it holds the distributions of the above stated methods of third party points.

Python is an allowance language for customizable requests which is additionally satisfactory. In python data types are applied in C or C++ and by the translator, absolutely new functions are extended.

Now this language is certainly prolonged with new functions and the data types are implemented in C or C++ that is called as python interpreter. This python is additionally satisfactory as an extension of language for the customizable applications.

The high-level, construed, collaborating, and object oriented scripting language is the python programming language. This language is an extremely legible language and it is intended in such a way. All the other languages uses punctuations but python is the language which often uses English keywords and all the other languages customs punctuations, also python has no other syntactical constructions than other languages.

4. RESULTS AND DISCUSSION

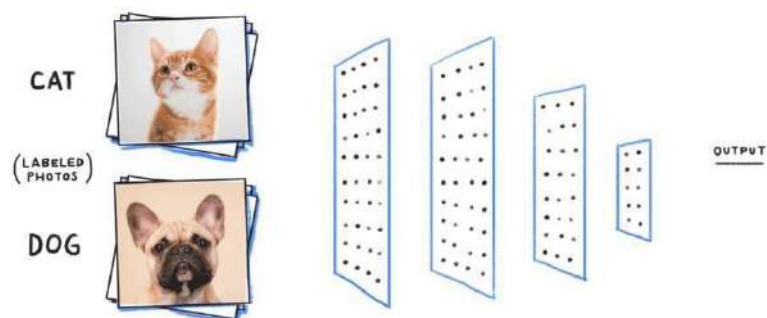


Fig2. Image Acquisition

The data from the online source are taken in this module represented in (Figure 2). And the images are resized for the future use. The resizing of images, rescaling of images and the image which has geometric image alteration, this rest on an image exclamation algorithm and it modifies the image size. The whole size of the image data's is adjusted and the image scaling process can upsurge or decline the determination of the objective image size.

Computers cannot understand the images like the way how we do, but they are able to perform the calculations on numbers. To understand images, the computers should change the imageries to numbers, these imageries will be transformed to gray scale the variety of the gray glooms are from white to black, Based on the darkness the value of each pixel is assigned in the computer. All this numbers are taken in the form of an array and this computer compute the values in that array. Then the array is feed in to the next step.

Data Preparation and Model construction

Train and test are the two data sets which are first divided by the people often. The later they put the test set away and they choose the train data that is X% randomly with the real train set and the outstanding (100-X)% and this is the long-established set, Here we say that X is the fixed number which is 80%, After this the model is trained and authenticated on the dissimilar sets iteratively. Hence we survey and use the same method to prepare the data for the analysis and training phase.

Convolutional neural network is the method which is used for building our model. In 1988 Yann LeCun proposed the superior architecture of artificial neural networks which are the convolutional neural networks(CNN). Visual cortex features are used by the CNN. After this we have completed the pre-processing and now we can start applying the neural network. There are 3 convolution layers which are 2*2 be very successful pooling we used.

Max-pooling

By considering the maximum pixel price of a network a technique is used to decrease the measurements of an image. This makes the classical more general and also helps to reduce the over appropriate. Then the fully 2 connected layers are established and added. The input which is given based on fully connected layers must be two dimensional, then the four dimensional complication layer is the output, and now it is needed to have a destruction layer between them. Here the softmax layer is the fully associated layer eventually.

After model construction it is now time for the model training. We were able to build an artificial convolutional neural network that can recognize images. Split the dataset into train and test dataset. Eventually we will build and train the model using the training dataset.

Model training

Now the next process is model training after classical construction. The network that can identify the imageries are built using the artificial convolutional neural network. The dataset is divided into train and test the data's. The model is built and trained by using this training dataset at the end.

Model testing and evaluation

Model analysis is likely to carry out by training the model. A test data set is loaded during this period. The true accuracy is verified here and the results are taken by the data set and

has never been seen by the model. In the real world the model which is saved can be used eventually. This is called model evaluation. The new data can be evaluated by the use of this given model.

5. CONCLUSION

To distinguish the malaria parasite from a smear blood cell images this is the approach used to positively build the classical which is intelligent and uses the computer vision. The peripheral blood parasite of the protoctist genus is the causes of serious infectious disease called malaria. Here the results are problematic to replicate and it is found that it is time consuming because it was occasionally tested by the conservative microscopy which is presently the gold common for the diagnosis of malaria. The assessment procedure is automaton highly important and it is used to postures serious total health problem. This is the progressive method recently and this scheme comes under the deep learning algorithm. The spitting image grouping is more suitable for the image handling particularly using the CNN algorithm. Hence this shows the 90% correctness by investigational result of the scheme which is established and concluded.

REFERENCES

- [1] W. World Health Organization, “Who Report 2015,” 2015.
- [2] N. R. Shet and N. Sampathila, “An Image Processing Approach for Screening of Malaria,” Canar. Eng. Coll. Mangalore, pp. 395–399, 2015.
- [3] D. A. Ghate and P. C. Jadhav, “Automatic Detection of Malaria Parasite from Blood Images,” Int. J. Adv. Comput. Technol., vol. 1, no. 3, pp. 66–71, 2012.
- [4] M. S. Suryawanshi and P. V. V Dixit, “Comparative Study of Malaria Parasite Detection using Euclidean Distance Classifier & SVM,” vol. 2, no. 11, pp. 2994–2997, 2013.
- [5] A. Anand, V. K. Chhaniwal, N. R. Patel, and B. Javidi, “Automatic identification of malaria-infected RBC with digital holographic microscopy using correlation algorithms,” IEEE Photonics J., vol. 4, no. 5, pp. 1456–1464, 2012.
- [6] V. Špringl, “Automatic Malaria Diagnosis through Microscopy Imaging,” Czech Tech. Univ. Prague Fac. Electr. Engineering, 2009.
- [7] CDC, “CDC - Malaria - About Malaria - Biology - Malaria Parasites,” USA Government, 2012. [Online]. Available: <http://www.cdc.gov/malaria/about/biology/parasites.html>.
- [8] N. Bhargava and R. Bhargava, “Threshold and binarization for document image analysis using otsu ’ s Algorithm,” Int. J. Comput. Trends Technol., vol. 17, no. 5, pp. 272–275, 2014.
- [9] Aimi Salihah Abdul-Nasir, Mohd Yusoff Mashor, and Zeehaida Mohamed, “Colour Image Segmentation Approach for Detection of Malaria Parasites Using Various Colour Models and k –Means Clustering,” WSEAS Trans. Biol. Biomed., vol. 10, no. 1, pp. 41–55, 2013.
- [10] W. Khan, “Image Segmentation Techniques: A Survey,” J. Image Graph., vol. 2, no. 1, pp. 6–9, 2013.
- [11] A. Verm, M. T. Scholar, C. Lal, and S. Kumar, “Image segmentation: Review paper,” Int. J. Educ. Sci. Res. Rev., vol. 3, no. 2, 2016.

THE ROLE OF BIG DATA IN ECOMMERCE

*A. Karthick Praveen Kumar¹,
Sr. Data Architect, AWS, USA*

*B. Sakthimala²
Assistant Professor, Department of BBA
Marudhar Kesari Jain College for Women, Tamilnadu, India*

ABSTRACT

Due to the information revolution, one of the biggest challenges of e-commerce is the huge data that needs to be processed and analyzed to gain benefits from it. Big Data Analytics aims to improve the decision-making process by analyzing and understanding big data, e.g., messages, social media posts, etc. Furthermore, there has been an increasing emphasis on big data analytics in e-commerce in recent years. However, it remains poorly-explored as a concept, which obstructs its theoretical and practical development. The paper also explains data transformations in business

Introduction

The goal of this paper is to discuss how data can be used to transform business, especially an ecommerce business. We focus on how to harness the power of data to transform a business, how to improve services provided, improve customer experience, reduce cost, improve target marketing, improve personalization, forecast and predict sales, identify buying patterns, detect fraudulent transactions, and ultimately increasing profit.

Many companies use data to understand their customer behaviour. Understanding customer behaviour will give them the power to know what the customer wants, and predict their needs. Predicting customer needs will give the power for companies to focus on building and delivering products that customer needs. Predicting customer behaviour used to be a very complicated task in the past, the limitations were mostly due to the limits imposed by technology and lack of data. With the growth in technologies like cloud, mobile, internet of things (IoT), companies are able to harness the power data much easily than before. This enables companies to predict customer behaviour.

The Role of Big Data

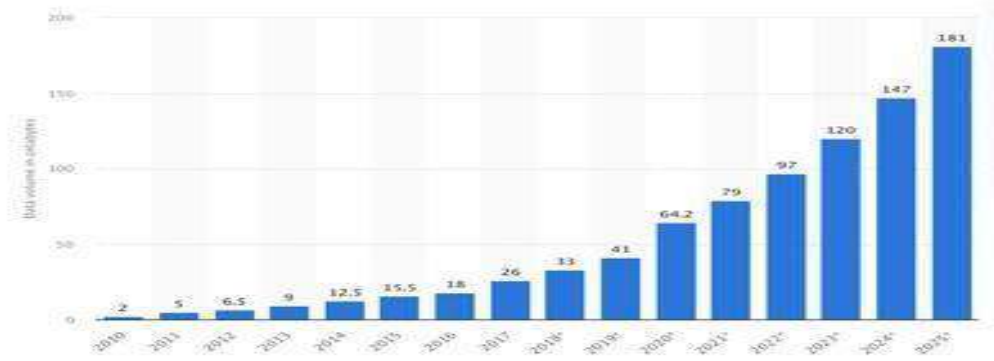
Data are individual facts, statistics or items of information, that are collected through observation. In a more technical sense, data are a set of values of qualitative or quantitative variables about one or more person or objects, while a datum is a single value of single variable. But how does data become more meaningful. For example, when I say “145” does this make any sense? It can be 145 apples or 145th street or something else. But when I add more

contexts to this data, it becomes more meaningful. For example, when we say “145th Street”, this piece of data becomes more meaningful. Now everyone can understand that the data represented here is an address (or part of an address). This context, which provides more



meaning to data, is called as metadata. In simple terms, metadata is data about data. Now that we know what data and metadata is, let's understand what big data means. Big data refers to data that has the following 5 V's. The following are the 5 V's which makes up big data: volume, velocity, variety, veracity and value.

Volume refers to the amount of data that exists. Volume is the base of big data; it indicates the amount of data collected by the business. Velocity refers to the rate which data is generated and how quickly it's moving. Velocity is important for business as it helps in making quicker decision. Variety refers to the types of data collected, with the technological advancement, companies can now store not just structured but also unstructured data, and semi structured data. Veracity refers to the quality and accuracy of the collected data. As data is collected, stored and transformed, there is a high chance that the quality of data may be affected, and become messy. Veracity is important in making business decision, because “no data is better than bad data”. The final V refers to the value. This is the value the data can provide to the business. With the technological advancement, the volume of data created, captured and consumed worldwide has significantly increased from 2010 to 2025 (in zettabytes) as shown in the below chart.



Cloud Computing in Big Data

Having data stored in hard drives will not help drive business outcomes, unless we process it to extract valuable information. In the past, companies need to build data centres (buildings), purchase computer hardware (servers, networking devices, etc), get utility connection (electricity and internet), hire highly skilled people to help connect these together and make them working. This would take several months and sometimes years to complete. Once the data centre is built, they need to purchase, install and configure software. Until the data centre is ready with required software's, companies were not able to run their business effectively. But with cloud computing, companies do not have to wait, they can spin up computing or storage resources in minutes. Also cloud computing companies provide virtually unlimited compute and storage options, which helps business to store and process their big data easily. With the basics covered, it's time to get into the crux of the paper to discuss about how cloud computing and big data has influenced ecommerce. As discussed earlier, it's important for companies to understand their customer to provide better service to them. While it's highly impossible to discuss every single use case of how big data and cloud computing has transformed e-commerce business, we will discuss some important and most common use cases in this paper. This will lay a strong foundation to understand the importance of data and its impact on business.

Personalization

As a customer, we feel better when businesses know what we need and when we need. As a customer this helps us to buy the right product rather than going over a million products online. For example, as a customer I would like to purchase hot chocolate during winter, and I like to purchase brand A hot chocolate. Without personalization, I will have to browse through hundreds of products and browse through hundreds of pages to find and purchase my product. With personalization, the ecommerce company can directly display my preferred hot chocolate brand in the home page when the winter starts. This is achieved by analyzing my purchase

history. Similarly, this personalization can be done for hundreds of millions of customers in a very short time. This is possible only by analyzing huge volume of previous purchase.

Predict Trends

Predicting trends helps business to change the way they do business, or adopt to the latest changes in the society. For example, say people are actively discussing about an issue in their social media, by analyzing social media data, companies can understand the sentiment of a particular demographics, and based on that they can change the way they recommend products to the customer or target advertisements, etc.

Better Forecasting

By combining different types of data, business can improve their forecasting. For example, by combining customer previous purchase data with weather data, we can understand that during the winter months people tend to play online games more as opposed to outdoor games. This will help online gaming companies to better plan for sudden increase in online gamers.

Improve customer service

When a customer calls a customer care centre with a problem, it would be better for the agent to know who is calling and their details. Thus, by just collecting very basic information the agent will be able to authenticate the customer. This makes it easy for the customer.

Target marketing

Companies that can better understand their customer will be able to deliver relevant advertisements to them. For example, when selling a mountain bike it's better to understand the type of customers who normally use mountain bike. Mountain bikes are generally used by people who are more active, who live in a particular geographical location and are of certain age group. By understanding this, companies can target and deliver to the right set of people and not to all people. This helps in saving marketing cost.

Detect and reduce online fraud

By understanding customer purchase pattern, e-commerce companies can reduce fraudulent transaction. This will benefit both the customer and the business. For example, based on the previous customer purchases, we can understand that the customer normally makes transaction between certain amount, say between \$10 to \$500 in a month. But if we suddenly see an increase in the purchase amount for the customer, say \$6472 this should raise an alarm. Someone else might have made this purchase. In this case, the ecommerce company can hold the transaction and call or email the customer to confirm the transaction before proceeding. This way both the customer and the business can improve their security posture. The customer

can change their login credentials to prevent unauthorized use and the company can save in shipping cost.

Better pricing

Using data to manage pricing is a very effective practice in e-commerce. For example, companies can analyze what products other business are selling and at what price. Based on the outcome, companies can now change their pricing structure. This will help companies to provide products at a very competitive price and thus providing more value to customer.

Conclusion

By using big data and the power of cloud computing, companies can now transform the way they operate. While this paper focused more the impact of big data and cloud computing on ecommerce, the same can be applied for other business too.

References:

1. <https://www.talend.com/resources/big-data-ecommerce/>
2. <https://ijarcce.com/upload/2017/december-17/IJARCCE%2035.pdf>
3. <https://www.bigcommerce.com/blog/ecommerce-big-data/>
4. https://www.researchgate.net/publication/347443987_The_Impact_of_Big_Data_Analytics_on_the_E-commerce_Industry
5. https://www.researchgate.net/publication/298739144_Big_data_analytics_in_E-commerce_a_systematic_review_and_agenda_for_future_research
6. <https://ijarcce.com/wp-content/uploads/2016/11/IJARCCE-ICRITCSA-28.pd>

**A STUDY ON SCOPE OF HYDERABAD AS THE IT
STARTUP HUB OF INDIA***T. Tejaswi¹**Research Scholar, Department of Management,**Dr. U. Devi Prasad²**Professor, Department of Management,**GITAM (Deemed to be University), Hyderabad Campus, Telangana*

ABSTRACT

In India, information technology is divided into two categories: Outsourcing of IT services and business processes. The sector's contribution to India's GDP has increased from 1.2 percent in 1998 to 7.7 percent in 2017. According to NASSCOM's IT industry report, the sector generated revenue of US\$180 billion in 2019, with export income of US\$99 billion and national revenue of US\$48 billion, a 13 percent increase. India will have 4.36 million IT workers by 2020. Accenture, Amazon (company), Deloitte, Tata Consultancy Services, Microsoft, HCL Technologies, Oracle Corporation, Google, Qualcomm, Dell, and Cognizant are just a few notable IT businesses. When you think of India's startup scene, you probably think of Bangalore or Gurgaon, but Hyderabad has recently emerged as a prominent destination for Indian entrepreneurs. Hyderabad's IT exports are expected to reach 128,807 crores (US\$15 billion) by 2020, and there are 1500 IT and ITES enterprises in the city, employing 582,126 people.

Keywords: Hyderabad, Information Technology, hi-tech city

Introduction

In recent years, The Indian start-up ecosystem has matured and taken off. The growth of start-ups has been fuelled by capital availability, consolidation efforts by several firms, expanding technology environment, and increasing demand within the domestic market. The numbers on start-ups speak for themselves: it is estimated that there will be 11,500 businesses by 2020, up from 3,100 in 2014. This is a groundbreaking trend! Moreover, it will alter the current state of affairs in India's markets. A group of quirky, one-room "work stations" are forging a new character for Hyderabad, a start-up hub, among the towering offices of international firms clogging the lanes of the hi-tech city. These start-ups are mushrooming across the city's IT centres, giving firms the jitters even in Delhi, Chennai, and Pune. They are managed by IT experts pulsing with out-of-the-box thoughts.

Hyderabad

This is the capital city of Telangana state in India. This city is notable for its HITEC City, Cyberabad, or Nawab's City, and to a lay Indian, Charminar is, of course, and by most yardsticks, it is true. It is the second biggest IT exporter of India, a critical global IT hub, and its largest bioinformatics centre. After the formation of Telangana state in 2014, Hyderabad entered the fray. However, according to recent research from start-up blink, cooperating with entrepreneurs, investors, government, and academics has already built the groundwork for a healthy start-up environment. The 'Cyber City' is rapidly improving its innovation game and has risen 115 places to become the 75th best city for start-ups in 2019. When it comes to IT and start-ups, it does not have the same clout as Bangalore. The reasons are mainly obvious: Bangalore has a much larger pool of IT companies, IT people, and R&D houses than other Indian cities, giving it a competitive advantage. This opinion piece will attempt to demonstrate why Hyderabad has the potential to become India's future start-up capital, as well as the numerous reasons that support that assertion.

Governance with vision and foresight

Over the last 1-1.5 years, Hyderabad has seen many changes. The most important of these is political stability, which led to the formation of Telangana, a new state with Hyderabad as its capital. The city had suffered due to the uncertainty, with some businesses deferring investments and others relocating to other places. Everything has now come to an end. All of this is now in the background.

Telangana today has a government that is both stable and forward-thinking. Shri K.T. Rama Rao, the state's exceptionally astute and charismatic IT Minister, has revolutionized people's perceptions of the state and the perceptions of external investors. He has the warmth of a statesman and the strategic acumen of a strategist. Srinivas Kollipara, the founder of IT-Hub, was inspired by the minister's forward-thinking concept.

The importance of infrastructure

In contrast to other cities such as Bangalore, Hyderabad boasts one of the best infrastructures. It also has a large pool of IT, BT, and R&D expertise, all of which are essential inputs for the success of any knowledge hub. It also has some of the most affordable commercial real estate in the country, making it a desirable business destination and offshore centre. It also has a progressive bureaucracy and some of the country's most stringent industrial and IT regulations.

Human Capital drives ecosystem

There are thousands of research institutes in Hyderabad, including CCMB, NIN, NARM, ICRISAT, DERL, DMRL, BEL, ECIL, IIT, IIIT, ISB, NALSAR, and others, all of which produce adaptable and talented individuals as well as highly quantitative and wise professionals.

This enables emerging knowledge-intensive industries such as IT, telecommunications, and, more recently, the start-up ecosystem to tap into sufficient home-grown expertise. Out of the 20,000 start-ups in the country, roughly 3000 are located in Hyderabad. It is one of the country's top three buzzing start-up ecosystems, with a slew of incubators, accelerators, and co working spaces that are hives of activity. These forums allow companies to connect with other start-ups, find talent for their team, find mentors or advisors, find angel investors, find venture funds, and various other things, all of which will help the start-up succeed.

T-Hub is India's largest start-up incubator

In terms of innovation, start-ups, and entrepreneurship, T-Hub has become Hyderabad's crown jewel. With 200+ start-ups housed on a beautiful 70k+ sq. ft. space with a seating capacity of over 700 people, numerous compact rooms and meeting rooms, cafeterias, fun zones, mentor garages, and a large event room, it is India's largest start-up incubator. T-Hub is a one-of-a-kind PPP model in which the government contributes to the Capex side while leaving day-to-day operations to an independent management group overseen by an experienced board of directors. It is now a highly active and bustling structure, brimming with energy and exuberance.

T-Hub is more than just a flashy structure. A vibrant community of companies, mentors, co- founders, and support services, as well as an unrivalled number of events, can be found on the campus. It is a genuine, breathing community where start-ups may access a wide range of services at a drastically discounted price, including business incorporation, HR services, financial bookkeeping, and legal services. All of this is available to businesses on campus rather than going out and finding it, which wastes time and resources that the start-up cannot afford to waste. T-Hub start-ups also have access to an impressive range of mentors from within the city and outside.

Connecting Business and Industry

T-Hub now hosts Microsoft, IBM, Intel, Amazon, Google, Samsung, Yes Bank, ICRISAT, Sales Force, Cisco, HPE, Nasscom, TiE, and a variety of other companies. They all provide a wealth of product and technological knowledge, as well as an incredible quantity of

credits to allow start-ups to use many of these resources for little or no cost. This is significant from the standpoint of a start-up, and we at T-Hub understand this and strive to bring all of this and more to their doorstep.

Raising Capital

In less than a year in the incubator, at least two dozen start-ups have raised the next round of their investment (at least \$100K) or secured significant partnerships with major corporations. The vast majority of new business funding in the city has been in the \$100,000 to \$150,000 range (angel round), with only a few in the \$0.5 million to \$1 million range (VC round). This will likely change as many of these companies fine-tunes their services and go-to- market strategies and begins focusing on higher-margin revenue opportunities, and we expect more big-ticket raises.

Partnerships

Collaborations between firms based on campus have also occurred, which is a significant benefit of a giant incubator with hundreds of start-ups. T-Hub has various corporate collaborations and is an incubator for those who want to participate with the start-up community through many structured initiatives. One popular option is the accelerator program, which a few companies have previously implemented.

'T-Hub' Rising Stars

Want me to mention a couple of outstanding start-ups that are attempting to break through the success barrier. Gayam Motor Works (maker of electric bicycles and automobiles), Hug Innovation (maker of multi-utility smart wearable watches), Loop Reality (maker of mixed reality cycling + VR), ATL (one of the highest peer-to-peer lending start-ups), and a slew of others are among the city's start-up saviours.

T-Hub in Future

There are already ideas for a far more significant and much more comprehensive T-Hub Step 2, which would be five times the existing version and will become the world's most giant start-up incubator once it opened in 2018. This will house the most start-ups of any site globally, making it the ideal melting pot for any and everything innovation and entrepreneurship. So by this time, there would probably be a few significant positive outcomes from the present crop of companies, which would begin to get a skyrocketing impact around the time the more excellent playground is prepared to start. Considering all of the above, it is only logical to predict that Hyderabad will produce the most significant quantity of start-ups and prosper as the nation's most extensive start-up environment. As a result, Mumbai has a

good chance of becoming the country's "start-up capital" and a model for all other cities to emulate, making India the world's "start-up capital."

Conclusion

I am confident that Hyderabad, the pearl city, is sitting on a hidden gold mine that is slowly scraped to the surface and glistening. In terms of talent, Hyderabad is still lagging behind Bengaluru. According to real estate industry projections, the IT infrastructure development of this city is progressing rapidly, with total office stock expected to expand by 60% by 2021. Furthermore, Hyderabad's technological development is concentrated in a single district, HITEC City, making city planning relatively simple. In comparison to Bengaluru, the cost of living is also lower. It shows that Hyderabad has evolved into a central IT hub, challenging Bengaluru for the nation's most significant IT destination.

Bibliography

1. Rama Iyer, Senior Vice President, Strategic Alliances & Innovation, T-Hub on why Hyderabad can become the future start-up capital of India-CFO INSIGHTS (https://www.yesbank.in/pdf/special_feature_hyderabad.pdf)
2. India Briefing from Dezanshira & Associates (<https://www.india-briefing.com/news/indias-tech-hubs-bengaluru-vs-hyderabad>)
3. Startups India - An Overview, by ASSOCHAM India and Grant Thornton ([https://www.grantthornton.in/globalassets/1.-member-firms/india/assets/pdfs - /grant_thornton_startups_report.pdf](https://www.grantthornton.in/globalassets/1.-member-firms/india/assets/pdfs-/grant_thornton_startups_report.pdf))
4. https://en.wikipedia.org/wiki/Information_technology_in_India#Hyderabad
5. <https://m.economictimes.com/news/company/corporate-trends/hyderabad-fast-turning-into-a-hub-for-startup-companies/articleshow>
6. <https://www.forbes.com/sites/sindhujabalaji/2017/07/24/hyderabad-started-slowly-but-the-city-is-now-set-to-become-indias-next-startup-hub>
7. <https://timesofindia.indiatimes.com/city/hyderabad/hyderabad-is-75th-best-global-startup-hub/articleshow>

SOCIAL DISTANCING DETECTION AND MONITORING THROUGH DENSITYDETECTION

*Ms.BHARATHI¹, Asst.Professor, Department of Master of Computer Applications,
Shannuga Industries Arts and Science College, Tiruvannamalai.*

*Dr.G.ANANDHARAJ², Head & Asst.Professor, Department of Computer Science,
Adhiparasakthi College of Arts and Science, Kalavai.*

*R.ANGELIN PREETHI³, Asst.Professor, Department of Computer Science,
Kamban College of Arts and Science for Women, Tiruvannamalai.*

Abstract:

Maintaining social distance in public areas is critical for reducing COVID-19 infection and viral dissemination. Many countries around the world have been forced to close their workplaces, schools, and public venues as a result of COVID-19. This has encouraged policymakers, venue managers, and local governments to look into viable technology-based mitigation measures for securely exiting the lockdown and restoring cities and public places. This paper introduces Crowd Tracing, a dynamic overcrowding detection system that encourages social separation and sends a dynamic and privacy-preserving alert to the venue, municipal council, or facility managers. The proposed method can also be used to detect when social distancing rules are not being followed, allowing for a quick response and restricting or reducing the virus's transmission. DBSCAN is a clustering algorithm that is unique. This strategy, as the name implies, concentrates on the proximity and density of data in order to build clusters. Clusters are formed by the average value of cluster components, but every data point contributes to the formation of the clusters. This means that the suggested Crowd Tracing technology was able to identify 85 out of 100 cases in which social distancing rules were broken.

I. Introduction:

The COVID-19 outbreak has altered people's perceptions of the world. A considerable section of the population now works or studies from home and makes purchases online. The issue has the potential to undermine local economies by hurting main streets, shopping malls, and educational institutions. As city halls reopen and residents return to work and school, the focus has moved from avoiding contact entirely to ensuring that social separation is maintained in order to prevent new COVID-19 breakouts. Crowd monitoring in the traditional sense focuses on improving footfall and activity in public spaces like high streets. Due of the recent explosion, however, the attention has switched to protecting city residents by ensuring compliance with social distancing rules and spotting "hotspots" of overcrowded activities[1].

Manually crowds counting and social distancing solutions have already seen extensive use in an effort to limit the spread of COVID-19 through Interaction Tracing systems in order to address the overcrowding problem in city settings [2]. While these systems are efficient in alerting persons who have just interacted with a recent positive case of COVID-19, they do so after the infection has already occurred. As a result, automated social distancing detecting systems are required to enable for a better and more convenient reaction. Conventional footfall monitoring techniques, such as webcams, have limitations with range, expense, and an incomplete data about individuals[3].

Another prospective crowd monitoring technology is the Global Positioning System(GPS) that allows mobile devices to really be tracked and tracked. A collaborative Geo location app was created to aid in the control of crowd mobility throughout large-scale gatherings [4]. The app drew in 28,000 users, who participated to 25 hundred million position modifications. While the study necessitated the downloading of an app, information was compiled to aid in the secure preparation and implementation events [5]. The fact that GPS requires an application to be loaded on the tracked user's mobile device highlights a fundamental limitation of the technology for crowd counting. Furthermore, GPS technology's correctness is extremely dependent on the surrounding and the frequency of GPS satellites in range, rendering it impracticable for inside locations. [6]

In this research, we offer Crowd Tracing, a consistent overcrowding detection system that uses a dense clustering technique to identify overcrowding and give alert messages of heavy crowd clusters. To identify overcrowding groups from sensor application records, we used density-based spatial clustering. The identification of clusters can then be used to alert event organizers or crowd controls to the presence of a potentially big crowd near the network overlay.

II. Related Work:

Depending on the statistics intercepted by pervasive access points, a multi-fold technique called CrowdTracing is offered as a means of identifying overcrowding and social separation in this study. The advancement of WIFI Probes, implementation through investigate intersection and multiple methods using a network of WiFi access points, fully automated partitioning of a large space, such as neighborhoods or cities, into smaller places, such as shopping malls or supermarket chains, to identify overpopulated places in which social distancing guidelines are not followed are the essential aspects of the system. Finally, a density-based clustering algorithm includes pre-defined social distancing and minimum grouping points to identify whether social

distancing norms are met in a given area, as well as a real-time alert system. The user identification and density-based clustering phase's assists in the discovery of arbitrary-shaped clusters and the identification of groups of individuals who do not adhere to social distancing. There are a variety of clustering strategies available that includes.

Density-based clustering: Identifying where data points are dense, while treating less dense areas and sparse statistical results as noise. In this case, the distance between grouping points and the density i.e., the required number of points for a collection of data points to be called a cluster can be modified according to the browser's needs. Density Based Spatial Clustering seems to be the most extensively deployed density-based clustering approach.[8]

Distribution clustering: Distribution Clustering is based on distribution assumes that the data is made up of probability distributions, the most common of which is the Gaussian distribution. The chance of a piece of information contributing to a distribution based cluster diminishes as the distance between it and the distribution's centre grows in distribution clustering. Gaussian Mixture Models are one example of this form of clustering.[9]

Hierarchical clustering: Hierarchical clustering uses a top-down or bottom-up strategy to group data. On the one hand, the top-down technique starts with each and every data point in a particular set in the same distinct cluster and divides the data points into smaller clusters depending on their closeness or dissimilarity. In the bottom-up technique, on the other hand, each data point is treated as a separate cluster that is a relevant fact combined with other clusters based on their similarity. When a particular stopping criterion is reached, both the bottom-up and top-down techniques are terminated.[10, 11].

Centroid Based Clustering: The similarities between data points is estimated as their proximity to the centroids of a user-defined collection of clusters k , where those centroids are continuously re-located to compute the distance, generally Euclidean distance, between them because the points associated with their respective clusters. Given the necessity to define k prior to the training phase, it is considered that an a purely rational understanding of the structure of the training data points exists. K-means is one of the most widely used Centroid-based clustering methods.[12, 13]

I. Methodology:

Density Based Clustering provides a convenient platform for detecting the maintenance of social distance regulations, given the properties of the above categories of clustering. This is based on the below major considerations. Initially Density Based Clustering can recognize clusters of any shape, indicating that persons in close proximity can be detected regardless of their

relative positions. Second, depending on the application, the parameter for ϵ the minimum distance between two points to be considered in the same cluster can be modified. As a result, it can be regularly updated in accordance with the most recent instructions issued by relevant government agencies or medical institutions.

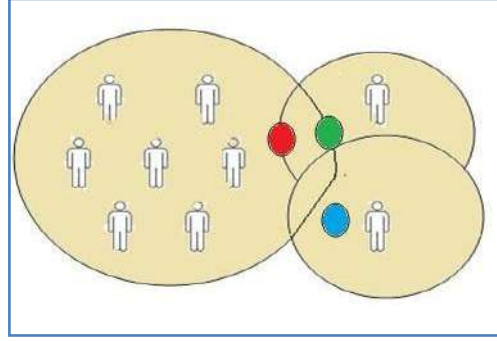


Fig 1: DBSCAN

The Red color represents the core point, next green color represents border point and finally blue color represents noise point within distance n from itself. The priority was focused on density clustering because of the suggested system's application, which is to detecting people gatherings. As a result, we use DBSCAN, a cutting-edge density clustering technique that uses two major characteristics: distance ϵ and the minimum level of points Minimum points that must be within such a radius for those same points to be designated a cluster. The two components, an estimation of the clusters' point density, are taken into account. DBSCAN can restrict the density and thus resolve the occurrence because of its predetermined minimal distance. To tests the suggested social distancing detecting algorithm. [14] The distances or lengths between each and every user n and each access point are determined using the Euclidean distance formula. After that, the distances are transformed to their corresponding Received Signal Strength Indicators (RSSI).

$$RSSI = -N_p - 2E_F \log_{10} S \quad (1)$$

Where N_p is the expected RSSI,

E_F is the constant environmental factor.

Each of the predicted RSSI values is given random noise from a standard deviation, and then they are transformed to distances using Equation 1. The new coordinates such as A noisy and B noisy

$$\left(\frac{A_1 + A_2 + A_3 + A_4}{4} \right) \left(\frac{B_1 + B_2 + B_3 + B_4}{4} \right) \quad (2)$$

of each and every user n are computed using tri-iteration averaging as the median parameters generated by four tri-iteration processes employing three points of entry at the present moment.

The suggested system is designed to determine whether density-based spatial clustering of people exists. Obtained using the proposed clustering algorithms, social distance norms are maintained in each separate scenario. DBSCAN is preferred over the other clustering methods because of its reliability and low overhead. Furthermore, DBSCAN's capability to establish a distance and minimum points, such as for the minimal number of individuals and the minimum distance among individuals, makes it ideal for detecting overcrowding in open places when individuals are not sustaining the acceptable social distance around themselves.

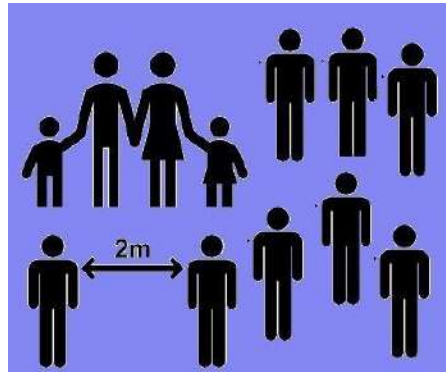


Fig 2: Recommended Social Distancing

As a result, the cluster is continuously extended until no new items are added to it, i.e., all of the objects examined in the recurrence stage are boundary objects. Objects or places that have been visited before. Let T be the minimal number of individual users within the radius ϵ that the neighborhood must have in order to be labeled non-compliant with social distancing laws. Using real-world coordinates as A noisy and B noisy. It evaluates if these individuals are maintaining social distance by subtracting 2 meters from ϵ each of them and substitute T by 4. The use of T is equal to 4 is intended to prevent system alarms from being activated by a small amount of individuals within the same neighborhood. While detecting social distancing, the noisy parameters as A noisy and B noisy, ϵ are employed as a hyper parameters of the similarity measure to optimize the similarity between the noisy situations and the genuine scenarios.

III. Result:

The system performance is evaluated as a function of the radius of the region, as described in the preceding section. The idea is to accommodate for the unwanted noise that is normally present in radio frequency communications received signal strength indicators. Although it is critical to accept responsibility for unwanted noise in the RSSI, permitting for a certain widening of the neighborhood radius when particularly in comparison to that used with genuine coordinates, and to recognize the maximum number of situations in which social distancing rules are broken, it is also critical not to produce a large number of false reports, which can cause

unnecessary disruption to monitoring or detecting institutions.

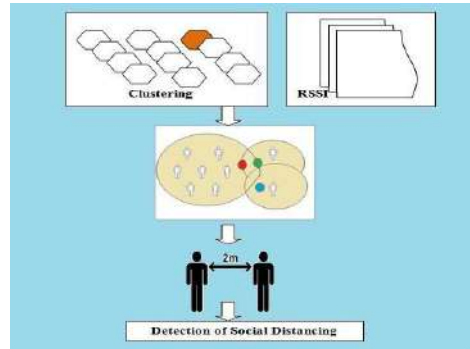


Fig 3: Framework of Crowd Tracing.

As a result, a fair balance between the system's precision, recall and F-measure measurements should be maintained. Increasing the radius of the neighborhood will undoubtedly result in higher recall levels; however, this increase may also result in an overly large number of false positives. As a result, the classification performance measures are used as a reference metric for reporting the system's best results. The Performance Measures of the radius of neighborhood is shown in the figure.

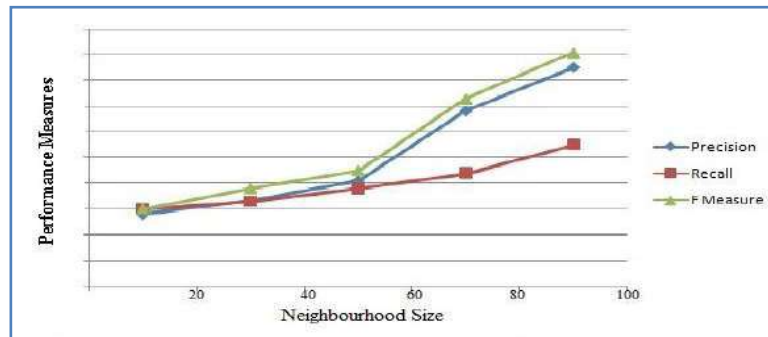


Fig 4: Performance Measures of the radius of neighbourhood.

IV. Conclusion:

The COVID-19 outbreak is extraordinary, and it has inflicted effects on the lives of millions of people all over the world. This outbreak has presented cities with a number of research problems and possibilities, which they must meet by implementing place-based preventive interventions the virus from spreading further. The preliminary findings obtained by the proposed Crowd Tracing method are encouraging, and it has the potential to aid in the detection of densely populated areas. Furthermore, with classifying results of the technique aimed at determining if social distancing rules are obeyed within particular places show that the same access points can be used to help to control social distancing in areas of interest.

References:

- [1]. Kieran Woodward, Eiman Kanjo, et al., Low cost wearable that acts as a social distancing reminder and contact tracer: Poster abstract. In Proceedings of the 18th Conference on Embedded Networked Sensor Systems, SenSys '20, page 758–759, New York, 2020.
- [2]. Balamurugan Soundararaj, et al., Estimating real-time high-street footfall from wi-fi probe requests. International Journal of Geographical Information Science, 34(2):325–343, 2020.
- [3]. Enrique Hernández-Orallo, et al., Evaluating how smartphone contact tracing technology can reduce the spread of infectious diseases: the case of covid-19. IEEE Access, 2020.
- [4]. Soyoung Hwang and Donghui Yu. Gps localization improvement of smartphones using built-in sensors. International Journal of Smart Home, 6(3):1–8, 2012.
- [5]. Ulf Blanke, Gerhard et al., Capturing crowd dynamics at large scale events using participatory gps localization. In 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), pages 1–7. IEEE, 2014.
- [6]. Soyoung Hwang and Donghui Yu. Gps localization improvement of smartphones using built-in sensors. International Journal of Smart Home, 6(3):1–8, 2012.
- [7]. Soyoung Hwang and Donghui Yu. Gps localization improvement of smartphones using built-in sensors. International Journal of Smart Home, 6(3):1–8, 2012.
- [8]. Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In Kdd, volume 96, pages 226–231, 1996.
- [9]. Arvind Krishna et al., Distributional Clustering: A distribution-preserving clustering method. Statistica Sinica 2019.
- [10]. Yayan wang et al., An Objective for Hierarchical Clustering in Euclidean Space and Its Connection to Bisecting K-means. Proceedings of the AAAI Conference on Artificial Intelligence, 34(04), 6307-6314. 2020.
- [11]. Wen-bo xie et al., Hierarchical clustering supported by reciprocal nearest neighbors. Information Sciences. Vol 527, pp 279-292. 2020.
- [12]. Santosh Kumar Uppada. Centroid Based Clustering Algorithms-A Clarion Study. International Journal of Computer Science and Information Technologies, Vol. 5 (6). 2014.
- [13]. Mohammad Rezaei, Improving a Centroid-Based Clustering by Using Suitable Centroids from Another Clustering. Journal of Classification. Vol 37. Pp 352-365. 2020.
- [14]. Abdulraheeb Alhammadi et al., A three-dimensional pattern recognition localization system based on a Bayesian graphical model. Int. J. of Distributed Sensor Networks. Vol 16 (9). 2020.

IOT BASED UNDERGROUND CABLE FAULT DETECTION

*Mr.Rajath V¹, Swetha Shekarappa .G²
Department of Electrical & Electronics Engineering,
Alliance University,Bengaluru,India*

ABSTRACT

The objective of the proposed work is to select the crevice of underground line cable blame from the foot office in term of distance and shown on the web. Underground line cable gadget could be unordinary put went with in vital locales in Urban locations. Where as blame happens for vital reasons, at that point the fathoming procedure related with that exact cable is difficult since of real obscure range of the blame with inside the cable. This IOT Innovation is utilized to find the exact region of the blame and to dispatch records in graphical format to our web location the utilize of an IOT module on the in distinguishable time it appears at the LCD screen. This paper makes utilize of the normal thought of Ohms law, i.e., whereas a moo DC voltage is actualized on the feeder haltthrua grouping resistor(Cablelines), at that point the modern may run depending up on the range of the blame with inside the cable since the resistance is corresponding to the hole . In case there may be a fast circuit(LinetoGround) ,the voltage all through collection resistors adjustments in step with the resistance that adjustments with separate .Usually at that point nourished to an ADC to broaden virtual records which the modified micro-controller of the 8051-claim circle of relatives appears in kilometres.

Keywords: Micro-controller, Relays, ADC, Underground cable Fault and 230v power supply.

Introduction

In city regions, the electric link runs underground rather than overhead lines. At whatever point the deficiency occurs across underground link it's miles vital to find the accurate space of the short coming for way of fixing that exact link. The proposed device recognizes the exact space of the shortcoming and with the guide of utilizing the methodology of IOT it's sequentially imparted toward worker. Since inconvenience that occurs in underground link is an enormous difficulty up to this point. As it's miles extremely intense to search the accurate region or faulty region physically, which at the same time impacts the exhibition of the link rope in view of misfortune sharpened. Till now numerous system shad effectively were completed to find short coming in link line. However, the difficulty arrived up is the best approach to find deficiency in link string while it's miles underneath grounded, and the be approach to get passage to or recover the ones records related with damaged region each time it

smiles required. To fill the ones holes, we executed the contraption which identifies the accurate space of the deficiency and by means of the methodology of IOT it's sequentially conveyed toward worker. Through going before explores numerous systems arrive dup which had been gainful to vanquish the difficulty as much as a couple of degrees. In one of the papers with the guide of utilizing K. Hasan. Says that- disappointment and debasing of air make wiring is a huge test that may moreover cause hearth place and smoke due to arcing. In any case, the proposed technique essentially dependent on TDR, wherein instruct of heartbeats are created to find the issue as discussed in [1]. Shortcoming in link might be named in gatherings: Open circuit shortcoming: Open circuit deficiencies is higher than brief circuit issue, because of the reality while this flaw happens present day streams through link will get zero. This type of flaw is a result of crush in endeavour way. Such blames emerge [2] while one or additional part conductor's crush.

Balanced flaw: Three-segment shortcoming is alluded to as even deficiency. In these every one of the 3 levels are brief circuited [4].

Unsymmetrical deficiency: In this shortcoming worth of present day isn't generally equal dislodged through method of method for 120degrees.[5]

Deficiency district approach: Fault locale procedures might be named as [6].

Online methodology: This methodology utilize strategy the tested voltages & present day to choose the flaw focuses. Online methodology of underground link is considerably less than over headlines.

Disconnected methodology: In this methodology one of a kind gadget is utilized to look at supplier of link with-inside the field. There are disconnected methods as shown below.

Tracer approach: Deficiency factor is recognized through method of method for by walking at the link lines. Shortcoming factor is demonstrated from discernible sign or electromagnetic sign. It is utilized to pinpoint flaw are a precisely.

Block Diagram and it's working

This model uses resistors to indicate fault location distance in kilometres. The resistors from RR1to RR5address R phase of the link. Additionally, RY1 to RY5 and RB1 to RB5 establish Y and B fragment of the links. RN1 to RN12 are utilized to re-frame the fair- minded link. To establish frequency of flaw in underground lines toggle switches are utilized. All fragments are snared with hand-off which in flip is snared to Port C of Microchip. When there is no issue, the LED's connected to each hand-off glows.

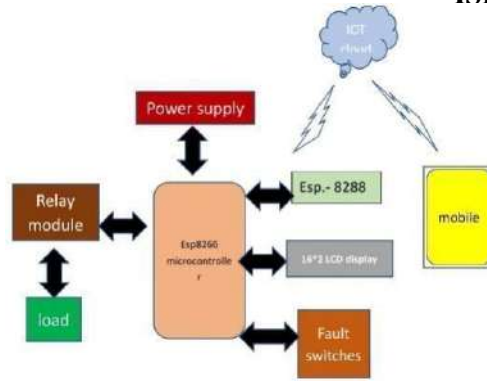


Fig 1: Block diagram of Underground line cable fault detection system

At the point when an exchange connected to a fragment is shut, the LED connected to the exact section with out help from any one else sparkles .The opposition connected to that exact fragment gives up and dipin voltage thus produced is given to Port of the Microchip. The dip in voltage is changed to remove as per Table 1 and is shown with-inside LCD. Moreover, the pin of Port C is connected to that exact LED is going over the top and the call of the blamed fragment is shown with-inside the LCD.

The main test is to search the hole of underground link flaw from the base location in kilometres. In numerous city regions, link shortcoming is a not strange spot issue. At the point when a flaw happens in view of a couple of reasons, the way of issue observing without understanding the spot related with that interesting link might be troublesome. The proposed device is intended to music the exact spot of the deficiency happened with-inside the link.

This test utilizes the Ohms Law thought, while a low voltage DC is completed to the feeder stop through a chain resistor, then, at that point the present day could change fundamentally based absolutely at the spot of deficiency happened with-inside the link. On the off chance that is there any fast circuit occurred from line to ground, then, at that point the voltage all through assortment resist or changes in like manner, then at that point it's miles took care of to a simple to virtual converter to grow certified information, which the pee-modified Arduino micro-controllers will show in kilometres.

The proposed contraption is planned with a fixed of resistors to meaning the term of a link in kilometres, and the flaw approach is planned with a fixed of toggle switches at each perceived distance (KM) to cross-investigate the precision of the equivalent. The issue going on at a specific distance and the extraordinary area is shown on a LCD interconnected to the Arduino micro-controllers. also, Blynk application.

Hardware Design

The hardware used in this proposed approach is given below:

Power Adaptor: 12V energy elements (or 12VDC energy elements) are one of the maxima not unusual place energy elements in use today. In general, a 12VDC output is acquired from a 120VAC or 240VAC enter the usage of a aggregate of transformers, diodes and transistors..Switching regulated 12VDC energy elements, occasionally called SMPS energy elements, switchers, or switched mode energy elements, alter the 12VDC output voltage the usage of a complicated excessive frequency switching method that employs pulse width modulation and feedback. Apopain switching regulated energy elements additionally appoint sizable EMI filtering and protecting to reduce each not unusual place and differential mode noise performed to the road and load. Galvanic isolation is trendy in our 12VDC switchers, affording our customers enter to output and output to floor isolation for max versatility. Apopain switching regulated energy elements are quite efficient, small and lightweight, and are to be had in each AC- DC unmarried and extensive-regulate output and DC-DC configurations. Our Low Profile extensive regulate output switchers may be voltage or modern-day regulated and are externally programmable.

Sl. No.	Name	Specifications
1.	Operating Voltage	5V
2.	Input Voltage (recommended)	7-12
3.	Input Voltage (limits)	6-20V

Arduino UNO: Arduino UNO is a microcontroller with product call Atmega328. Arduino UNO performs a completely essential position with inside the in verter stage. The gate drivers which drive MOSFET's are managed via way of means of this micro controller. The software within side the microcontroller can extrude the frequency of the output, sort of pulse, pulse width etc. The essential specifications are provided in table 1. The Arduino Uno is a microcontroller board primarily based totally at the ATmega328 (datasheet). It has 14virtual input/output pins (of which 6 may be used as PWM outputs), 6 analogue inputs, a sixteen MHzceramic resonator, a USB connection, a energy jack, an ICSP header, and a reset button. It consists of the whole thing had to help the microcontroller; actually, join it to a laptop with a USB cable or energy it with an AC-to-DC adapter or battery to get started. The Uno differs from all previous forums in that it does now no longer use the FTDI USB-to-serial motive force chip.

The pin diagram of Arduino is presented in figure2.

Table 1: Microcontroller specifications

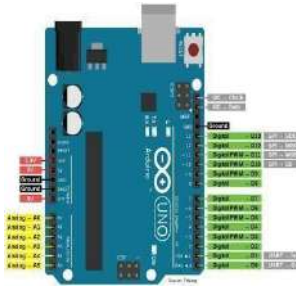
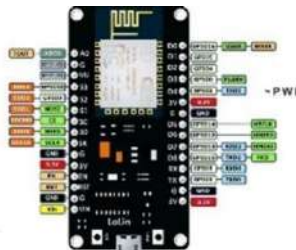


Fig. 2: Pin diagram of Arduino UNO

ESP8266-12e: The ESP8266 WIFI Module is a self-contained SOC with included TCP/IP



protocol stack that could supply any microcontroller get right of entry to in your Wi-Fi community. The ESP8266 is able to both web website hosting an software or offloading all Wi-Fi networking features from every other software processor. This module comes with AT instructions firmware which permits you to get capability like Arduino WIFI shield , but you may load distinct firmware's to make your personal software at the modules' reminiscence and processor. It's a complete lymonetary module and hasa big and developing network support. This module has onboard 80Mhz low energy 32-bit processor which may be used for custom firmware's. This additionally manner that you may host small web pages with none outside controller. For greater information see: NODEMCU. The ESP8266 helps APSD for VoIP programs and Bluetooth co-life interfaces; it carries a self-calibrated RF permitting it to paintings below all working conditions, and calls forno outside RFparts. ESP8266 is remodelling the arena with it slow value and excessive capabilities which makes it a really perfect module for Internet of Things (IOT). It may be utilized in any software in which you want to attach a tool in your nearby community or internet. This module ESP-12 has 7GPIOs.

Four-Channel Relay Module: The 4-channel relay module consists of 4- 5V relay and the related switching and setting apart additives, which makes interfacing with a microcontroller or sensor clean with minimal additives and connections. There are terminal blocks with six terminals every, and every block is shared through relays. The terminals are screw type, which makes connections to mains wiring clean and changeable. The 4 relays at the module are rated for 5V, because of this that the relay is activated while there's about 5V throughout the coil. The contacts on every relay are specific for 250VAC and 30VDC and 10A in every case, as marked at the frame of the relays.

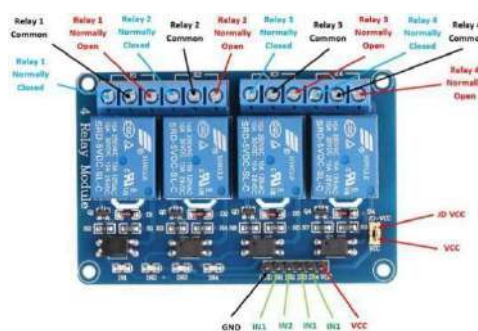


Fig. 4. Four-Channel Relay Module Pin Diagram

LCD Module:

LCD modules are very usually utilized in maximum embedded projects, the purpose being its reasonably-priced price, availability and programmer friendly. Most folks might have stumble upon those shows in our everyday life, both at PCO's or calculators. The look and the pin outs have already been visualized above now allow us to get a chunk technical.

Features of 16×2 LCD module

- Operating Voltage is 4.7V to 5.3V
- Current intake is 1mA without backlight



IV Software Design

Arduino is open-supply bodily processing that is primarily based totally on a microcontroller board and integrated improvement surroundings for the board to be programmed. Arduino profits some inputs, for example, switches or sensors and manipulate some more than one output, for example, lights, engine and others. Arduino application can run on Windows, Macintosh and Linux running structures (OS) contrary to maximum microcontrollers' frameworks which run most effective on Windows. Arduino programming is straightforward to examine and follow to novices and amateurs.

There are several exclusive microcontrollers and microcontroller systems available for bodily computing. Parallax Basic Stamp, Net media's BX-24, Phi gets, MIT's Handy board, and several others provide comparative usefulness. These apparatuses take the chaotic diffused factors of microcontroller programming and wrapitupina easy to-make use of bundle. Arduino moreover rearranges the technique of operating with microcontrollers; moreover, it gives a few benefits for instructors, students, and intrigued individuals:

Inexpensive - Arduino forums are fairly reasonably-priced as compared with different microcontroller forums. The most inexpensive model of the Arduino module may be accrued through hand, or even the preassembled Arduino modules price brief of what\$50.

Cross-platform - The Arduino programming runs more than one running structure Windows, Macintosh OSX, and Linux operating frameworks. So, we finish that Arduino has a bonus as maximum microcontroller frameworks are restrained toWindows.

Straight forward, clean programming method-The Arduino programming surroundings is straight forward to apply for novices, but sufficiently flexible for slicing ideclients to journey as well. Fore ducators, its favourably engaged across the Processing programming surroundings, so understudies locating methods to recognize the way to application in that surroundings may be acquainted with the character of Arduino. Open supply and extensible programming. The Arduino application language is to be had as open supply, to be hadfor

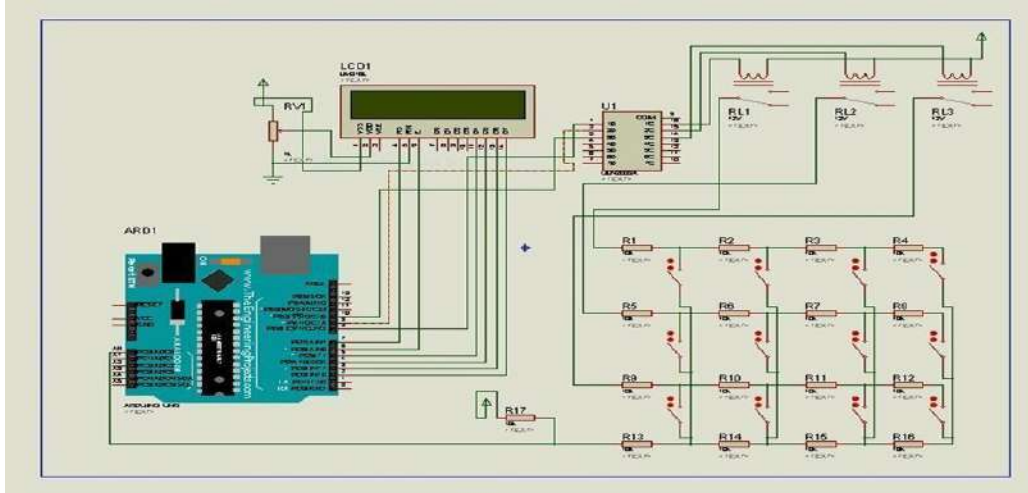


Fig. 6: Under Ground Cable Fault Detection System Circuit design

improvement through skilled engineers. The lingo may be reached out thru C++ libraries, and those looking forward to recognize the precise functions of various pastimes can take the plunge from Arduino to the AVR C programming language on which it's far primarily based totally. Basically, you may comprise AVR-C code honestly into your Arduino packages when you must Open supply and extensible hardware - The Arduino is focused round Atmel's Atmega8 and Atmega168 microcontrollers. The plans for the modules are circulated below a Creative Commons license, so skilled circuit designers could make their personal specific interpretation of the module, extending it and enhancing it. Barely green clients can construct the bread board variant of the modulerem embering the completed goal to understand the way it capacities and store money.

Cross-platform - The Arduino programming runs more than one running structures Windows, Macintosh OSX, and Linux operating frameworks. So we finish that Arduino has a bonus as maximum microcontroller frameworks are restrained to Windows.

Result and Discussion

The working and output for 3 phase underground cable with fault is shown in figure 7. The control lersensesthe change in resistance and changes the ADC value. In the figure 7, hardware setup for fault identification at every 1 Km length of the cable 1 K Ω resistor is used. and for each KM one switch used to create fault.



Fig. 7. Hardware setup for the proposed work*Fig. 8. 4km fault creating in Y phase by manual**Fig. 9. Output retrieved from the display.*

Table 2: Analytical Tabulation Result

Switch shorted	Voltage across series resistor (V)	Distance where Fault occurred (Km)	Resistance of the cable ($K\Omega$)	ADC Output
SW1	3.30	1 Km of first cable	3	180
SW5	4.00	2Km of second cable	2	204
SW9	4.40	3 Km of third cable	1	230

Table 2 shows the values of voltage, distance, resistance and ADC output of the controller under different fault conditions. For instance, if we look from the leftmost switch, the third switch is pressed, it will create short circuit fault and exclude the $3K\Omega$ resistances from the circuit so the fault distance is 3km from the feeder end. Figure 10 shows the same information with cable resistance is display in the LCD.



Fig. 10. Output retrieved from Blynk Application

Conclusion and Future Scope

Underground cables provide a cheap and inexpensive solution for key components and, in some cases, the entire high-voltage overhead line. Using the right technology in the right place can minimize the impact of underground cables on the environment. The purpose of this project is to use an Arduino microcontroller to locate the accurate location of the circuit deficiency in the underground cable in kilometres from the end of the feeder. The Arduino microchip functions based on the output impedance of the wire. Health line. This design only identifies the location of the line break in the buried cable link, but it can also be modified to locate an open circuit. Change the impedance and manipulate the distance to the fault

REFERENCE

1. Xiaoyun, Qu, Kang Xiaoning, Zhang Chao, Jiang Shuai, and Ma Xiuda. "Short-term prediction of wind power based on deep long short-term memory." In 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), pp. 1148-1152. IEEE, 2016.
2. Cheung, Gilbert, Yuan Tian, and Tobias Neier. "Technics of locating underground cable faults inside conduits." In 2016 International Conference on Condition Monitoring and Diagnosis (CMD), pp. 619-622. IEEE, 2016.
3. Sain, Nikhil Kumar, Rajesh Kajla, and Mr Vikas Kumar. "Underground Cable Fault Distance Conveyed Over GSM." International Organization of Scientific Research Journal of Electrical and Electronics Engineering 11, no. 2(2016).
4. Bharatiraja, C., S. Jeevananthan, and Josiah L. Munda. "A timing correction algorithm-based extended SVM for three-level neutral-point-clamped MLI in over modulation zone." IEEE Journal of Emerging and Selected Topics in Power Electronics 6, no. 1 (2017):233-245.
5. Jaradat, Manar, Moath Jarrah, Abdelkader Bousselham, Yaser Jararweh, and Mahmoud Al-Ayyoub. "The internet of energy: smart sensor networks and big data management for smart grid." Procedia Computer Science 56 (2015):592-597.

- 6.Dharani, Dhivya. "A and Sowmya. T,""Development of a Prototype Underground Cable Fault Detector"." International Journal of Electrical, Electronics and Computer System (IJECS) 2, no. 7 (2014):17-21.
- 7.Islam, Md Fakhrul, Amanullah MT Oo, and Salahuddin A. Azad. "Locating underground cable faults: A review and guideline or new development. "In 201222nd Australasian Universities Power Engineering Conference(AUPEC),pp.1- 5. IEEE,2012.
- 8.Fonseca_Badillo, M., L. Negrete_Navarrete, A. González_Parada, and A. Castañeda_Miranda. "Simulation and analysis of underground power cables faults." Procedia Engineering 35 (2012):50-57.
- 9.Pandey, Abhishek, and Nicolas H. Younan. "Underground cable fault detection and identification via fourier analysis." In 2010 International Conference on High Voltage Engineering and Application, pp. 618-621. IEEE,2010.
- 10.Neier, Tobias. "Cable fault location practical experience." HV Technologies, version-1(2006).
- 11.Shi,Qinghai,UweTroeltzsch,andOlfaKanoun."Detectionandlocalizationofcablefaultsbytimeandfrequencydomain measurements." In 2010 7th International Multi-Conference on Systems, Signals and Devices, pp. 1-6. IEEE,2010.
- 12.Lee, Duck-Su, Xia Yang, and Myeon-Song Choi. "A line to ground fault location algorithm for underground cable system." The Transactions of the Korean Institute of Electrical Engineers A 54, no. 6 (2005):267-273.

HYBRID APPROACH OF SET PARTITIONING IN HIERARCHICAL TREES WITH CDF WAVELET FOR MEDICAL IMAGE COMPRESSION

P.Jeyanthi¹

Research Scholar, Department of Computer Science

Dr. K.Prabavathy²

Assistant Professor, Department of Computer Science

Rathinam College of Arts & Science, Coimbatore, TN, India

ABSTRACT

The intensive usage of computers in the medical field necessitates the primary need for significant handling of storage space and band width, as huge volume of medical images have to be processed in runs together. Image compression has become indispensable to lessen the bandwidth needed for communication and reduce the memory space & power of transmission. In spite of the existence of plentiful approaches for data compression, nevertheless, none of the methods are capable of consistently attaining the theoretical best-case compression ratio. Hence there is a need for more effective algorithms to fulfil the expectations. Challenging aspect of medical image compression is to preserve the quality of the images regardless the level of compression, since any loss of vital information during the course of the compression procedure would be a life threatening concern. The primary objective of this paper is to present an approach that attains these merits in lesser time duration by combining together the most successful Set Partitioning in Hierarchical Trees (SPIHT) with Cohen–Daubechies–Feauveau (CDF) Wavelet. The results are assessed against Discrete Wavelet Transform, one of the most prevalent existing approach

Index Terms – Data compression, Image compression, SPIHT, DWT, CDF.

INTRODUCTION

Digital image compression is indispensable due to the progressive increase in several happenings such as multimedia, satellite communication, internet teleconferencing, high definition television technologies and lot more. It has become a great deal to manage with the increased storage requirement and transmission bandwidth. The way out to this challenge is to compress the information and Image compression is achievable because of the existence of information redundancy in the images. Hence the accomplishment of compression is owing to redundancy and irrelevancy reduction where redundancy represents the duplication and

irrelevancy implies the portion of the information contained in the image which the human visual system that will perceive.

In the domain of image compression, the wavelet transform are the most influential and extensively applied tool. Owing to its competence in efficiently representing and assessing the data, wavelet transforms have emerged as prevalent technique for image processing in the previous decade. Image compression algorithms that are grounded on discrete wavelet transform like Embedded Zero Wavelet (EZW) compress the images into a bit stream in an intensified precision using the method of progressive coding. Set partitioning in hierarchical trees (SPIHT) is a wavelet based technique which is considered as one of the best image compression technique that provides a high speed computation and achieves excellent compression ratios, improved image quality and speedy execution.

SET PARTITIONING IN HIERARCHICAL TREES (SPIHT)

It is considered by the researchers that the application of wavelets produce exceptional results in comparison with that of the results by DCT or fractals. Among the wavelet based approaches, SPIHT has succeeded in reconstructing the image as much undifferentiated as the original image. Apart from this, in the process of encoding colour images, the impact of SPIHT is highly prominent since the bits are automatically assigned for local optimality amongst the colour components. In contrast, all the other encoding schemes operate on the colour components individually on the basis of the global statistics of every single component.

SPIHT is one of the contemporary approaches meant for optimal progressive image transmission by which a fully embedded coded file is generated in such a way that the image quality at any instant of the process is the best obtainable for the number of bits received up to that moment of time. Hence, when there is a need for the user to review hastily the image and end up with the conclusions upon whether the image is adequate in quality or else should undergo improvement, SPIHT made this possible by embedded coding optimization and also by first coding the highly significant information every time. The better speed of coding and decoding is another most cherished advantage of this method and also it is almost symmetric in the sense that it takes nearly the same time for both encoding and decoding. The features that are exiting in most of the images are exploited by SPIHT and is utilised for the encoding of natural images such as portraits, medical images like CT scan, X-rays and so on, scientific data, elevation maps and many more.

Coder and Decoder in SPIHT

The pictorial representation of the SPIHT coder and decoder is as shown below in Fig 1. The original image is given as input to the decomposition block first and the compressed file which is achieved as the output would then be delivered to the reconstruction block for the image to be recovered at the receiver end.

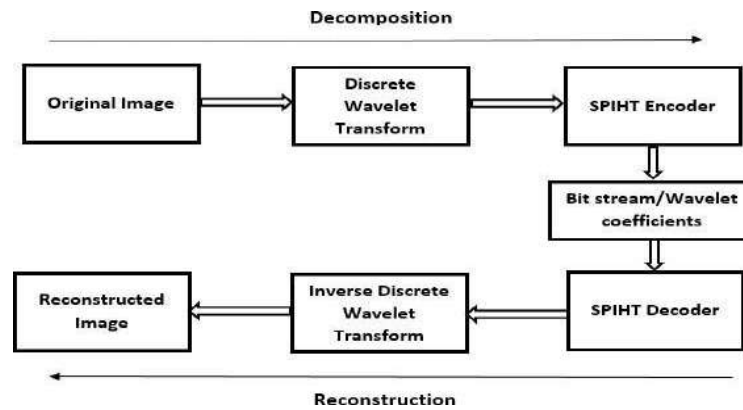


Fig. 1 Block diagram of SPIHT technique

The image to be processed is forwarded to the block of DWT in which the image decomposition happens and the wavelet coefficients of the image are delivered as the output from this block. The coefficients are then passed to the SPIHT encoder in which the coefficients are encoded and the output is delivered in the form of a bit stream. This stream of bits is then supplied to the decoder block wherein the decoding of the bit stream happens. Then comes the play of the Inverse DWT (IDWT) block that produces the reconstructed image.

Working principle of SPIHT

SPIHT works under the principle of self-similarity across the scale and is accompanied with DWT. This algorithm is meant for the process of encoding and decoding the wavelet transform of an image. It is one of the strongest algorithms known for image compression. The decomposed wavelet is partitioned into significant and insignificant elements by the algorithm based on the function.

$$S_n(T) = \begin{cases} 1 & \max_{(i,j) \in TM} \{ |C_{(i,j)}| \} \geq 2^n \\ 0 & \text{otherwise} \end{cases}$$

Where, $S_n(T)$ - significance of a set of coordinate T $C(i, j)$ - coefficient value at coordinate (i, j)

The Parent-child relationship between the sub-bands of the image is depicted beneath:

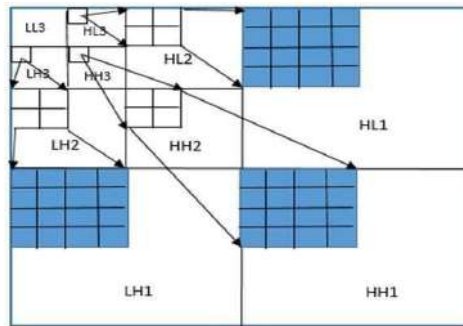


Fig. 2 Sub-band parent-child relationship

The scanning order of the coordinates of the coefficients in the sub-bands of the transformed image by SPIHT is as shown in the following fig.

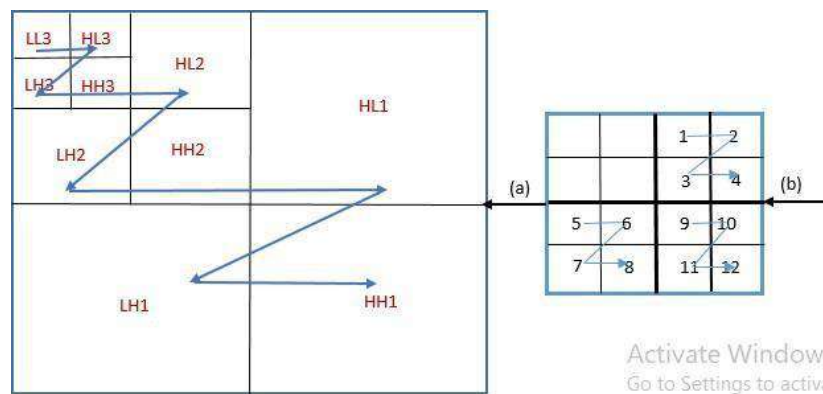


Fig. 3 Scanning order in SPIHT

Sets of coordinates:

The algorithm works based on four categories of the sets of coordinates of coefficients which are explained as follows:

H - set of all root nodes

O (i, j) – set of coordinates of the off springs of the wavelet coefficient at the location (i,j)

D (i, j) - set of all descendants of the coefficients at the location (i,j)

L (i, j) - set of coordinates of all descendants of the coefficients at the location (i,j)

Lists of wavelet coefficients:

Apart from the sets maintained, there are also three lists of wavelet coefficients being provided by SPIHT with each having its own functionality been explained as follows:

List of insignificant pixels (LIP) – Coordinates of individual coefficients that are insignificant ie., with magnitudes lesser than the current threshold are maintained in this list

List of significant pixels (LSP) – Coordinates of significant coefficients ie., possessing magnitudes greater than the current threshold are in this list

List of insignificant sets (LIS): - Coordinates of roots of all insignificant sub-trees are contained in this list The functioning of the SPIHT algorithm is as shown in the following fig:

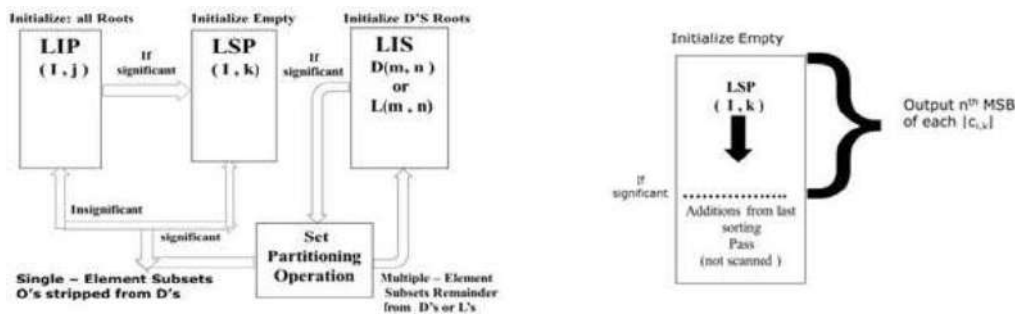


Fig. 4 SPIHT functioning

SPIHT Algorithm

Here comes the work flow of the SPIHT algorithm:



Fig. 5 SPIHT workflow

Step 1: Initialization:

Threshold and order list initialized

All root node coefficients in low-pass sub-band are assigned to LIP

All trees assigned to LIS

LSP initialized as an empty set.

Step 2: Sorting pass: meant for encoding significant coefficient of current bit

All wavelet coefficients in LIP are analysed to establish whether they are significant or not

If yes, then output "1" and the sign bit (positive - 0; negative - 1); Move the coefficient from LIP to LSP

If no, output "0"; Coefficient not moved from LIP

Based on the type of trees, all significant trees in LIS are analysed

Step 3: Refinement pass:

For each element in LSP, if it is not added to the list in the current sorting phase, then nth most significant bit of each element is transmitted as output

Step 4: Threshold updating

The maximum number of bits needed to denote the largest coefficient in the spatial orientation tree is determined and represented by n , where

$$n = \lceil \log_2 c_{\max} \rceil$$

where, c_{\max} - maximum value of coefficient

Threshold decremented by 1 for each pass and go to step 2

Cohen–Daubechies–Feauveau (CDF) Wavelets

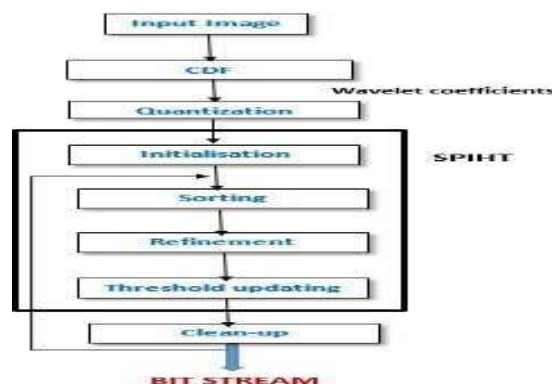
Cohen–Daubechies–Feauveau wavelets are a group of bi-orthogonal wavelets which was made a commonly accepted approach by Ingrid Daubechies. The CDF family wavelets has two concurring numbering systems as one for the low-pass filters' smoothness factors or equally for the high-pass filters' vanishing moments and the next one for the size of the filters employed for low-pass or equally for the high-pass.

In a filter bank, if the size of the filters are 7 & 9, they can be of the vanishing moments as 6 & 2 in case of the trivial factorization or else with the vanishing moments as 4 & 4 in case of JPEG 2000 wavelet. Then the wavelet could be denoted on the basis of filter sizes as CDF 9/7 or on the basis of vanishing moments as bi-orthogonal 4, 4. In the CDF 9/7 wavelet, the low pass filters have $p=9$ coefficients in the analysis block and accordingly $p=7$ coefficients in the synthesize block as given in the following table.

The lifting scheme is a procedure meant for wavelet designing and it is sensible to incorporate this phase in the wavelet transformation for the designing of the wavelet filters. Actually in the discrete wavelet transform for a single signal, quite a few filters are applied separately, whereas, in the lifting scheme, the signal is split alike a zipper and runs of convolution–accumulate processes are applied over the signal that is divided.

Proposed hybrid approach of SPIHT with CDF Wavelet

In this section the proposed hybrid methodology for medical image compression that combines the Cohen–Daubechies–Feauveau (CDF) Wavelets with Set Partitioning



Hierarchical Trees is described. The following fig represents the process of image compression in the proposed method.

Fig. 6 Workflow of proposed approach

The sample medical image and its corresponding first level decomposition achieved by the hybrid algorithm is as given below:

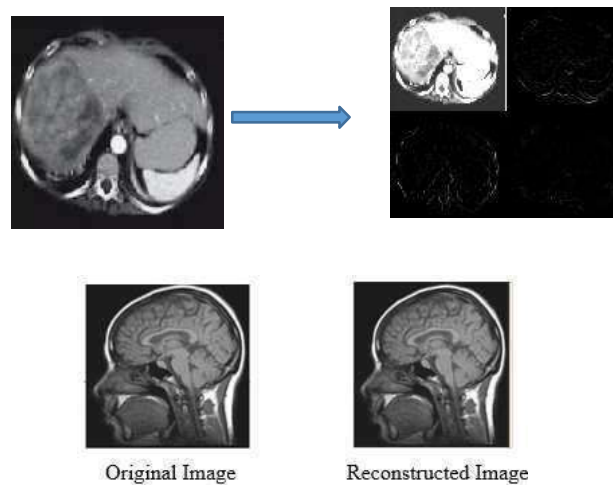


Fig. 7 First level decomposition of medical image

Experimental results and discussion

The results of an extensive set of simulation tests are shown for the proposed technique which is compared with the existing DWT for different metrics. The performance of the proposed approach is evaluated using eight different metrics for different compression rates starting from 0.1 to 1 and the results are obtained for two level decomposition. The medical images are taken from the CT scan and MRI image datasets. The sample reconstructed image after compression is as shown below. The results obtained by the convetional DWT algortihm and proposed method is given in Table 1 and Table 2 respectively.

Table1 ResultsobtainedbyconventionalDWT

COMPRESSION RATE	COMPRESSION RATIO	TIME ELAPSED	PEAK SIGNAL TO NOISE RATIO	MEAN SQUARE ERROR	AVERAGE DIFFERENCE	CORRELATION	NORMALIZED CROSS CORRELATION	NORMALIZED ABSOLUTE ERROR
0.1	29.5914	0.733221	9.0275	81.345	72.0634	0.401084	0.193135	0.917236
0.2	26.048	0.848732	11.3621	47.5192	46.8407	0.756142	0.627197	0.728397
0.3	18.1992	1.07943	13.5108	28.9734	30.4227	0.829978	0.800109	0.555804
0.4	14.2738	1.62799	15.7289	17.3856	16.6356	0.876602	0.927611	0.425877
0.5	11.9193	2.00916	18.7155	8.7404	13.0536	0.934647	0.947694	0.302774
0.6	10.3493	2.07013	21.0106	5.15256	6.05458	0.940987	0.98061	0.234893
0.7	9.22797	2.80206	22.4266	3.71895	5.3622	0.953401	0.977935	0.193289
0.8	8.38705	2.62595	24.44	2.33925	4.75478	0.969410	0.977712	0.144051
0.9	7.73298	3.18573	26.7438	1.37627	2.0829	0.977499	0.986259	0.110566
1.0	7.73298	4.06129	29.5710	1.0106	1.7289	0.936994	0.989974	0.110172

Table21 Results obtained by hybrid approach

REFERENCES

1. Allaeldien Mohamed G. Hnesh , Hasan Demirel, "DWT-DCT-SVD based Hybrid lossy image compression technique", IEEE IPAS'16: International Image Processing Applications And Systems Conference, November2016
2. Bhaskar Mishra, Pradeep Tripathi, "Lossless Medical Image Compression Based On Hybrid Technique", International Journal of Emerging Technology and Advanced Engineering, Volume 5, Issue 6, Pages 371 – 377, June2015
3. D. Cerná, V. Finek, "Discrete CDF 9/7 Wavelet Transform for Finite-Length Signals",2011
4. J. Malý, P. Rajmic, "DWT-SPIHT Image Codec Implementation", IEEE Transactions, Pages 340-344,2012
5. Jianxiong Wang, Fuxia Zhang, "Study of the Image Compression Based on SPIHT Algorithm", International Conference on Intelligent Computing and Cognitive Informatics, June2010

6.Karimella Vikram, Garlapati Narayana, Niraj Upadhayaya, “Medical Image Compression”, International Journal Of Electronics And Communication Technology, Volume 3, Issue 2,2012

7.M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform,” IEEE Trans. Image Processing, Volume 1, Pages 205–220, April1992

8.Reman N. R, Binu Ani Oommen, Mythili P, “Image compression using SPIHT with modified spatial orientation trees”, International Conference on Information and Communication Technologies, ScienceDirect,2014

9.Venkata Anjaneyulu, P. Rama Krishna, “FPGA Implementation of DWT-SPIHT Algorithm for Image Compression”, - International Journal of Technology Enhancements and Emerging Engineering Research – Vol 2, Issue 3,2014.

SECURED AND ENERGY EFFICIENCY ORIENTED CLUSTER ROUTING ALGORITHM FOR FANET BASED INTERNET OF THINGS

*Ms.Revathi B¹, Mr.Arulanandam K²
Department of Computer Applications,
Govt, Thirumagal Mills College, TN, India.*

Abstract

Internet of Things (IoT) is presently a hot technology for intelligent communication between resources. IoT enables digital interaction between the physical stuff using communication and heterogeneous networks. IoT is the common platform for various sensor devices for emitting data. IoT allows objects to be sensed or controlled remotely across existing network infrastructure, creating opportunities for more direct integration of the physical world into computer-based systems and resulting in improved efficiency, accuracy and economic benefit in addition to reduced human intervention for collecting and sharing physical devices in modern days. Flying Ad - hoc Networks (FANETs) with IoT is the technological phenomenon for communicating worldwide frameworks through the internet. This advanced combination facilitates communication with greater mobility and minimum costs. But achieving energy efficiency is still an open challenge in the networking environment. In this paper, we proposed FIOTCR (FANETs-IOT Clustering Routing Algorithm) for achieving energy-efficient utilization. To prove the efficiency of the proposed system, comparison work is carried with FANETs-WSN [11]. The obtained result shows the performance dominance of the proposed FIOTCR in terms of energy consumption and packet delivery ratio than the FANETs-WSN.

Keywords - IoT, FANETs, Cluster & Energy-efficient routing.

Introduction

IoT attracts the researchers as well as the industrialists because of the enormous applications which are very much useful to automate the industrial jobs as well as home environments and the challenges which enable the researchers to perform research on improvement and betterment of the protocols and algorithms. The different applications of IoT are categorized under consumer, commercial, industrial, and infrastructure. Some of the examples of the applications are like smart home, internet of medical things, and automation of manufacturing jobs, smart agriculture, environmental monitoring, and internet of battlefield

things for military purposes, etc. The IoT shows a lot of opportunities for directly incorporating the physical world into computer-based systems which leads towards a fully automated era.

In the digital world, invention on the internet enables devices had rapid growth. Internet connectivity allows reliable interfaces between the devices. IoT plays a vital role in establishing an interconnection between the global network infrastructures. K. Ashton coined IoT in 1999 [1], which is defines as the future internet. The significant feature of IoT is connecting the number of electronic devices through Internet connectivity. IoT uses communication technologies for identifying and providing communication between the physical or virtual things within time [2]. The significant feature of IoT is incorporating the interconnection between small objects like sensors, etc. Wireless Sensor Network (WSN) plays a vital role in IoT system, especially in data sensing, data collection, data connectivity, and data processing. The advancement of WSN was put forth for the evolution of the Flying ad-hoc network (FANET) [3]; FANET is a multi-hop network that is wireless and self-configuring. FANET connects the number of nodes through the router. The combination of FANET and IOT paves the way for intelligent and reliable technology.

The main reason for the successful combination of the FANET-IoT systems is energy balancing. IoT travels through multiple wireless sensors, which increases energy consumption. FANET protocols minimize energy consumption by choosing short and efficient routes in the network. There are several researchers involved in discovering energy-efficient routing protocols. The primary key for achieving minimum energy consumption is efficient routing protocols, maintaining the multi-hop wireless network connectivity for prolonging the network's lifetime. In FANET-IoT networks, the wireless network protocols cannot apply directly due to human interface with nodes, computational speed, and network node density.

In this paper, a cluster agent-based energy-efficient routing algorithm is proposed. Clustering is the concept of a grouping that enables routing energy-efficient and straightforward. The clustering mechanism with FANET-IoT increases the network lifetime than the other technologies. The main thing in our proposed FIOTCR is the selection of cluster agents. There are several nodes in the cluster, and each node is capable of selecting as a cluster agent. The only criteria for a cluster agent are the node with a higher energy level. Because the entire network communication is done under the supervision of a cluster agent, it decides if any nodes are limited to the transaction, which is the next node to continue the communication. FIOTCR performance is determined in NS-2.

This paper is organized as follows: Section 1 describes the introduction, section 2 describes the literature review, section 3 illustrates the structure and workflow of the proposed system, section 4 describes the result and discussion, and finally section 5 describes the conclusion of this work.

Literature Review

FANET

A Flying Ad-hoc Networks (FANETs) is such kind of network that consists of a group of small UAVs connected in ad-hoc manner, which are integrated into a team to achieve high level goals. Mobility, lack of central control, self-organizing and ad-hoc nature between the UAVs are the main features of FANETs, which could expand the connectivity and extend the communication range at infrastructure-less area. On one hand, in case of catastrophic situations when ordinary communication infrastructure is not available, FANETs can be used to provide a rapidly deployable, flexible, self-configurable and relatively small operating expenses network; the other hand connecting multiple UAVs in ad-hoc network is a big challenge. The ad-hoc sensor network contains a large structure of operations, and it does not require and centralized controller; it communicates directly [4]. This level of coordination requires appropriate communication architecture and routing protocols that can be set up on highly dynamic flying nodes in order to establish a reliable, efficient and robust communication. Routing protocols for FANETs are described below

Topology - based routing

Topology based routing protocols are

Proactive routing protocols

Reactive routing protocols

Hybrid routing protocols

Location - based routing

Location-based routing uses the node's actual position for making routing decisions.

They are

Global Positioning System (GPS)

Location - Aided Routing (LAR) protocol

Internet of Things (IoT)

The Internet of Things (IoT) is the network of physical devices, vehicles, home appliances and other items embedded with electronics, software, sensors, actuators and network connectivity which enables these objects to get connected and exchange data. Each thing is uniquely identifiable through its embedded computing system but is able to inter-operate within the existing Internet infrastructure [5].

The Interaction between Internet of Thing and MANET

The sensor revolution increases remote monitoring usage in various domains such as medical, industries, etc. Wireless Sensor Networks (WSN) are significant sources of remote monitoring as they do not require any wireless medium. In WSN, several sensors are connected for acquiring the node's sensed data and transmission. WSN is a centralized network [6], enabling communication to a wide range through IoT systems. IoT performance is depends on high power consumption and scalability [7]. For improving the efficiency of WSN, routing protocol plays an important role. But still, challenges like path elimination, dead sensor are existing in FANET, which had a greater influence on the Quality of Service (QoS) [8].

Cluster-based routing protocols

Alamerie et al. [9] proposed combined FANET protocols and WSN routing principles for data transmission. The proposed system includes dynamical monitoring with a predefined threshold value. The vital component of the proposed approach is the cluster head, and the cluster head selection is based on obtaining comparative results on the node's there hold value.

Yalda Akbari & Shayesteh Tabatabaei et al. [10] proposed a novel fuzzy logic and reinforcement learning for achieving high reliable routing system. The proposed system working procedure is based on the node's available bandwidth, node's energy in routes, and distance to the sink nodes.

The proposed system shows better performance in network lifetime and power consumption. Li, G., Zhang et al. [11] proposed a multicast routing algorithm for achieving quality of service. Routing links are estimated based on the collected aggregation information. The deterministic algorithms for finding a multi-thread tree are not required.

Shah, J., & Mishra, B et al. [12] proposed a content-aware method for clustering sensor nodes. The proposed system is based on the Ant Clustering algorithm. The clustering sensor nodes are created as semantic sensor networks (SSON). Next, the proposed ANTCLUST is

applied for building the cluster based on the content of the nodes. This approach minimizes the node searching process and improves network performance.

Qiu et al. Proposed the delay iterative method (DIM) to achieve real-time IoT response. The proposed system applies Delay Estimates for avoiding the neglecting of valid routes. This new routing method decision based on the information has enhanced the security level on transmission. Additionally, for load balancing in the network REPC method is implemented.

1. Proposed System

A. FIOTCR (FANETs-IoT CLUSTERING ROUTING ALGORITHM)

FIOTCR is a cluster-based routing algorithm, and cluster formation is very effective in identifying the shortest path in the network. The proposed system and its working mechanism is described below

The proposed mechanism begins with node deployment. The node deployment model comprises of implementing nodes within a range. The deployed nodes have different energy levels, such as low and high. Initially, the network is done among the deployed nodes. For this message, the broadcast is done, by which each node broadcasts a "Hello" message to its neighbour nodes. The replying nodes are considered active nodes, and only the active nodes take part in transmission. Based on the active nodes, cluster formation is done, and next is the cluster agent selection process. The node within the cluster network with a high energy level is selected as the cluster agent (CA). CA is the central core of the proposed system, and hence the entire network transmission is done under its supervision. Next, CA does another message broadcast to all its neighbour nodes. Based on the reply, each node's distance and its energy levels are noted. The nodes with the high energy level are selected as Cluster Head (CH) CH1, and the node with the next energy level is the CH2. Based on the energy level and its distance, a routing path is discovered. CA is monitoring the CH1 and other participating nodes in the transmission. CH 1 decides the next node to take part in the path. If the CH1 energy level is drained, then CH2 becomes the next CH, and this process cycle through the transmission until it completes.

Steps for implementing the proposed system are stated below:

Step 1: Implementing network model (Node Deployment)

Step 2: Executing CAMIOTCR for selecting Cluster Agent (CA)

Step 3: Nodes in the network are analyzed, and finally, the node with a high energy level is selected as CA.

Step 4: CA progression begins

Step 5: CA broadcast the "Hello" message to all the neighbour nodes and track its energy level.

Step 6: Next, CA executes the CH selection process. The node with a high energy level is selected as CH.

Step 7: The Selected CH begins its transmission process with its neighbouring nodes in the network.

Step 8: The entire network transmission is under the supervision of CA. It looks at the responsibility of monitoring the CH energy level.

Step 9: If the current CH energy level drains, CA allocate the next node with a higher energy level as CH. This process continues till transmission completes.

Step 10: Network Performance Evaluation.

Results & Discussion

In this research, the performance of the proposed work is executed in the simulation environment NS-2. Network simulator – 2 is the best-known platform for research execution. The experimental setup consists of 50 nodes between the ranges of 1000 x 1000. The performance comparison is made between the proposed FIOTCR and FANET-WSN [11]. The performance metrics taken for comparison are energy consumption and packet delivery ratio. The execution time is calculated in sec. The observation obtains from both algorithms is plotted in graphical representation for deciding the performances. The other parameters used in the experimental setup are described in below Table 1.

Parameter	Value
Number of nodes	50
Simulation environment	1000*1000
Radio transfer range	250 m
Packet size	1024 bit
Send type	Constant
Simulation time	200 s
Mac layer	IEEE 802.15.4
primary energy value	200-450 Jul

Table 1: Experimental parameter values

Above Fig 2 illustrates the energy consumption achieved by each algorithm. The energy consumption is calculated initially with 10 nodes at each stage; the input nodes are increased in the count of 10 and the observations are noted. The X-axis describes the number of nodes inputted in the experiments, and Y-axis represents the energy consumed by each set of nodes, respectively. The proposed FIOTCR energy consumed for 10 nodes is 7500 sec, 20 nodes are 13000 sec, 30 nodes is 16000 sec, 40 nodes are 22000 sec, and 50 nodes is 25000 sec. Whereas FANET-WSN energy consumed for 10 nodes is 15000 sec, 20 nodes are 18000 sec, 30 nodes are 23000 sec, 40 nodes is 25000 sec, and 50 nodes is 26000 sec. This graphical representation clearly shows energy consumed by the proposed FIOTCR is minimum than the FANET-WSN. Thus proves the proposed system performance efficiency in the term of energy consumptions.

Above fig 3 illustrates the Packet Delivery Ratio (PDR) achieved by each algorithm. PDR determines the ratio of successful packet delivery in comparison to the total packets transmitted within the time. The PDR is calculated initially with 10 nodes at each stage; the input nodes are increased in the count of 10, and the observations are noted. The X-axis describes the number of nodes inputted in the experiments, and the Y-axis describes the PDR achieved by each set of nodes. The proposed FIOTCR achieved PDR for 10 nodes is 25%, 20 nodes are 18%, 30 nodes are 75%, 40 nodes are 90%, and 50 nodes are 94%. Whereas FANET-WSN PDR for 10 nodes are 22%, 20 nodes are 42%, 30 nodes are 56%, 40 nodes are 70% and 50 nodes are 68%. This graphical representation clearly shows that PDR achieved by the proposed FIOTCR is far better than the FANET-WSN.

Conclusion

In this paper, FIOTCR (FANET-IoT Clustering Routing Algorithm) is proposed for attaining a better data routing path. The main motto of this paper is to enhance the energy performance of IoT with FANET. For this cluster-based network model is introduced, which makes the network transmission free from the complex process. To prove the performance of the proposed work, a comparison is made between the proposed FIOTCR with FANET-WSN in terms of energy consumption and packet delivery ratio. From the observation, it is clearly shown that the proposed FIOTCR is far better than the FANET-WSN. The hierarchical and chain selection process in CH achieves quick transmission with a highly successful packet delivery rate.

References

1. (IOT): A vision, architectural elements, and future directions, Future Generation Computer Systems, 29(7) (2013) 1645–1660.
2. ITU-T/Recommendation Y.2060, Overview of the Internet of Things, available at: <https://www.itu.int/rec/T-REC-Y.2060-201206-I>, 2013
3. Muhammad Asghar Khan, Alamgir Safi, Ijaz Mansoor Qureshi, Inam Ullah Khan, Flying ad-hoc networks (FANETs): A review of communication architectures, and routing protocols, IEEE Xplore, 2018.
4. Z. Hussain, R. Balakrishna, A survey on Fanets Types, Characteristics, Applications and Protocols used, National Conference on Frontiers and Advances in Information Science and Technology, 2013.
5. Dr. Jeeva Jose, Internet of Things, Khana publishing, 2018.
6. P. Bellavista, G. Cardone, A. Corradi, L. Foschini, Convergence of MANET and WSN in IoT Urban Scenarios, IEEE Sensors Journal, 13(10) (2013) 3558-3567.
7. G. S. Yovanof, G. N. Hazapis, An Architectural Framework and Enabling Wireless Technologies for Digital Cities & Intelligent Urban Environments, Wireless Personal Communications, 49(3m) (2009) 445-463.
8. N. Bessis, F. Xhafa, D. Varvarigou, R. Hill, M. Li, (eds), Internet of Things and Intercooperative Computational Technologies for Collective Intelligence, Springer Verlag, (2013).
9. S. I. A. Alameri, MANETS and Internet of Things: The Development of a Data Routing Algorithm, Engineering, Technology & Applied Science Research, 8(1), 2018, 2604

MOTIVATING FACTORS OF WOMEN ENTREPRENEUR IN CHENNAI CITY

*Dr.R.Dharmaragini¹, Dr.K.Rajamannar², R.Subitha rani³
Dept of Commerce & Research Centre,
Annai Veliankanni College for Women,Tamilnadu India*

INTRODUCTION

Entrepreneurship is a business-related field in which leaders are not afraid to take chances and implement unique strategies to improve their company's success and maximize profits. Business opportunities are the focus of entrepreneurship, business plan formulation, exploitation and monitoring, start-up formation, management, team-building, and market strategies, leadership, Human resource management, creativity, intellectual property, the economy, networking, and market analysis are only a few examples.

Women Entrepreneurship is concerned with the position of women in society also the role of women entrepreneurs in that society. Women have motivated by the many factors it consist of Experience in the field, Passion for business, Hobby, Social Prestige etc. The Success factors of Women Entrepreneur have Self motivation, Human relation, hard work, Spousal support, Eager to Achieve etc.

OBJECTIVES

1. To find out the factors motivating the women entrepreneurs and assisting them to remain successful.
2. To analyze the challenges faced by women entrepreneurs.
3. To find out Success factors of Women Entrepreneurs.

REVIEW OF LITERATURE

The purpose of this chapter is to briefly review selected streams of major academic research relevant to Motivating women and Success factors of the women entrepreneurs in Chennai City. Entrepreneurship has transformed many entrepreneurs into successful business persons and generated income in Chennai City.

Moore and Buttner (1997)¹ In their study titled "Women's Organizational Exodus to Entrepreneurship: Self-Reported Motivations and Correlates with Success," Buttner and Moore (1997) found that "Women's Organizational Exodus to Entrepreneurship: Self-Reported Motivations and Correlates with Success" Examining why 129 female executives and professionals left large corporations to become entrepreneurs, as well as how they calculate success, self-determination, shows that they did so in search of a challenge, and a way to balance work and family obligations.

Vatharkar (2012)² In his study, he looked into the issues that women entrepreneurs face at various levels in Pune, as well as the factors that encourage them to start their self businesses.

Saikia, (2017)³, the numerous motivating factors influencing women to start their own business, such as educational qualifications, self-worth, economic needs, independence, and so on As a result, it is said that many motivating factors forced women into starting their businesses.

METHODOLOGY

The study concentrated on primary data. The researcher collected the secondary data from Journals, Articles, Books, Websites, etc.

FACTORS MOTIVATING WOMEN ENTREPRENEURS

Motivation factors are vital for the Women Entrepreneurs in keeping them stay in their entrepreneurship activity. Here in this study experience in the field, passion for business, means of survival, venture something new, family business, hobby, social prestige, high income in business, not interested to be an employee, government incentives and concessions, innovative ideas, friends and relative in the field are taken into consideration.

Mean Ranking of the factors motivating Women Entrepreneurs

A questionnaire with a five-point scale was prepared and given to the respondents and they were asked to select in the respective column with response ranging from “strongly agree” to “strongly disagree”. A weighted average, also known as a weighted mean, is a little more complicated to figure out than regular arithmetic mean. As the name suggests, a weighted average is one where the different numbers you’re working with have different values, or weights, relative to each other. The mean value in table below reflects the importance of each observation and is more descriptive.

Mean Ranking of the factors motivating Women Entrepreneurs

S.N O	Motivational factor	Mean Rank	Mean	SA	A	N	DA	SD A
1	Experience in the Field	1	4.29	150	118	46	4	1
2	Passion for business	2	4.08	98	161	48	10	2
3	Means of Survival	5	3.85	91	106	10 7	12	3
4	Venture something New	8	3.71	62	144	76	31	6

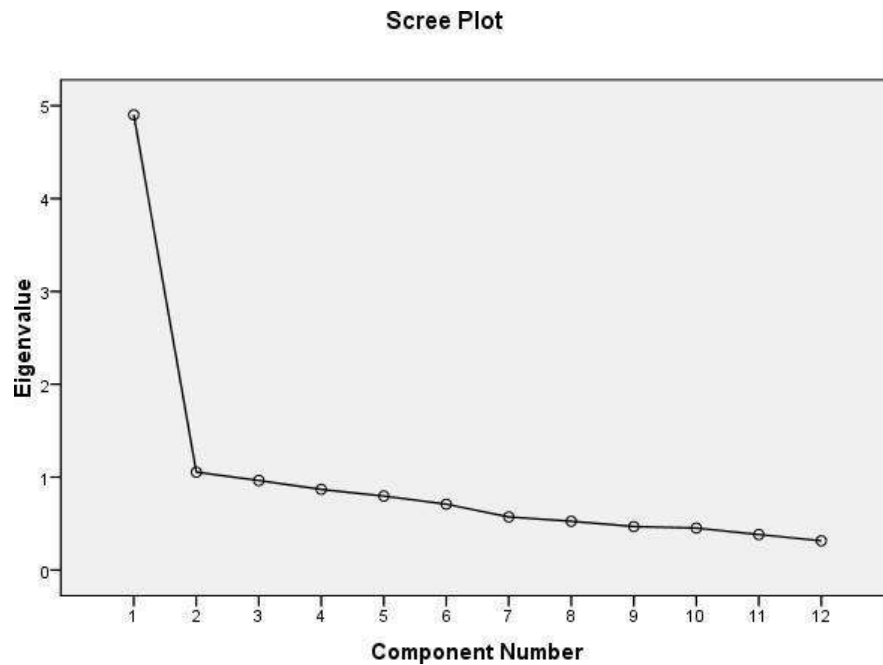
5	Family Business	12	3.56	87	89	85	32	26
6	Hobby	4	3.87	85	143	64	19	8
7	Social Prestige	6	3.8	89	117	80	26	7
8	High income in Business	10	3.65	67	117	94	37	4
9	Not interested to be an employee	7	3.77	94	109	75	30	11
10	Government Incentives and Concessions	13	3.5	60	118	78	46	17
11	Innovative Ideas	3	3.95	100	123	80	13	3
12	Friends and relative in the field	9	3.71	77	129	74	23	16
13	Other Reasons	11	3.65	74	110	94	30	11

Source: compiled from primary data

SA - Strongly Agree, A - Agree, N-Neutral, DA- Disagree, SDA-Strongly Disagree

It is inferred from the above table that factors experience in the field, passion for business, innovative ideas, hobby, means of survival are the most influential factors of motivating women entrepreneurs. Social prestige, not interested to be an employee, venture something new, friends and relative, in the field are the next influencing factors then comes the family business, high income in business, government incentives and concessions and other motivating factors. It can be concluded that women entrepreneurs are more motivated towards their passion rather than government incentives and concessions.

The scree plot graphs the eigenvalue against the factor number. In the scree plot components are plotted in the X-axis and the corresponding eigenvalue at the Y-axis. The below graph exhibits the eigenvalue against the motivating factors of women entrepreneurs.



Graph Screen plot on the motivating factors of women entrepreneurs

We can see these values in table immediately above. From this third factor on, we can see that the line is almost flat, meaning each successive factor is accounting for smaller and smaller amounts of the total variance. The first two components are considered as their eigenvalues are greater than 1 and sharing maximum variance hence they are essential in the present study.

The rotated factor loadings, which show both how the variables are weighted for each factor and the correlation between the variables and the factor, are found in Rotated Component Matrix table. Because they are correlations, the range of possible values is -1 to +1. We used the options blank (.30) on the /format subcommand, which tells SPSS not to print any correlations with a value of .4 or less. This makes the output easier to read by reducing the clutter of low correlations that are most likely unimportant.

Conclusion

In many nations throughout the world, entrepreneurship is critical to economic growth. Women entrepreneurs have always been there and unpredictably shaped the economy. Women entrepreneurs' greater performance is mostly due to their motivation. Several driving elements enable Chennai women to become entrepreneurs and succeed in business. We believe women are motivated by various reasons to step into the entrepreneur world.

The results of the mean ranking exhibit that experience in their field (i.e) in the field of doing business, is the major motivating factor to be successful as it gives them more knowledge and creativity in their business. Passion for doing business is also a major motivating factor. Entrepreneurship gives a lot of women the freedom to work on what they love to do. Some women are into entrepreneurship to start something with what they love to do. In the highly competitive world that we live in, innovative ideas are what will separate women entrepreneurs from the rest hence it is the next factor that attracts women entrepreneurs.

Many hobbies have the potential to provide benefits that can improve the level of success as a women entrepreneur. The term "means of survival" is used to characterize a firm that is always struggling to stay afloat. It is also a powerful motivator for women entrepreneurs. . A person's power can be shown in the social order with their status, in the business. They believe that the person's social prestige, class position, and means of doing business also motivate women entrepreneurs in doing their business.

Not interested to be an employee, Venture something new, Friends and relative in the field, High income in Business, Family Business and Government Incentives and Concessions are the minimal motivational factors for the women entrepreneurs on being successful.

References:

1. Buttner, E. and Moore, D. (1997). Women's Organizational exodus to Entrepreneurship: Self-Reported Motivations and Correlates with Success. *Journal of Small Business Management* 35(1)(pp.34to47).
2. Krishnaveni Motha. (2004). Women Entrepreneurship in Rural Areas of India. *Kurukshetra*, 15(1),11-14.
3. Poonam Vatharkar, (2012). A Study of Constraints and Motivating Factors for Indian Women Entrepreneurs in Small Scale IndustriesIntentional Conference on Business and Management, Phuket – Thailand, pp 473-488.
4. Sreenivasa Behara And Niranjan, (2012), Rural Women Entrepreneurship in India IJCEM International Journal of Computational Engineering & Management, Vol. 15 Issue 6, November 2012 ISSN (Online):2230-7893.
5. Balhara, Shruti & Singh, Ajmer, (2015), ‘Women Entrepreneurship: A Big Motivation’. *Business and Economic Research*, vol.5, no.2.4.
6. Jayakumar And Vincentsahayaraj (2015). “ Inclusivegrowth Of Thiruchirapalli District Through Farm Sectors Industries”*International Journal Of Accounting And Financing.Management Research (Ijafmr)* Issn (P):2249-6882 Vol7.

USING A MODIFIED MINIMUM SPANNING TREE TO IMPROVE CLUSTER BASED FEATURE SELECTION

Dr.A.Suresh¹

*Assistant Professor and
Head Department of
Computer Science
Tamilnadu india*

Mr. A.Kaleemullah²

*Research Scholar
Department of
ComputerScience,
Tamilnadu india*

Dr.P.Rizwan Ahmed

*PostDoc Researcher
Dept. of Computer Applications
Tamilnadu india*

ABSTRACT

The issue of feature subset selection for data of high dimensionality in classifying opinions has been investigated in this paper. Feature selection schemes select the crucial feature subset producing similar or superior classification outcomes compared to the original feature set obtained. Despite their higher efficiencies, wrapper-based feature selection schemes have high overheads of computation. However, the issue is the problem is Non-deterministic Polynomial (NP) hard. This work suggests the clustering scheme based on MST optimized by the Group Search Optimization (GSO) for efficient selection of features. The Amazon camera review dataset is used for evaluating the suggested scheme. Experiments are conducted to evaluate the proposed method and compared with the MRMR feature selection, FCM clustering, and MST based clustering.

Keywords: *Sentiment analysis, Feature Selection, Minimum Spanning Tree (MST), Data Analytics, Clustering and Group Search Optimization (GSO)*

1INTRODUCTION

Feature selection in machine learning is also referred to as changeable subset selection. This refers to the method of choosing a subset of pertinent features for building a prototype. When correlated models are being built, the following advantages are exhibited by the feature selection techniques: Interpreting the implicit meaning using the improvised prototype. Smaller duration of training and increased generalization by over fitting are other benefits. Feature selection detects crucial features for forecasting and aids in the data analysis process. When some of the good features are selected as per the target concepts, feature selection can effectively decrease dimensionality, eliminate redundant data, enhance the precision of learning and improvise the understanding ability. There are many irrelevant features as well as repeated ones contained in data of high dimensionality. While the repeated features do not provide any more useful information than the chosen ones, the irrelevant features do not provide, under any situation any useful information [1].

For training the opinion mining system, there are several techniques used by the researchers. Obtaining an effective feature set in most machine language applications is the crucial step in sentiment classification. There are many applications in which there is a huge use of sentiment analysis. There are many viable tasks it can perform including detecting the attitude or the opinion of the customers regarding the product or the services. The product feedback and the service reviews will help make reasonable decisions. For instance, a new person visiting a city can be aided by the restaurant reviews for locating a good place. In the same manner, a movie can be decided as worthy of watching or not by referring to the movie reviews [2].

2RELATED WORKS

For high dimensional data, an innovative feature subset selection algorithm on the basis of clustering has been suggested by Kumari & Naidu [7]. The features of this algorithm are: (i) eliminating redundant features (ii) Building minimum spanning tree from comparative ones (iii) After portioning the MST, the representative features are selected. There are features comprised in a cluster. Every cluster is taken to be a single feature and there is a reduction in dimensionality. Five of the popular feature selection schemes exist with which the performance of the suggested algorithm can be compared. These include FCBF, ReliefF, CFS, Consist, and the FOCUS-SF on thirty-five openly accessible image microarray as well as text data from 4 various sections of the chosen features. For RIPPER, C4.5 and

Naïve Bayes, the suggested algorithm has the optimal selected features, runtime and accuracy of classification.

Different techniques for decreasing the dimensionality in high dimensional microarray cancer data have been presented by Hira & Gillies [8]. There is an increase in the quantity of data which is to be analyzed. This has made it essential to allow decrease in dimensions for obtaining pertinent outcomes. There are various feature selection as well as feature extraction schemes that have been described and compared. Along with them, their benefits and disadvantages are also contrasted. Additionally, this work has also presented various schemes for including the prior knowledge from different biological sources which is a technique of enhancing the precision and decreasing the computational complexity of the schemes that are already present.

A feature selection scheme has been designed by Kaveri& Asha [9]. This scheme makes use of the accurate relevance schemes. This work uses relevance measures “Symmetric Uncertainty (SU)” for effectively selecting the relevant attributes. The chosen attributes are then partitioned into clusters on the basis of “graph theoretic” clustering scheme by making use of a relevance measure referred to as “Conditional Mutual Information (CMI)”. To choose the attributes corresponding well to target class and also the one which represents the cluster in the best possible way, the relevance measure “symmetry uncertainty” has been selected. This leads to the accurate and independent subset of features. This technique helps to produce smaller as well as more precise subset of features improving the act of machine learning functions like the Naive Bayes classifier.

Pullela et al., [3] suggested clustering scheme having two stages for feature selection. The features are clustered in the first stage and in the second, for decision making and further use in data mining, the most representative features are extracted. This data of high dimensions is used as input and for formulation of clusters, clustering has been performed. There are many features in every cluster. The selected subset of features is the best features that are representative of features. The input used is the high dimensional data. For the formation of clusters, clustering process is employed. There are several features in each cluster and the subset of features selected contains the best features that are representative of the features chosen. This has several utilities that can be used in real world applications. These chosen features help find newer avenues scope for further data mining purposes. The experimental outcomes have been promising. A significant direction for future work is the construction of a prototype comprising generalized features that can help adapt to different high dimensional datasets for various domains.

A feature subset selection algorithm for huge dimensional data which has been based on clustering has been suggested by Elankavi et al., [4]. These are the features of the algorithm- 1) eliminating irrelevant features. 2) Using relative trees for construction of minimum spanning tree and 3) dividing the MST and choosing delegate features. There are features contained in the selected algorithm. After treating each cluster as a single feature, there is a drastic decreased of dimensionality. The best proportion of the selected features has been obtained from the suggested algorithm, including the optimal runtime, the optimal classification precision for Naïve Bayes, C4.5, RIPPER including the second best classification precision for IB1.

3METHODOLOGY

Extracting opinion of the users regarding a product or a service from the review documents is referred to as sentiment classification. When machine learning techniques are used in sentiment classification, the feature vectors have high dimensionality problems. Hence there is a need for the feature selection scheme to remove the unrelated and loud features from the feature vector for the effective functioning of the machine learning algorithms. In this section, the MRMR feature selection, FCM clustering, MST based clustering and GSO optimized MST clustering methods are discussed.

3.1 Minimum Redundancy Maximum Relevance (MRMR) Feature Selection

The discriminate features of a class can be recognized using the MRMR feature selection. Features having the maximum relevancy/high dependency to the class and minimum redundancy / minimal dependence among the features are selected by the MRMR techniques. At times there is a lot of redundancy among the features for the pertinent features with maximum relevancy with the class. Given two features that are redundant, eliminating one of the features will not cause much difference in the discrimination of class. The correlation or the dependence among the features and the class attributes and amongst the features can be calculated using the mutual information [12]. Features that comprise enhanced mutual information/maximum relevance with class attribute are selected by the mRMR technique while the ones having high mutual information and high correlation among themselves/ minimum redundancy are eliminated.

Some of the benefits of mRMR are: estimation of both relevance and redundancy are problems of lower dimensions which have only about two variables. Thus, it is easier when compared to direct estimation of multivariate density or mutual information in huge dimensional area. This estimation is not only quicker but also more consistent. MRMR is an optimal first order approximation of $I(.)$, maximization, and relevance-only ranking only maximizes $J(.)$.

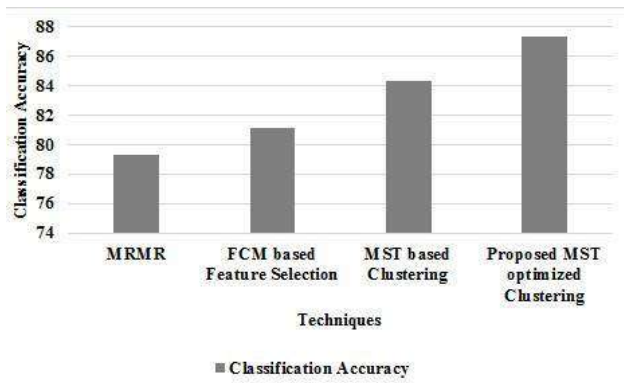
Minimum Spanning Tree (MST) Based Clustering

MST is employed for generating highly complex clusters and is a graph based prototype. The MST edges can be either chosen or discarded by it. A spanning tree is an acyclic sub-graph of graph G comprising all of G 's vertices. The least weight spanning tree of a graph is the MST of that weighted graph. The expense to construct a minimum spanning tree in conventional MST is $O(m \log n)$. Here, n is representative of vertices and m is representative of edges. Regarded as a hierarchical clustering algorithm, MST can follow divisive clustering scheme.

The divisive clustering scheme is tracked by MST clustering which is a hierarchical clustering algorithm. For constructing MST of point set and for deleting conflicting edges, clustering algorithm on the basis of minimum and maximum spanning tree have been examined. Their weights are expansively greater compared to the standard weight of the closest edges of the tree. The objective is maximizing the minimum inter cluster distance.

RESULTS AND DISCUSSION

Hence, the MRMR, FCM based feature selection, MST based clustering and proposed MST optimized clustering methods are used. Table 1 shows the summary of results. The classification accuracy, positive predictive value for positive, neutral and negative opinion and hitrate for positive, neutral and negative opinion as shown in figures 1.

Figure 1 Classification Accuracy for Proposed MST Optimized Clustering

	MRMR	FCM based Feature Selection	MST based Clustering	Proposed MST optimized Clustering
Classification Accuracy	79.26	81.11	84.33	87.37
Positive predictive value for Positive Opinion	0.8049	0.8385	0.8441	0.8834
Positive predictive value for Neutral Opinion	0.7989	0.8056	0.8352	0.8606
Positive predictive value for Negative Opinion	0.7766	0.794	0.8509	0.8777
Hitrate for Positive Opinion	0.7333	0.75	0.8422	0.8422
Hitrate for Neutral Opinion	0.7944	0.8056	0.8444	0.8778
Hitrate for Negative opinion	0.85	0.8778	0.8433	0.9011

From the figure 1, it can be observed that the proposed MST optimized clustering has higher classification accuracy by 9.73% for MRMR, by 7.43% for FCM based feature selection and by 3.54% for MST based clustering.

5CONCLUSION

In the classification problems that involve high dimensional data, there are many features. However, not all of them are suitable for classification. Better classification performance can be achieved using filter selection which involves selecting small number of related features. The presence of unrelated, noisy and redundant features may degrade the performance. An important objective of feature selection is increasing the classification performance and minimizing the number of features. The input space can be segregated into decision regions using fuzzy entropy which can select appropriate features with good partitioning for classification. This work suggests the GSO based feature selection optimized MST clustering algorithm. Clusters with irregular edges can be detected using the MST clustering algorithms. GSO is motivated by the searching behaviour and the group living theory of the animals. PS model forms the basis of the algorithm in which it is assumed that the group members 'finding' are producers or the ones that are 'joining' are scroungers. The algorithm is developed using the animal scanning scheme like vision. Also, for evading the local minima and for performing random walks, rangers can be employed. Results show that the proposed MST optimized clustering has higher classification accuracy by 9.73% for MRMR, by 7.43% for FCM based feature selection and by 3.54% for MST based clustering.

REFERENCES

1. Khedkar, S. A., Bainwad, A. M., & Chitnis, P. O. (2014). A survey on clustered feature selection algorithms for high dimensional data. *Int J ComputSciInfTechnol (IJCSIT)*, 5(3), 3274-3280.
2. Manek, A. S., Shenoy, P. D., Mohan, M. C., & Venugopal, K. R. (2017). Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier. *World wide web*, 20(2), 135-154.
3. Pullela, V., Kumar, V. S., & Yadav, C. R. (2014). A Framework for Mining High Dimensional Data for Feature Subset Selection. *International Journal of Computer Science and Mobile Computing*, 3 (12), 50-55.
4. Elankavi, R., Kalaiprasath, R., & Udayakumar, D. R. (2017). A fast clustering algorithm for high-dimensional data. *International Journal Of Civil Engineering And Technology (Ijciet)*, 8(5), 1220-1227.
5. Song, Q., Ni, J., & Wang, G. (2013). A fast clustering-based feature subset selection algorithm for high-dimensional data. *IEEE transactions on knowledge and data engineering*, 25(1), 1-14.
6. Torkestani, J. A., & Meybodi, M. R. (2012). A learning automata-based heuristic algorithm for solving the minimum spanning tree problem in stochastic graphs. *The Journal of Supercomputing*, 59(2), 1035-1054.
7. Kumari, B. S., & Naidu, M. D. (2014). Feature Subset Selection Algorithm for Elevated Dimensional Data By using Fast Cluster. *International Journal Of Engineering And Computer Science*, 3(07).
8. Hira, Z. M., & Gillies, D. F. (2015). A review of feature selection and feature extraction methods applied on microarray data. *Advances in bioinformatics*, 2015.
9. Kaveri, B. V., & Asha, T. (2015). An Ameliorated Methodology for Feature Subset Selection on High Dimensional Data using Precise Relevance Measures. *International Journal of Computer Applications*, 127 (7).
10. Singh, A. (2009). An artificial bee colony algorithm for the leaf-constrained minimum spanning tree problem. *Applied Soft Computing*, 9(2), 625-631.
11. Gu, S., Cheng, R., & Jin, Y. (2018). Feature selection for high-dimensional classification using a competitive swarm optimizer. *Soft Computing*, 22(3), 811-822.
12. Elawady, R. M., Barakat, S., & Elrashidy, N. M. (2014). Different feature selection for sentiment classification. *International Journal of Information Science and Intelligent System*, 3(1), 137-150.

13. Manjula, S. P., & Jagadeesan, J. (2014). Fuzzy C Means Clustering Algorithm for High Dimensional Data Using Feature Subset Selection Technique. IOSR Journal of Computer Engineering (IOSR-JCE), 16 (2), 64-69.
14. Jeyarani, D. S., & Pethalakshmi, A. (2016). Optimized Feature Selection Algorithm for High Dimensional Data. Indian Journal of Science and Technology, 9(31).
15. Nagendrudu, S., & Reddy, V. R. (2015). Enhanced Clustering of High Dimensional Data Using Fast Cluster Based Feature Selection. International Journal of Science, Engineering and Computer Technology, 5(5), 113.
16. Alapati, Y. K., Sindhu, K., & Suneel, S. (2015). Relevant Feature Selection from High-Dimensional Data Using MST Based Clustering. International Journal of Emerging Trends in Science and Technology, 2(03).

NUTRITIONAL STATUS AND RISK OF TYPE2 DIABETES AMONG PRE-MENOPAUSAL WOMEN IN KOLAGUR

V. Shivani Abirami¹,
Assistant Professor, Dept of ND
Marudhar Kesari Jain College for Women, Tamil Nadu, India.

Abstract

India has one of the largest tribal populations in the world nearly, 7.5% of the population is classified into tribes. In India, Tamil Nadu state has a tribal population of 7.21 lakh as per census report in 2011. Salem is one of the districts of Tamil Nadu have the highest tribal population of 1.2 lakh, especially Kolagur- a small Village/hamlet in Yer caud Block in Salem District has a total population of 1371 people. Women play an important role in the community by preserving the culture and traditions of the country thus, nutritional status of women is an important determinant of the nutritional status of the child in society. Tribal women in Kolagur lack knowledge about balanced diet and nutrients present in their local foods, hence they suffer from deficiency diseases and non-communicable diseases such as hypertension, diabetes and thyroid dysfunction are highly prevalent among the women in kolagur. The study aims to assess the nutritional status of pre-menopausal women using an exploratory study design and simple random sampling technique. The mean anthropometric parameters and the mean biochemical parameters such as blood haemoglobin level, packed cell volume, fasting blood glucose and serum lipid profile were found to be normal. A few subjects suffered from hypothyroidism due to lack of iodized salt in their diet. The cultivation, consumption of millets and a balanced diet can be encouraged by the repeated nutrition education program.

Keywords: Tribes, Nutrition, Premenopausal, Type2 diabetes, Vitamin, Calcium

1. Introduction

In India, tribes are called as “Adivasi” (‘Adi’ means oldest and ‘vasi’ means inhabitant) meaning indigenous people or original inhabitants of the country. India has one of the largest tribal populations in the world nearly, 7.5% of the population is classified into tribes. In India, Tamil Nadu state has a tribal population of 7.21 lakh as per census report in 2011. Salem is one of the districts of Tamil Nadu have the highest tribal population of 1.2 lakh, especially Kolagur. In Salem, Kolagur is a small Village/hamlet in Yercaud Block, comes under the Nagalur Panchayath and is situated in North 24 kilometers away in the direction of Salem District Headquarters has a total population of 1371 people with local language as *Tamil*. The common occupation of the people of Kolagur village is agriculture mainly as laborers in coffee estate and tea plantation. Nearly 41% of the tribal women of Kolagur village belong to lower income group.

The people of the tribal community are unable to maintain good health because of lack of knowledge in nutrition, they take nutrition synonymous with food and as a means of satiation for their hunger. Only 41.7 % of the women have toilet facilities at home while rest of the population lack proper sanitation. Poor sanitation is another major reason for the tribal community to easily get trapped by diseases like malaria, measles, tetanus, whooping cough, gastroenteritis, and other major diseases. The changing environment, lifestyle and rapid growth in urbanization are the reasons for the occurrence of many other major diseases such as hypertension, cancer and diabetes. On the other hand, there may be the risks of nutritional deficiency diseases like anemia, hypothyroidism, hyperthyroidism, hypercholesterolemia and many others.

The tribal people of Kolagur village have good access to pulses, rice and oil which are distributed to them through Public Distribution System (PDS). In the local market of Kolagur, lots of locally cultivated vegetables and fruits are available but the fruit intake of the tribal people is less. Few years back, the staple food of Kolagur was millets but availability of surplus amount of rice through Public Distribution System made the people to consume rice as their staple food now. They perform high levels of physical activity such as walking 3 to 4 km per day for their work, schools, sometimes to Yercaud to visit doctors, buying groceries and apparels.

Risks associated with nutrition among tribal community women

There are many risks associated with the nutritional status of the foods in the tribal women’s community. The lack of daily nutrition needs has an impact on the wellbeing and health of the women thus reduces the ability to live an active and enjoyable life (Gupta et al. 2018).

Furthermore, poor nutrition intake can lead to many diseases such as diabetes, cancer, and many others. They require many health nutritional factors like carbohydrates, protein, fat, vitamins, calcium, magnesium, potassium, cobalt and many other elements daily (Mahajan et al. 2019). Due to malnutrition, the tribal women have the risk of anemia and malnutrition in pregnant women is a major problem of the tribal women in India. On the other hand, it also causes many major diseases like obesity, high cholesterol, heart disease, Type2 diabetes, osteoporosis, high blood pressure (BP) and many others (Ahirwar and Mondal, 2019).

2. Materials and Methods

This study was approved by the Institutional Ethics Committee, the study was an exploratory study design and a simple random sampling method was adopted for the study. Sixty-eight pre-menopausal women of Kolagur, Yercaud who were willing to participate were chosen for the study while Pregnant and Lactating Women were excluded from the study. A detailed and standardized interview schedule was prepared to obtain information on the socio-demographic profile and sanitary habits. The interview schedule was translated into the local language and a face-to-face interview with the subjects was done to obtain the information. The anthropometric parameters such as Height, Body Weight, Waist circumference, Hip circumference, Body fat were measured. Blood samples for the analysis of the selected biochemical parameters such as blood haemoglobin level, fasting blood glucose level, Serum lipid profile, thyroid function tests (T3, T4, TSH) were taken by qualified lab technicians from the internationally recognized laboratory **Lister Metropolis** in Salem. The clinical parameters such as blood pressure levels and pulse rate were assessed and recorded using sphygmomanometer. Food Frequency Questionnaire was prepared to assess the intake of the foods and 24-hour dietary recall was done for three nonconsecutive days (two weekdays and one weekend) to assess the food and nutrient intake. With the use of **DietCal** Software, the daily average intake of nutrients and energy was calculated which was developed by Gurdeep Kaur in the year 2015 and established for the statistical analysis. Indian Diabetes Risk Score was used to identify the high-risk subjects of Diabetes Mellitus (Arun, 2016).

3. Results and Discussion

Anthropometric Parameters of the Subjects

The anthropometric parameters include body weight, height, waist circumference, hip circumference and body fat percentage are listed in the following Table 1:

Table1: Mean Value of Anthropometric Parameters (n=68)

Parameters	Mean \pmSD
Height (cm)	152.53\pm5.97
Body Weight (kg)	48.97\pm10.44
Body Mass Index (kg/m²)	21.28\pm4.58
Waist Circumference (cm)	73.13\pm 18.5
Hip Circumference (cm)	83.22\pm17.77
Waist to Hip Ratio	0.86\pm0.20
Percentage Body fat	27.95\pm7.85

Table 1 reveals the mean values of anthropometric parameters such as body weight, height, waist circumference, hip circumference and body fat percentage of the subjects. The mean values of height and body weight of the subjects within normal limit. According to WHO's (2004), BMI classification, the mean Body Mass Index of the subjects were normal. The mean values of waist circumference, hip circumference and Waist to Hip Ratio were also within the normal range prescribed by WHO, (2008). The mean body fat percentage of the subjects was within the normal range according to the reference value of body fat percentage given by Gallagher et al., (2000).

The mean values of all the anthropometric parameters of the subjects are found to be normal which clearly reveals that an adult performing physical activity daily with the consumption of a diet high in vegetables, low in processed meat, red meat, fast foods, white bread, soft drinks, reduced fat dairy and whole grains have normal Body Mass Index, waist circumference and body fat percentage (Mathur et al., 2009; Romaguera et al., 2011). Tribal women in Kolagur are physically active, they walk for 3 to 4 km per day for their work, sometimes to Yercaud to visit doctors, buying groceries and apparels. They consume one serving of vegetable three times in a day and the consumption of processed meats, red meat, fast foods, white bread and soft drinks are less as they are rarely available which was the major reason for the normal mean Body Mass Index, waist circumference and body fat percentage.

Biochemical Parameters of the Subjects

The biochemical parameters include blood haemoglobin, Fasting blood glucose levels, serum lipid profile and thyroid function test of the pre-menopausal women in Kolagur were listed in the following Table 2:

Table2: Mean Values of Biochemical Parameters

Parameters	Reference Value	Mean \pm SD
Blood Haemoglobin Levels	11.0 – 12 g/dL	11.46 \pm 1.43
Packed Cell Volume	37%- 47%	37.64 \pm 3.34
Fasting Blood Glucose	<100 mg/dL	95.24 \pm 12.03
Total Cholesterol (mg/dL)	< 200 mg/dL	153.96 \pm 33.19
Triglycerides (mg/dL)	< 150 mg/dL	98.72 \pm 62.93
HDL-Cholesterol (mg/dL)	< 40 mg/dL	41.06 \pm 10.96
Non-HDL-Cholesterol (mg/dL)	< 130 mg/dL	112.90 \pm 30.51
LDL-Cholesterol (mg/dL)	< 100 mg/dL	92.46 \pm 30.38
VLDL-Cholesterol (mg/dL)	6-38 mg/dL	19.74 \pm 12.59
LDL/HDL Ratio (mg/dL)	2.5-3.5 mg/dL	2.38 \pm 0.92
CHOL/HDL Ratio (mg/dL)	3.5-5 mg/dL	3.94 \pm 1.06
T3	70-190 ng/dL	124.68 \pm 19.44
T4	5 - 12 μ g/dL	7.60 \pm 1.59
TSH	0.4- 4.2 μ IU/mL	2.95 \pm 2.62

The table 2 gives the mean values of biochemical parameters such as Blood Hemoglobin levels, packed cell volume, Fasting blood glucose, Total cholesterol, triglycerides, HDL- Cholesterol, Non -HDL - Cholesterol, LDL- Cholesterol, VLDL- Cholesterol, LDL/ HDL ratio, CHOL/HDL ratio, T3, T4 and TSH. The mean hemoglobin value and mean Packed Cell Volume were found to be in normal range according to WHO,(2011) and Estridge, (2000) respectively. The study reveals that high consumption of green leafy vegetables and eggs have positive relation with haemoglobin concentration (Kankanamge et al., 2017). Green leafy vegetables such as semanchi keera and manathakali leaves (kamanchi keera) are rich source of iron and consumed frequently in the form of gravies or kootu by the tribal women in Kolagur.

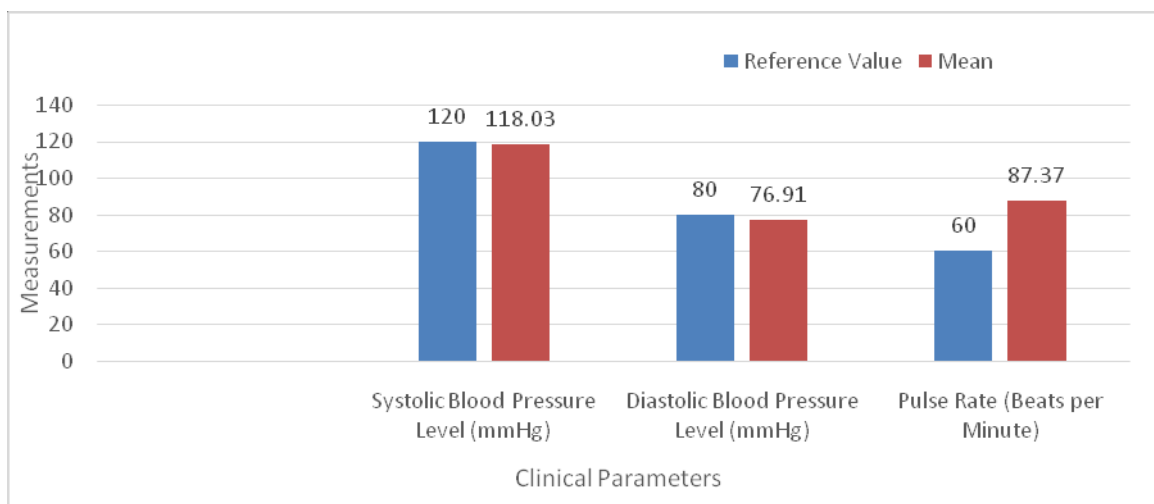
According to American Diabetes Association (2017), the mean fasting blood glucose of the subjects was found to be in normal range which reveals that most of the subjects are not diabetic. Generally, a diet rich in protein and good physical activity lowers the fasting blood glucose level (Gannon et al., 2003; Sakung et al., 2018). Tribal women in Kolagur consume high protein diet consisting of black bean in the form of gravy or kootu and as mentioned before they are physically active- walk for 3 to 4 km per day for their work and other daily needs.

With regard to serum lipid profile, the mean values of total cholesterol, triglyceride, HDL- Cholesterol, LDL Cholesterol, VLDL-Cholesterol, Cholesterol/HDL and LDL/HDL ratio of the subjects were also within the desired level. With regard to thyroid function test, the mean values of T3, T4 and TSH were also within the desired level. Studies reveal that high levels of physical activity, lower intake of foods rich in saturated fat and higher intake of monounsaturated fatty acid rich foods decrease the total cholesterol levels, triglycerides and increase the HDL-Cholesterol levels (Da Silva et al., 2016 ;Schwingshackl and Hoffmann , 2013). Tribal women in Kolagur consume less or never consume fat and fat rich foods such as beef, processed meat such as sausages, samosa, chips, palm oil and coconut oil since they are rarely available. Ground nut oil which is rich in monounsaturated fatty acid is used for cooking and high levels of physical activity is attributed to favourable serum lipid profile and thyroid function test.

Clinical Parameters of the Subjects

The clinical parameters include blood pressure levels and pulse rate of the pre-menopausal women in Kolagur are graphically represented in the following Figure 1:

Figure 1: Mean Clinical Parameters of the Subjects (n=68)



From the figure 1, the mean systolic blood pressure level (118 mmHg) and mean diastolic blood pressure level (76 mmHg) are within the normal range. The mean pulse rate is 87 beats per minute which are in the normal range given by AHA/ACC Hypertension Guidelines, 2017. Tribal women in Kolagur are living in high altitude for many years, consumption of vegetable three times a day which makes their diet rich in fiber and less intake of sodium rich foods such as processed foods and high fat dairy products such as cheese and butter as they are unavailable in Kolagur which made the blood pressure level of the subjects falls within normal range.

Dietary Assessment of the Subjects

The dietary parameters include Macronutrients and micronutrient intakes of the pre-menopausal women in Kolagur are listed in the Table 3.

With regard to macronutrients, the high energy intake in tribal women in Kolagur is due to the increased consumption of ca

Rbohydrates which is needed to meet the requirement of high levels of physical activity carried out every day. The high protein intake is due to consumption of pulse variety which is colloquially called Kottai every day in their diet. The low fat intake is due to the low intake of fried foods and processed foods, meats like mutton and chicken.

With regard to vitamins, the mean intake of Vitamin B1 is as per requirement due to frequent intake of green leafy vegetables. However, the mean intake of Vitamin B2, Vitamin B6, Vitamin B9 and Vitamin C is higher than the Recommended Dietary Allowance due to high consumption of green leafy vegetables. The mean intake of Vitamin A (Beta-Carotene) and Vitamin A (Retinol) is lower than the RDA due to low intake of foods rich in Vitamin A such as carrots, organ meats such as liver and nuts.

With regard to minerals, the mean intake of zinc and iron was as per RDA. The mean intake of Magnesium and Calcium was higher than RDA.

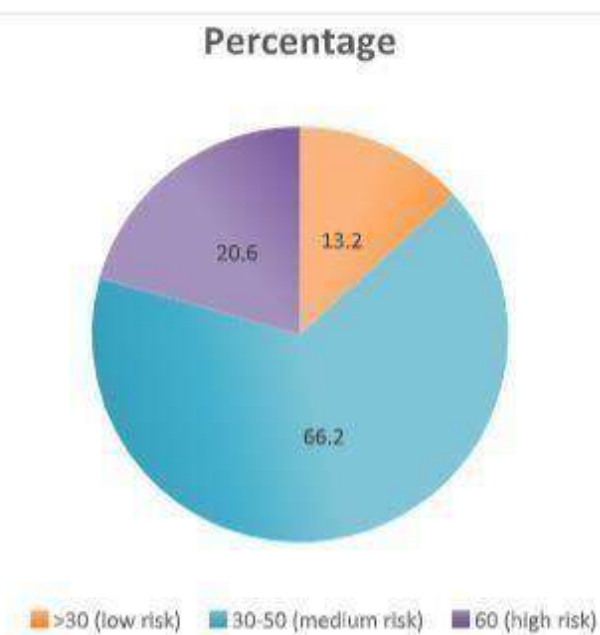
Table 3: Mean Nutrient Intake of the Subjects (n=68)

S.N O	Nutrient	RDA	Mean \pm SD	Mean Increase /Decrease
1.	Energy (Kcal)	2230	2560.79 \pm 599.31	14.8 \uparrow
2.	Carbohydrate (gm)		527.35 \pm 131.09	–
3.	Protein (gm)	55	77.81 \pm 71.45	40 \uparrow
4.	Fat (gm)	25	14.60 \pm 11.57	40 \downarrow
5.	Thiamine (mg)	1.1	1.50 \pm 0.46	36 \uparrow
6.	Riboflavin (mg)	1.3	0.63 \pm 0.29	–
7.	Pantothenic Acid (mg)		4.96 \pm 1.32	–
8.	Pyridoxine (mg)	2.0	1.83 \pm 0.54	8.5 \downarrow
9.	Biotin (mg)		8.86 \pm 6.63	
10.	Folic Acid (mg)	200	288.06 \pm 258.18	44 \uparrow
11.	Cobalamin (μ g)	1	2.25 \pm 1.75	125 \uparrow
12.	Ascorbic Acid (mg)	40	64.02 \pm 94.25	60 \downarrow
13.	Beta-Carotene (mg)	600	161 \pm 267	73 \downarrow
14.	Cholecalciferol (μ g)		0.53 \pm 0.53	
15.	Tocopherol (mg)		1.89 \pm 1.51	

16.	Phylloquinone (mg)		151.83 ± 213.90	
17.	Menaquinones (mg)		10.71 ± 0.00	
18.	Calcium (mg)	600	504.68 ± 705.10	16 ↓
19.	Magnesium (mg)	310	365.49 ± 178.30	18 ↓
20.	Phosphorous (mg)		1232.93 ± 437.43	
21.	Potassium (mg)		2597.79 ± 1169.43	
22.	Cobalt (mg)		0.04 ±0.03	
23.	Copper (mg)		2.69 ± 0.92	
24.	Chromium (mg)		0.06 ±0.05	
25.	Iodine (µg)			
26.	Iron (mg)	21	21.74 ± 6.00	3.5 ↑

Indian Diabetes Risk Score

Figure 2: Percentage Distribution of Subjects According to Indian Diabetes Risk Score



From the figure 2, nearly 20.6% of the subjects fall in high risk category for diabetes followed by 66.2% of the subjects in medium risk category and 13.2% of the subjects are at low risk category for diabetes. Among the high-risk category of diabetes, 4.4% of subjects were diabetic and few of them in medium risk category were also diabetic due to age and they had family history of disease.

4. Conclusion

The tribes of India display a wide spectrum of India which we hardly know about in a different angle. Tribes are a part of our national life with their cultural heritage, they have been contributing to complex amalgam. Tribal women in Kolagur belonged to low socioeconomic status with low literacy rate and are well nourished due to good access to rice, pulses and oil through the Public Distribution System. Rice is consumed as their staple food as it is easily available through Public Distribution System. Black bean which is called as “Kottai” colloquially is consumed continuously on daily basis. Even though the availability of fruits is abundant through local markets, the consumption of fruits is less among them. Meat and meat products such as chicken, mutton and eggs were consumed as per the availability in the local blood haemoglobin level, packed cell volume, fasting blood glucose and serum lipid profile were found to be normal mainly due to the intake of diet nutritionally rich with low fat and extensive physical activity of walking for 3 to 4 km every day for work and other necessities such as visiting doctor, buying apparels, schools and colleges. A few subjects suffered from hypothyroidism due to lack of iodized salt in their diet. The cultivation, consumption of millets and a balanced diet can be encouraged by the repeated nutrition education program.

REFERENCE

1. Ahirwar, Rajeev & Mondal, Prakash. (2018). Prevalence of obesity in India: A systematic review. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*. 13. 10.1016/j.dsx.2018.08.032.
2. American Diabetes Association., *Standards of Medical Care in Diabetes-2018*. (2017). Retrieved from [Www.Copyright.Com](http://www.Copyright.Com).
3. Arun, D., Gupta, P., & Sachan, D. (2016). Indian Diabetes Risk Score (IDRS), A Strong Predictor of Diabetes Mellitus: A Hospital Based Cross Sectional Study of Lucknow. *International Journal of Medical Science and Innovative Research*, 1(7), 135–138. Retrieved from www.ijmsir.com.
4. Estridge, H.B., Reynolds, P.A., Walters, J.W. (2000). *Basic Medical Laboratory Techniques*. USA: Thomson Learning.

5. Gannon, C.M., Frank Q Nuttall, AsadSaeed, Kelly Jordan, Heidi Hoover . (2003) An increase in dietary protein improves the blood glucose response in persons with type 2 diabetes, The American Journal of Clinical Nutrition, Volume 78, Issue 4, October 2003, Pages 734–741, <https://doi.org/10.1093/ajcn/78.4.734>
6. Gallagher, D., Heymsfield, S. B., Heo, M., Jebb, S. A., Murgatroyd, P. R., & Sakamoto, Y.(2000). Healthy percentage body fat ranges: an approach for developing guidelines based on body mass index The American Journal of Clinical Nutrition, 72(3), 694–701. <https://doi.org/10.1093/AJCN>.
7. Global Database on Body Mass Index - World Health Organization. (n.d.). Retrieved March 23, 2020, from <http://www.assessmentpsychology.com/icbmi.htm>.
8. Guideline for High Blood Pressure in Adults (2017) - American College of Cardiology. (n.d.).Retrieved March 23, 2020, from <https://www.acc.org/latest-in-cardiology/ten-points-to-remember/2017/11/09/11/41/2017-guideline-for-high-blood-pressure-in-adults>.
9. Gupta, P. N. S. (1980). Food consumption and nutrition of regional tribes of India. Ecology of Food and Nutrition, 9(2), 93–108.
10. Kankanamge,S.U., Ariyaranthna,S. and Perera,P.P.R.(2017). Association between Dietary Pattern and Hb Concentration among Young Adult Female. International Journal of Current Research Vol. 9, Issue, 01, pp.45497-45500, January, 2017, Available online at <http://www.journalcra.com>.
11. Kolagur Village Population - Yercaud - Salem, Tamil Nadu [Internet]. [cited 2020 Mar 31]. Available from: <https://www.census2011.co.in/data/village/634073-kolagur-tamil-nadu.html>
12. Kumar Kshatriya, G. (2014). Citation: Gautam Kumar K. Changing Perspectives of Tribal Health in the Context of Increasing Lifestyle Diseases in India. In J Environ Soc Sci (Vol.1).
- 13.Mahan, L.K., Sylvia, E.S., Raymond., L.J., Krause, V.M. (2012). Krause’s Food & the Nutrition Care Process. USA: Elseiver Saunders.
14. Mathur N, Kesavachandran C, Bihari V. (2009). Can physical activity maintain normal grades of body mass index and body fat percentage? Int J Yoga. 2009;2(1):26.
15. Ministry of Tribal Affairs, Statistics Division, Government of India. Statistical Profile of Scheduled Tribes in India 2013 Ministry of Tribal Affairs Statistics Division Government of India www.tribal.nic.in [Internet]. 2013. Available from: www.tribal.nic.in.
16. Schwingshackl L, Hoffmann G. Long-term effects of low glycemic index/load vs. high glycemic index/load diets on parameters of obesity and obesity-associated risks: a systematic review and

- meta-analysis. Nutr Metab Cardiovasc Dis. 2013 Aug;23(8):699-706. doi: 10.1016/j.numecd.2013.04.008. Epub 2013 Jun 17. PMID: 23786819.
17. Thyroid Association, A. (2017). AMERICAN THYROID ASSOCIATION ®. Retrieved from www.thyroid.org.
18. WHO/Europe | Nutrition - Body mass index - BMI. (n.d.). Retrieved March 23, 2020, from <http://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi>
19. WHO. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and Mineral Nutrition Information System. Geneva, World Health Organization, 2011 (WHO/NMH/NHD/MNM/11.1).
20. World Health Organisation. (2008). Waist Circumference and Waist-Hip Ratio: Report of a WHO Expert Consultation. World Health Organization, 8–11. <https://doi.org/10.1038/ejcn.2009.139>.

IOT BASED SECURITY SYSTEM FOR HOME AUTOMATION

Ms. R.Janani¹

*Dept of Computer Application
MKJC,Tamil Nadu, India.*

Ms. S.Sneka²

*Dept of Computer Application
MKJC,Tamil Nadu, India.*

Ms. K.Nivetha³

*Dept of Computer Application
MKJC,Tamil Nadu, India.*

Abstract:

Internet of things is a fast becoming disruptive technology for business opportunity. It is a standards emerging primarily for wireless at communication between the devices and gadgets in day to day human life, in generally it referred to as things. In our research paper we represent a system that uses the internet networking to provide an easy path to our user to control the different appliances at home. IOT refers to the infrastructure of connected physical devices which is growing at a rapid rate as huge number of devices and objects are getting associated to the internet. The home automation can be implement using different types of wireless communication techniques such as ZigBe, Wi-Fi, Bluetooth, GSM, etc. These existing methods have the drawback because they work in short range. To overcome this drawback we are going to implement this projects on IOT based security system for home automation. By sending an captured image through E-Mail . IOT focuses on building a smart wireless home security system which sends alerts to the owner by using internet in case of any trespass and raises an alarm optionally . The same can be utilized for home automation to make use of set of sensors. The microcontroller used in the current prototype is the TI-CC3200 launch pad board which comes with an embedded micro-controller and an onboard Wi-Fi shield making use of which all the electrical appliances and managed.

Keyword:

IOT, Home Automation System , Wi-Fi, USB, TI CC3200

Introduction:

As we know that today world is progressing very fast, things are becoming very easy than before. People are considering the automation devices than the manual devices. People want the easier approach to the devices. Now a days automation is growing very fast. Internet is the basic part of universe communication. Internet is also used to communicate between the devices to operate accordingly .

Home automation:

It aims to bring the control of operating system to our every day home electrical appliances to tip of our finger. It gives the user affordable lighting solution and better energy conservation with optimum use of energy .Apart from just lighting solutions, the concept also further extends to have a overall control over our home security as well as build a centralized home entrainment system and so on. The basic definition are it creates a networking between home appliance and devices so that all the appliances in the houses can be controlled through a single machine . The internet of things based on home automation as more beneficial because of some reasons they are it is less costly for installation , it had a wide range of control of appliances ,energy saving system , better security of our home and it is easily access able to our cell phone or pc.

TI CC3200:

TICC3200 launch pad consist of applications microcontroller, Wi-Fi network processor, and power management subsystem. The TI CC3200 device is a complete platform solution including software, sample applications, tools, user and programming guides, reference designs, and the TI E2E support community. It uses ARM cortex M4 core processor at 80MHzand embedded memory including RAM (256 kb). The features of TI CC3200 are USB interface using FTDI USB drivers, board is powered through USB for the launch pad and the external booster pack, operated from 2 AA batteries , the TI CC3200 supports the four wire JTAG and two wire SWD, GDB supports over open chip debugger ,two20 pin connectors enable compatibility with the booster pack , flash memory is updated through USB using simple link programmer. The board can be programmed through Energia IDE over the USB cable the accessible Wi-Fi is used to the board that proxyless and it can be WAP type.

STM32L151:

The STM32L151/152 devices are designed for medical, industrial and consumer applications and feature an onboard fast12-bit 1 MSPS ADC, USB 2.0 FS and capacitive touch sensing. We may extend the memory range using the SDIO and FSMC interfaces. The STM32L151/152 devices extend the ultra-low-power concept with no compromise on performance, using the Cortex-M3 core and a flexible CPU clock from 32 kHz up to 32 MHz As well as the dynamic run and low-power run modes, two additional ultra- low-power modes bring you very low power consumption while keeping an RTC, backup register content and low-voltage detector. To reduce application cost and design, the STM32L151/152 offer an integrated LCD driver with up to 8 x 40 segments. The STM32L151/152 portfolio offers from 32 to 512 Kbytes of Flash and from 48- to 144-pin packages.

Wi-Fi:

Wi-Fi is the wireless technology used to connect computers, tablets, smart phones and other devices to the internet. Wi-Fi is the radio signal sent from a wireless router to a nearby device which translates the signal into data and we can access them and use it. It is a family of [wireless network protocols](#), based on the [IEEE 802.11](#) family of standards which are commonly used for [local networking](#) area for devices and [Internet](#) access. It allow nearby digital devices to exchange data by [radio waves](#). These are the most widely used computer networks in the world. It is used globally in [home and small office for networks](#) to link [desktop](#) , [laptop](#) , tablet, [smart phones](#), smart TV's, [printers](#), and [smart speakers](#) together to a [wireless router](#) to connect them to the Internet. In [wireless the access points](#) in public places like coffee shops, hotels, libraries and airports to provide the public Internet access for mobile devices. Most commonly it uses the 2.4 gigahertz (120 mm) [UHF](#) and 5 gigahertz (60 mm) [SHF](#) radio bands. These bands are subdivided into multiple channels. Channels can be shared between networks but only one transmitter can locally transmit on a channel at any moment in time. It include the technology for standards mobile phone, such as [2G](#), [3G](#), [4G](#), [5G](#) and [LTE](#).

ESP8266:

The ESP8266 Wi-Fi module in an ease Wi-Fi chip with full TCP/IP stack and MCU (Micro Controller Unit) capacity was created by Chinese .Primary arrangement of modules made with ESP8266 by outsider producer AI-Thinker and remain the most generally available. The large alluded to as ESP-xx modules and shape a workable advancement framework and it require the extra parts .A particularly serial TTL-to-USB connector and outside 3.3 volt control supply.

USB:

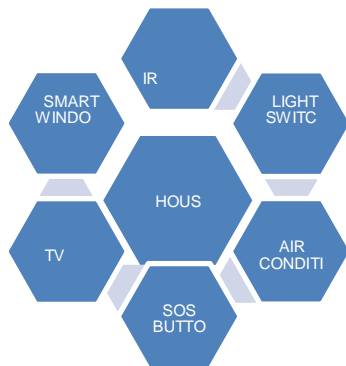
Universal Serial Bus is an industry standard that establishes specifications for cables and connectors and protocols for connection, communication and power supply between computers, peripherals and other computers. USB is an attractive option for connectivity and controls because it is easy to use prototype in an application by buying off the shelf chips and plugging them together. It should take few hours to prototype an IOT application by plugging together like USB cameras, sensors and a Wi-Fi link. Implementing USB in an IOT platform chip means that features can be added to IOT product either by adding a chip to the system PC or by plugging in USB peripheral. USB host drivers are widely available for Linux and other open source operating system .

It supports either in the form of integration the USB gives an advantage over interfaces such as 12C. Many IOT platform chip designers are now using the USB 2.0 interfaces specification and the USB type-c connector as an integration standard for IOT designers at the board level for finished IOT design at the system level.

Literature survey:

Home automation system, Device to device broad caste, Device to cloud broad caste, Device to gateway mode, Internet of things problems, Android mobile current generation

Home automation system:



Its cover the software, hardware, sensor, protocols, architecture and platforms. Application of IOT enabled connectivity are home security, air quality monitoring ,infotainment delivery, smart lock etc., Home automation process ,aims to Control all the device of your smart home through internet protocols or cloud based computing. What really would compel other to actually developed a product which is a finish internet of things based on home automation system? Could it be want to develop the safety of your house, could it be the desire to live a jetson like life that millennial always dreamt of device.

Device to device board cast:

General configuration of IOT communication model may include communication between device, as well between cloud and gateway. Device to device can be directly connected and can communicate with each other over IP network of the internet and can use protocols like Bluetooth. It's is advanced data transmission technology developed to increase the efficiency of network.

Device to cloud board cast:

Device to cloud communication involves and IOT device connecting directly to internet cloud serve like an application services provider to exchange data and control message. A use of cellular based devices to cloud would be smart dog that track your dog while you're not around , which would need wide area cellular communication because you wouldn't know where the dog might be.

Gadget to gateway model:

An IOT gateway is a solution for enabling IOT communication, usually device to device communication or device to cloud communication. The gateway is typically hardware device using housing application, software that essential task. A gateway facilitates the connection between different data sources and destinations. Features sets: Data coaching, buffering, Some data aggregation, System diagnostics. A gateway is a physical device or software program that serves as the connection point between the cloud and controllers, sensor and intelligent electronics.

IOT problem (home automation):

common home automation issues and our solution: Sub system not integrating, Insufficient features and functionality, Too many home automation control apps, Lengthy delays during the construction of your project. Lack of physical security, attackers may sometimes, make physical changes in IOT devices that are located in remote place for long time. Botnot attracts, Lack of visibility, Industrial espionage, Hijacking, Incorrect access control.

Android mobile current generation:

Without mobile we cannot access any data in current generation. Only mobile will access the all data at any time, anywhere we can use with the help of internet.

Hepatic feedback sensors:

Uses case where smart phone can be used to an IOT device include personal emergency response, fitness tracking, location based asset tracking and natural vision processing.

PIR sensor:

A passive infrared sensor (PIR sensor) is an electronic sensor that measures infrared (IR) light radiating from objects in its field of view. They are most often used in PIR-based motion detectors. PIR sensors are commonly used in security alarms and automatic lighting applications. PIR sensors detect general movement, but do not give information on who or what moved. For that purpose, an imaging IR sensor is required. The term *passive* refers to the fact that PIR devices do not radiate energy for detection purposes. They work entirely by detecting infrared radiation (radiant heat) emitted by or reflected from objects. A PIR-based motion detector is used to sense movement of people, animals, or other objects. They are commonly used in burglar alarms and automatically activated lighting systems.

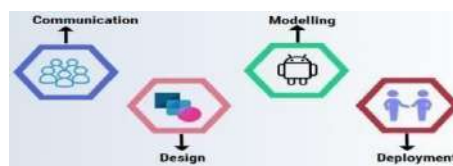
The most common models have numerous Fresnel lenses or mirror segments, an effective range of about 10 meters (30 feet), and a field of view less than 180°. Models with wider fields of view, including 360°, are available, typically designed to mount on a ceiling. Some larger PIRs are made with single segment mirrors and can sense changes in infrared energy over 30 meters (100 feet) from the PIR. The PIR sensor is typically mounted on a printed circuit board containing the necessary electronics required to interpret the signals from the sensor itself. The complete assembly is usually contained within a housing, mounted in a location where the sensor can cover the area to be monitored. The housing will usually have a plastic "window" through which the infrared energy can enter. Despite often being only translucent to visible light, infrared energy is able to reach the sensor through the window because the plastic used is transparent to infrared radiation. The plastic window reduces the chance of foreign objects (dust, insects, rain, etc.) from obscuring the sensor's field of view, damaging the mechanism, and/or causing false alarms. The window may be used as a filter, to limit the wavelengths to 8-14 micro meters , which is closest to the infrared radiation emitted by humans.

NOTES:

PIR sensor is used in IOT that measures the light radiating from object to field of view. It is used in thermal sensing application like security alarm, motion detection alarm, automatic lighting.

Prototype:

The prototype is the first, crucial step to building an Internet of Things product. The purpose of building a prototype of your IOT product is manifold: to test the concept with your target audience, to check that the technical requirements match your expectations, and ultimately to validate your business case. we regularly work with customer projects that start with an early prototype and end up as a successful business. This article is a condensation of our experience with IOT prototyping and the underlying building blocks. An IOT prototype consists of: A user interface, such as a smart phone app or web frontend, A hardware device ,Backend software, that implements the business logic, Connectivity.



Communication:

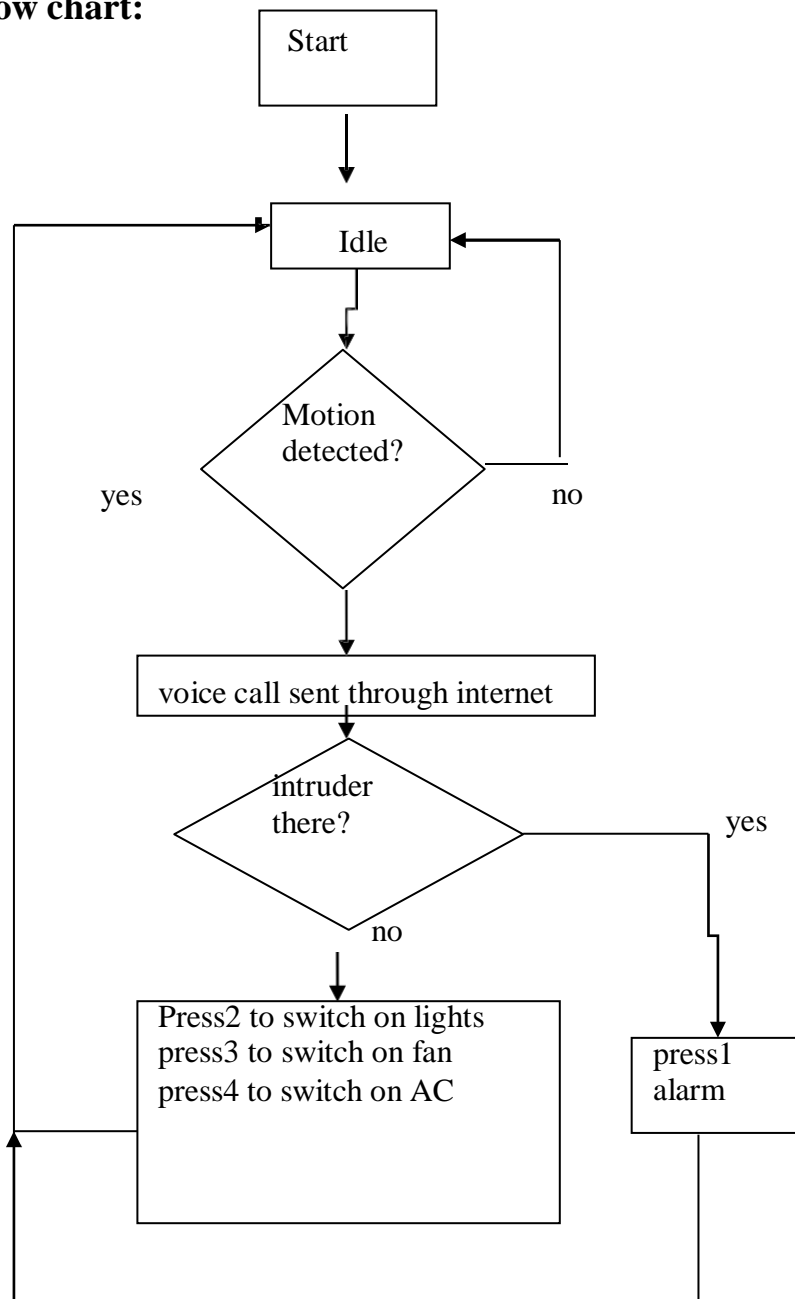
This is the phase where the developer and Client set up the meeting and talk about the objectives that need to be achieved for the software. **Design:** The design has been done quickly as the requirements are cleared from both ends in providing and receiving end. It is used to construct the prototype. It includes the important aspects of the software that is input and output but mainly focused on visible aspects than planned activities.

Modeling:

It provides a better idea of the requirement to develop the software and a better understanding of the software product.

Deployment:

Before deployment, the client evaluates the software, and if the client is not satisfied, then it is refined according to the client's requirements. This process goes on until the point client requirements do not accomplish. After the client's satisfaction with the product, the product finally deploys to the production environment. It is thoroughly evaluated and tested, and maintenance is done on a regular basis. The prototype model mainly works at the high end in the below format: Objectives determination, Develop the code, Communication and refine, Demonstrate, Test, Implement.

Flow chart:

Conclusion:

The IOT based home security and automation is very useful for remote users. Every successful IOT based product, system, or service needs to be prototyped to validate the business and technology choices. Some of the decisions along the way are different than what they are for the final product. For example, the primary purpose of the user interface is not to let the user interact with the system, but to prove the value to the customer. This must be taken into account then designing and building the prototype. When building an IOT prototype, there is no need to go too far down the hardware and connectivity rabbit holes – not until after the prototype has proved the business case. Any home can be monitored and controlled by using the prototype implemented. The prototype which provides reliable, cost effective and efficient to the IOT world solution.

References:

1. Umer Ijaz, Usama Ameer , Dr.Badar Islam (2016), IOT based home security and automation
2. Sudha Kousalya , G.Reddi Priya , R.Vasanthi , B.Venkatesh(2018), IOT based smart security and smart home automation.
3. A.Anitha (2017), home security system using internet of things.
4. Ravi Kishore Kodali, Vishal Jain , Suvadeep Bose, Lakshmi Boppana (2016), IOT based smart security and home automation system.
5. A.Madesh , Dr.M.Kowsigan , R.Kishore, Dr.P.Balasubramanie (2020), IOT based home automation and security system.
6. Rizwan Majeed, Nurul Azma Abdullah, Imran Ashraf , Yousaf Bin Zikria, Muhammad Faheem Mushtaq and Muhammad Umer (2020), an intelligent, secure, and smart home automation system.
7. Pranay Pratim Das, Indranil Bhattacharjee (2019), smart security and home automation using internet of things.

Review Paper on Various Software Testing Techniques & Strategies

Ms. K. Lavanya¹

*Dept of Computer Application
MKJC, Tamil Nadu, India.*

Ms. S. Dhatchayani²

*Dept of Computer Application
MKJC, Tamil Nadu, India.*

Abstract

Software checking out is the procedure of strolling an software with the purpose of locating software program bugs (mistakes or different defects). Software programs call for has driven the nice guarantee of evolved software program in the direction of new heights. It has been taken into consideration because the maximum vital degree of the software program improvement existence cycle. Testing can examine the software program object to discover the disparity among real and prescribed situations and to evaluate the traits of the software program. Software checking out results in minimizing mistakes and reduce down software program costs. For this purpose, we talk various software program checking out strategies and strategies. This paper pursuits to take a look at various in addition to improved software program checking out strategies for higher nice guarantee purposes.

Keywords: checking out strategies, checking out tools, verification, validation, degree of checking out, debugging, software program checking out objectives, software program checking out principles, software program checking out strategies, debugging, checking out methodologies, software program checking out existence cycle.

1. Introduction

Software improvement includes growing software program in opposition to a fixed of requirements. Software trying out is had to affirm and validate that the software program that has been constructed to satisfy those specifications. Software trying out allows withinside the prevention of mistakes in a system. It refers back to the system of comparing the software program to find out the mistake in it. It is likewise used to investigate the software program for different components of the software program like usability, compatibility, reliability, integrity, efficiency, security, capability, portability, maintainability, etc. Software trying out ambitions at reaching precise dreams and principles which might be to be followed. In easy words, trying out is the system of finding mistakes with inside the program.

Software Testing is executing the software program to (i) perform verification, (ii) to discover the errors and (iii) to achieve validation.

- i. **Verification:** It is the system of checking the software program regarding the specification.
[Verification: Are we making the product, right?]
- ii. **Error Detection:** It is the system of deliberately acting the incorrect inputs to test the system's performance.
- iii. **Validation:** It is the system of checking software program regarding the customer's expectation.
[Validation: Are we making the right product]

Testing is characterised as a manner of evaluation that both the definitive gadget meets it's distinct ful fillments to start with or not. It is particularly a gadget beset with the validation and verification manner that whether or not the advanced gadget meets the fulfillments described via way of means of the customers. Therefore, this hobby draws the distinction among the real and anticipated result. So, that is an evaluation that offers the buddies with the right expertise approximately the best of the product. Software Testing also can be described as a risk-based hobby. The checking out fee and mistakes may be observed in a courting in determine 1. It is seemingly demonstrated that fee rises dramatically in a checking out (purposeful and nonfunctional). The compelling checking out purpose is to do the surest quantity of checks in order that extra checking out undertaking may be minimized. From Figure 1, we will say that Software checking out is an crucial element in software best assertion

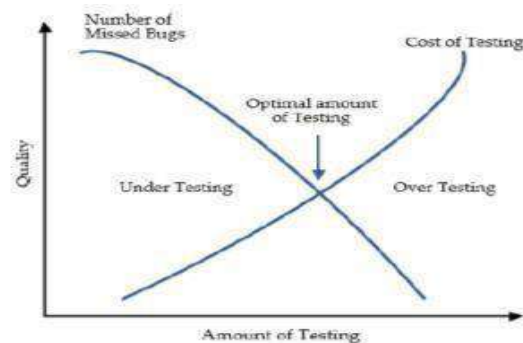


Figure 1: Every Software Project has optimal test effort
(Courtesy)

a) Basic Terminology

Mistake, Error: Human makes a mistake. A suitable synonym is an error. It variations among the actual output and the predicted output.

Bug: It method that once a developer makes the error at the same time as coding. It is the nation that's accountable for the failure of the precise function.

Fault: It is the illustration of the error, where illustration is the suggest of expression that can be diagrams, waft charts, etc.

Failure: Failure happens whilst a fault is executed.

b) Software Testing Objectives The essential goal of software program trying out is to offer a first-rate product concerning the reliability estimation and whole verification and validation of the product. The secondary goal of trying out includes executing a application with the reason of locating mistakes and generating a take a look at case that is able to detecting the undiscovered blunders as yet.

II. Literature Review

In this section, we can define the preceding works of Software Testing. According to "The Theory of Software Testing", trying out is the way of displaying the presence of mistakes withinside the software that can both be accomplished manually or automatically. It additionally consists of the primary terminology of trying out which include computerized trying out, failure, trying out team, and incorrect check case selection. This paper makes a speciality of the manner that ought to be observed to check the overall performance of latest software program and the entire system. The end of the thing is the complete view of software program trying out, initial trying out, and user attractiveness trying out.

III. Software Testing Methodologies

The significance of software program trying out to software program best can not be overemphasized. After the improvement of the code, it's miles obligatory to check the software program to discover all of the mistakes, and that they have to be debugged earlier than the discharge of the software program. Although it is not possible to discover and debug all of the mistakes withinside the huge software program at each phase, it's miles attempted to remove all of the errors as possible. Testing facilitates in locating the bugs; it can not finish that the software program is bug- free. We extensively categorised trying out strategies into two parts:

1. Static Testing
2. Dynamic Testing

1. **Manual Testing (Static Testing):** It refers to the method of testing where the code is not executed. It does not require highly skilled professionals since the actual execution of the system is not done in this process. It starts with the initial phase of the Software Development Life Cycle (SDLC); hence, it is also known as verification testing. The main objective of static testing is to enhance the quality of software products by helping software professionals to Identify and resolve their errors early in the software development process. Static testing is performed on the documents like Software Requirement Specification (SRS), design documents, source code, test suites, and web page content. It is performed before code deployment. As a result, it provides the evaluation of code as well as documentation. Static testing techniques include:
 - i. **Inspection:** It is primarily done to locate defects. Moderators conduct the code walkthrough. In this type of formal review, a checklist is prepared to check the working document.
 - ii. **Walkthrough:** It is not a formal process. The authors lead this process. The Author advises the participants through the document according to his or her thought process so that they can accomplish a common perception. It is especially useful for higher-level documents like requirement specification, etc.
 - iii. **Technical Reviews:** A professional round of review is operated to check if the code is made in consonance with the technical specifications and norms which may include the test plans, test strategy, and test scripts.
 - iv. **Informal Reviews:** It's the static testing technique in which the document is reviewed unofficially, and helpful comments are implemented.
2. **Automated Testing (Dynamic Testing):** Dynamic Testing is a kind of software testing technique in which the dynamic behavior of the code is analyzed. In dynamic testing, also known as validation where the actual system is considered. It requires the highly skilled professional with the proper domain knowledge. Dynamic testing involves testing the software for the input values, and output values are analyzed. Progressive testing is divided into two categories:
 - A. White Box Testing
 - B. Black Box Testing
 - C. Grey Box Testing

A) White Box

It's trying out Internal specs and systems of the machine are created conspicuous. So, it is acutely cost effective in detecting and resolving problems. Bugs will be discovered earlier than they reason bickering. Thus, we will summarize this method as trying out software program with the statistics of its inner shape and coding. White container trying out is likewise acquainted as particular container evaluation or white container evaluation or glass container trying out or obvious container trying out, and structural trying out. It's an method for locating mistakes inside which the tester has entire statistics. This method isn't always used lots for debugging in large structures and networks. Different forms of white container trying out encompass foundation path, loop trying out, manage shape trying out, etc. White-container trying out checks inner buildings or workings of a application because programming talents and the home context of the machine are used to layout check cases.

White-box Testing consists of the subsequent approaches:

- 1. Application Programming** - Interface checking out checks the utility the usage of public and personal APIs by developing checks to meet a few standards for code insurance.
- 2. Fault Injection Methods** – Introducing faults to gauge the efficacy of checking out techniques intentionally.
- 3. Code insurance equipment** can check the integrity of a check suite that turned into created with any method, which includes black-field checking out. This offers the software group to test the elements of a device which are rarely examined and assures that the maximum critical function factors were verified.

Advantages:

1. It exposes an blunders this is hidden in code by casting off greater traces of code.
2. Maximum insurance is received at some stage in check outline writing.
3. The developer discreetly offers motives for implementation.

Disadvantages:

1. An skilled tester is needed to perform this method due to the fact expertise of inner shape is needed.
2. Many paths will continue to be untested because it is hard to investigate each professionals and con.

b) Black Box checking out A black container checking out is a checking out wherein inner specifics and workings are not recognised or reachable to its customer. It helps specs and output needs. The essential reason is to pick out the requirements of the system. Black container checking out has little or no or no information at the internal logical shape of the system. Thus, it completely tests the simple capabilities of the system. It assures that each enter is as it should be familiar and outputs are effectively produced. Black-container checking out tests the capability with none expertise of the inner implementation. The testers most effective have an expertise of what the software program is meant to do, now no longer the way it does it. This is only performed primarily based totally on customers' aspect. The tester most effective is aware of the set of inputs and unique outputs. Black-container checking out techniques include:

i. Equivalence Partitioning: This approach divides the enter area of a application into equivalent lessons from which take a look at instances may be derived. Thus, it could reduce the range of take a look at instances.

ii. Boundary Value Analysis: It objectifies the checking out at borders, or in which the intense boundary values are chosen. It contains of minimum, maximum, blunders values, and ordinary values.

iii.Fuzzing: This technique takes random enter to the application. It is used for characterizing implementation bugs, the use of malformed or seminal formed information injection in an automatic or semi-computerized session.

iv. Orthogonal Array Testing: In this approach, the enter area is minimum however too massive to house exhaustive checking out.

v . Cause-Effect Graph: This checking out approach begins with the aid of using producing a graph and organising the relation among impact and its causes.

vi. All Pair Testing: The essential goal is to have a set of take a look at instances that covers all of the pairs. Here, take a look at instances are designed to execute all possible discrete combos of every couple of enter parameters.

vii. State Transition Testing: This checking out technique is beneficial for the navigation of a graphical user interface.

Advantages:

1. Testers do not now no longer want to realize unique programming languages. Testing is performed from a user's factor of view.
2. It facilitates to discover any ambiguities or inconsistencies with inside the requirement specs
3. Programmer and validators each are unbiased of every other.

Disadvantages:

1. Test instances are hard to layout without fair stipulations.
2. Probability of getting the repetition of exams that are already executed with the aid of using the programmer.
3. Here, a few components of the returned cease aren't examined at all.

c) Grey-Box Testing

Gray field trying out is the approach of trying out the software program with confined information of the inner structure and layout of the application. It is described as a trying out software program bundle which has a few statistics of its inner good judgment and underlying code. It makes use of inner information systems and algorithms. This technique holds important undertaking integration trying out among or extra modules of code written with the aid of using completely different developers. This technique consists of opposite engineering to training session at the boundary values. Grey field trying out is impartial and non-intrusive. Grey-field Testing has the information of inner statistics systems and algorithms for functions of designing exams even as executing the ones exams on the user level. The tester does not now no longer have complete get admission to to the software program's supply code. The following are a few subtypes of grey-field trying out:

- i. State-Model-Testing: It examines every approach of an object, transition & transition paths at every kingdom of an object.
- ii. Class-Diagram Testing: It examines all of the derived instructions of the bottom class.
- iii. Sequence-Diagram Testing: It examines all of the strategies taking place withinside the collection diagram.
- iv. Thread-Based Testing: In this technique, all of the instructions of unmarried Use-Cases are incorporated and then trying out is performed. This technique is going on till all of the guides of all Use-Cases have been considered.

Advantages:

1. It implements the combined benefits of black-box and white-box testing techniques.
2. In grey box testing, the tester can design high test scenarios.
3. Testing is unbiased.

Disadvantages:

1. Test coverage is limited because the access to source code is not available.
2. Many program paths remain untested.
3. The test cases can be redundant.

2. Non-functional Testing Techniques:

It is a type of testing which is performed to test various attributes of a system like stress, load, etc. Non-functional testing is performed at all test levels. It is concerned with the non-functional requirements and is explicitly modeled to assess the readiness of a system according to the various criteria which are not covered by functional testing.

- a) Performance Testing:** It assesses the whole overall performance of the machine. It is used to assess the machine's overall performance beneath Neath a selected workload.
- b) Load trying out:** A load take a look at is accomplished to make sure the weight taking the capability of a selected order. Load trying out is accomplished to decide the conduct of the machine beneathneath everyday in addition to height load conditions.
- c) Endurance Testing.** It is the sort of trying out accomplished to decide the machine's conduct after a selected time. For example, a way is running flawlessly nice withinside the preliminary first hour, however the overall performance reduced after 3 hours of execution.
- d) Stress Testing:** It is accomplished to decide the machine overall performance past fashionable operational capability, regularly to a breaking point. It is involved with the machine's load taking ability.
- e) Security Testing:** Security trying out is completed to evaluate that the machine is secure or not. It is a procedure this is involved with the reality that an facts machine protects statistics and continues capability as intended.

f) Recovery Testing: Recovery trying out is accomplished to test the restoration of the machine after the crash or hardware failure. In this type, the software program is compelled to fail beneathneath a given situation after which finally, the restoration is tested.

g) Compatibility Testing: Compatibility trying out is involved with checking the machine's compatibility with the relaxation of the environment. It assessments the concord of the evolved machine with the diverse different additives which includes hardware, extra software program, DBMS and working machine, etc.

IV. Software Testing Strategies

Software Testing strategies provide a method of integrating software test case design methods into a well-planned Series of steps that can result in the successful construction of software. It provides the road map for testing. The software testing Strategy should be pliable enough to develop a customized testing approach. The software testing strategy is actually produced by project managers, software engineers, and testing specialists. There are four different types of software testing strategies:

- 1) Unit testing
- 2) Integration testing
- 3) Acceptance/Validation testing
- 4) System testing

Unit testing :

Unit is the smallest testable part, i.e. the most modest collection of lines of code which can be tested. Unit testing is done by the developer as the proper knowledge about the core programming designing is required. Generally, unit testing is considered as a white-box testing class because it is partisan to evaluate the code as implemented rather than assessing conformance to some set of requirements.

Benefits of Unit Testing:

- 1) Cost-effective testing technique.
- 2) Simple testing technique because the smallest testable unit of the code is tested here.
- 3) Individual parts are tested when necessary, without waiting for another part of the system.
- 4) Unit testing can be performed in parallel by fixing problems simultaneously by many engineers.

- 5) Detection and removal of defects are much costeffective compared to other levels of testing.
- 6) Be able to take advantage of several formal testing approaches available for unit testing.
- 7) Clarify debugging by limiting to a small unit the possible code areas in which to search for bugs.

V. Discussion

In this section, the difference between testing and debugging is shown. Software testing is a process of executing a program in order to identify whether it meets the specified requirement or not. The test case design is conducted, a strategy can be defined, and results can be evaluated against prescribed expectations. Debugging is the process of fixing the problems that are found during the software testing process. It is a consequence of successful testing. That is when testing uncovers a defect, debugging is the process that leads to the removal of the error. The debugging process can be done by the programmer or the developer. Based on the defect that is reported, the developers try to find the root causes of the error to Global Journal of Computer Science and Technology Volume XIX Issue II Version I 48Year 2 019 () C © 2019 Global Journals Review Paper on Various Software Testing Techniques & Strategies make the system defect free. Debugging is done in the three phases of software development. The first one is during the coding process when the programmer translates the design into executable code. In the coding process, the errors that are made by the programmer when writing in code needs to be detected quickly and fixed before going to the next phase of development. In these cases, the developer performs unit testing to identify any defects at the module or component level. Then debugging is performed during the last stage of testing. These involve debugging multiple components or a complete system, when unexpected behavior such as wrong return codes or abnormal program termination may be found.

VI. Conclusion

To conclude our survey, we would like to find that Software testing is an essential activity of the Software Development Life Cycle (SDLC). We can never say that a product is "Perfect". Testing is a never-ending process. Testing only shows the presence of errors, not the absence. It is time-consuming and an intensive process, therefore, upgraded techniques and innovative methodologies are necessary to maintain the quality of the software. This leads to performing Automated Testing implementation before and during the testing process. This paper aims to describe in detail various software testing techniques, the need for software testing, software testing goals, and objectives. Software testing is often less formal and meticulous than it should. In order to perform software testing effectively and

efficiently, those who are involved in testing must be familiar with the basic concept, goals and principle of software testing.

We further discuss different software testing techniques such as white-box Testing, black-box Testing, grey-box Testing, Security testing etc. Hence, the future work in relevance with the testing process will be much more technology-dependent harnessing the simulation and automated testing model-based approach, not only expediting the testing life cycle but also providing optimum bug prevention and efficient quality assurance.

References

1. Kaner, Cem (November 17, 2006). "Exploratory Testing" (PDF). Florida Institute of Technology, Quality Assurance Institute Worldwide Annual Software Testing Conference, Orlando, FL. Retrieved November 22, 2014.
2. Glenford J. Myers, "The Art of Software Testing, Second Edition" Published by John Wiley & Sons, Inc., Hoboken, New Jersey. 1.1. P. Mathur, "Foundation of Software Testing", Pearson/Addison Wesley, 2008.
3. P. Mitra, S. Chatterjee, and N. Ali, "Graphical analysis of MC/DC using automated software testing, in Electronics Computer Technology (ICECT), 2011 3rd International Conference on, 2011, vol. 3, pp. 145 –149.
4. F. Saglietti, N. Oster, and F. Pinte, "White and greybox verification and validation approaches for safety- and security-critical software systems," Information Security Technical Report, vol. 13, no. 1, pp. 10–16, 2008.
5. "Software Testing Techniques and Strategies" Abhijit A. Sawant¹, Pranit H. Bari² and P. M. Chawan³ Department of Computer Technology, VJTI, University of Mumbai, INDIA.
6. Mohd. Ehmer Khan, "Different Forms of Software Testing Techniques for Finding Errors", IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 3, No 1, May 2010.

FILE SYSTEMS FOR VARIOUS OPERATING SYSTEMS: A REVIEW

*Ms. L. Shalini¹,
Programmer
MKJC, Tamil Nadu, India.*

Abstract:

The hardware and software program each technology have become advanced, there are modifications being introduced everyday. Therefore the garage strategies are had to be up to date for those modifications to be compatible and for green utilization. A plethora of garage strategies or document structures had been introduced. A survey of those document structures is provided right here on this study, which covers the purpose, working, advantage and barriers of the protected document structures

Keywords: Cluster FS, dispensed FS, flash reminiscence FS, journaling FS, log dependent FS, garage structures

INTRODUCTION

File device is the manner of storing the statistics on physical storage gadgets like disk, magnetic tapes, compact disk, flash drives and many others or the hierarchical agency of data established via way of means of working device. In computing environment working device is answerable for data organization and record structures management. With the passage of time the garage wishes are modified and the amount of statistics increased. A record device have to be reliable, persistent, secure, green, fault tolerant and scalable. To achieve those residences and to maintain tempo with changing computing wishes and garage requirements, different techniques and record structures are added via way of means of time. To make the green use of our statistics record device is very important. Therefore, a survey for the record structures is presented on this paper. The classes protected are: cluster record device, disbursed record device, parallel filesystem, flash record device, journaling record device, log filesystem, cellular record device, multimedia record device and network record device

Comparison table parameters: The distinctive parameters are used to represent and examine the document system's performance. Load balanced, scalable, cryptography, adaptable, anonymity, persistent, garage type, read/write, platform characterizing parameters used here (Ragib teal., 2005).

Cluster file systems : When more than one item primarily based totally storage devices are connected to a community, patron's records acquired in patron's request is quickly saved to the patron's memory after which written to item primarily based totally garage device but every now and then the reminiscence of patron is erased before writing to item primarily based totally garage tool due to smaller space to be had on item primarily based totally garage tool than the size of document to be written. OASIS-OSD is suggest set of rules to clear up this problem. For writing operation, an object primarily based totally garage tool is selected. The contents offile are saved on written except an mistakes message is generated from item primarily based totally garage tool that the spaces finished. Then final pages of document which are not written but are saved after which written to some other object based garage tool with to be had space, the proposed method was implemented with OASIS (Myung-Hornet al., 2007). The cluster generation achieves high performance through grouping reasonably-priced servers to cluster. Though these cluster document structures acquire higher overall performance buta change with inside the software program is wanted on patron side, which limits their extensive application. Therefore a network file get entry to interface is mixed with cluster document system to deal with this problem. CFS-SI includes three components: document server node, metadata server node, I/Onoda. The fashionable community document machine runs at the fileserver node and it accepts the requests of customers in standard community document get entry to patterns. File server node saves all metadata on metadata node and does now no longer save any records of CFS-SI. Then this stored metadata is used to access I/O node. Likewise the entire procedure of the network document machine is completed (Jun et al., 2009). Most supercomputers are **located** in shape of huge clusters nowadays which want for compact, more desirable and dispersed metadata processing techniques. An best meta data processing coverage calls for the automated stability of namespace locality and even Distribution with none manual processing. DDG walls the call area into hierarchical devices dynamically the usage of triple defined distribution granularity. Another method S2PC-MP is proposed for move over operation's consistency.

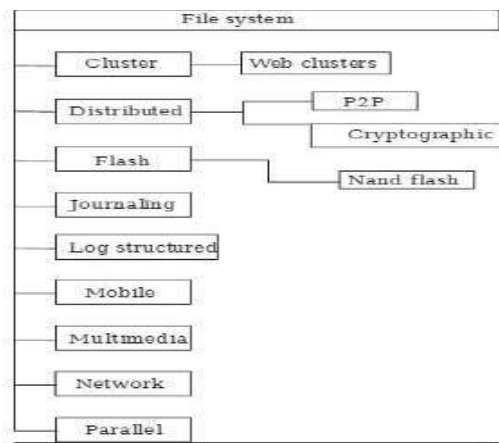


Fig. 1: DFD file systems

DISCUSSION

Web cluster file systems:

The concept of TH-Cliffs is largely originated from the operating of excessive speed networks generally get entry to the faraway statistics as rapid as they are accessing the statistics positioned locally. This emphasis the need of an I/O balancing method for cluster documents structures. Initially the documents are transferred from busy nodes to free nodes to proportion the I/O load. Finally, disk cache and memory are blended to make record cache of reminiscence disk. Unique cache plays the characteristic of I/O balancing at file cache level. This simply plays the performance of net server clusters in comparison to different traditional methods (Wei et al., 1999) (Table 1). Web proxy servers have an essential vicinity in today's net architecture. Disk I/O is a hassle the various most important troubles appeared against the overall performance of proxy servers as proven by previous studies. Traditional record structures do now no longer have good overall performance for proxy servers and feature large overheads. UCFS is a machine this is evolved for enlightening the I/O overall performance significance of proxy servers. UCFS keeps the tables for metadata available in reminiscence and approximately the all load of metadata updates and searches are eliminated. It displays a clustered record system that makes use of large disk transfers for the development of performance of writing operation of disk. Clustered filesystem additionally complements the study operation and does not create garbage. Consequently UCFS improves the performance of proxy servers as proven through end result of experiments (Jun et al., 2002).

Distributed file systems:

There are many instances in which application want the one-of-a-kind get admission to to a positive document but privileges given to a few document aren't pre-empt chapin a position and it limits the get admission to to that document. Local document machine offers with this hassle through simply assigning the rights and assessments the instances of an open document for greater opens towards some more requests. This required greater overhead through having the need for one greater server for registering the times of opened files. To cope with one-of-a-kind, get admission to problem semi pre-empt chapin a position locking mechanism is introduced; the server is unknown of any nation of a document opened globally. This nation is maintained at patron facet individually. When a request of opened document comes the server sends the messages to patron to remedy the battle while client releases the held lock then the soliciting for patron makes use of it (Randal et al., 2000). The improvements in reminiscence and processor era continuously emerge the disks with stronger processing electricity and compact cache Remini scene. By this improved processing electricity disk is permitted to perform greater operation than best the not unusual place disk operations. The latency of information manipulation can be reduced through giving software processing offloading part to disk. Suck disks are referred to as energetic disks. ADFS is a file machine wherein information server is energetic disk based. Datafiles resided on those energetic disks have the competencies of operations and making objects.

Advantage and limitation table for cluster file system

Technique	Advantage	Limitation
TH-CluFS(Wei et al., 1999)	Combines the advantage of parallel I/O systems and distributed file system and is suitable for cluster file system..	Cannot be implemented on distributed file system independently due to the volume dispersion conflict as in distributed file systems volumes stay on a single server but TH-CluFS spans multiple servers which are specified by system automatically.
UCFS (Jun et al., 2002)	User space is used to run the whole system Therefore, it is not difficult and not expensive as well from the implementation point..	Tested and implemented as study projects, no real world implementation yet.
OASIS-OSD (Myung-Hoon et al., 2007)	Uses two techniques namely i-prevention & ii- detection and recovery to ensure sufficient space on OSD for a file and write of file to the OSD respectively.	A locking technique is used for imposing mutual exclusion on the process in writing operations which creates load to the system.

Peer to peer file system: SDA-DFS is report allocation, the 2 strategies of replication and fragmentation are used at servers. The safety of report along with secrecy and integrity are secured even a subset of servers are sufferers of a few safety attack.

The set of rules is of adaptive nature due to alternate in examine write styles of report allocation and client's region in community alternate (Alessandro et al., 2003). A peer to see set of rules designed for the functions of obvious examine and write from the garage gadgets the usage of an interface named FUSE, Provides functions of excessive statistics availability through the usage of the replication and excessive fault tolerance through the usage of decentralization. Due to DHT calls it's miles scalable.

Cryptographic file system: A disk machine connected with the stable community machine is provided which suggests that cryptographic safety may be applied with the allotted record machine. The form of those structures is lower priced with the excessive pace processors of today. The maximum dependable and stable approach for the consumer is signing the checksum of every block for which they may be the usage of encryption primarily based totally on public key and on disk facet is to test each block for authentication earlier than writing to it. Three schemes are used scheme1 is slower due to signature technology and checking sum.

Flash file systems: It displays the success of fee effectiveness as it's far the maximum ideal and essential factor throughout improvement method of cell patron gadgets. Applying the compression mechanism is straightforward however an powerful technique closer to the success of fee effectiveness. An evaluation of strategies for compression of cell gadgets for consumers, at record machine stage is given on this study. Traditional record structures of compression are optimally used for disk orientated machine and feature wealthy assets of computing, they're now no longer well applicable for cell gadgets, as with susceptible strength of processing and little memory (Seunghwan et al., 2007). Due to the common writing of small statistics inserts an opening among span of lifestyles and persistency.

Journaling file systems: To compare the overall performance in phrases of robustness of journaling document machine the approach is proposed below failure of disk writes. Constructs fashions how journaling document machine orders writes of disks below severa modes and those fashions are used to restore write failures (Vijayan et al., 2005). Dual FS keeps statistics and metadata in separate gadgets and manages them differently.

Metadata is controlled as log structured document machine and statistics is controlled in groups. It greatly decreases the I/O time in workloads taken through document machine (Juan et al., 2007). Robustness of journaling document systems is evaluated via this approach below disk write failures. Models are built for distinctive journaling document machine modes and are used to inject the faults into machine (Vijayan et al., 2005). Dual journaling approach is used in this technique, twin in a feel that sorts of statistics is saved from starting and finishing to the centre element of garage device (Jeong-Ki et al., 2006).

Mobile file systems: With the improvement of cell gadgets into transportable facts gadgets it's miles becoming important for the designing of a allotted record gadget for such conversation gadgets. In this venture one have to keep in view the to be had little storage, little cappotential of computation and unreliable mobile networks. Mobile code era separates the not unusualplace record structures into elements one is purchaser component that consists of a few documents requiring little computing and a few operation; the other component is server component containing a masses of operations and documents with heavy computing abilities. The record gadget of transportable tool is the a part of server's record gadget (Yurong et al., 2001). The transportation of massive documents from a purchaser to server on susceptible connection is a essential hassle of cell record structures.

Multimedia file systems: There is a vast call for for video on call for utility with the growing networking and computing. Minimum possibility of blocking off is invented on this paper as load sharing is required for those systems. For placement of multimedia report the genetic set of rules and bin packing set of rules are combined (Kit-Sang, 2001). Solution to permit difficult disk drives with a unmarried important circulate to play and record more than one video streams, handles exclusive disk sorts and exclusive streams with exclusive bit rates (Li et al., 2003). The hassle of Broadcast report machine primarily based totally on digital garage media-command and manage protocol, is solved through the answer offered on this paper, the protocol is innovatively modified to shape a hierarchical shape that may be transmitted. To differentiate broadcast report machine from the conventional machine optimized techniques are brought for caching and receiving information from network (Zhang et al., 2004). By nature actual time multimedia files are accessed sequentially it favors the information placement in searching for style optimally.

CONCLUSION

It will now no longer be incorrect to mention that pc machine operations can't be accomplished with out garage operations. Every operation desires the help of number one or secondary memory. A massive element of performance and overall performance of working machine relies upon at the garage machine because it controls the performance of I/O also. For an green Operating machine it's far very important for it to undertake a sturdy and green garage machine to undertake. Therefore document structures are very crucial and critical. With the passage of time and converting computing desires and hardware document structures desires also are changed, consequently some of strategies are proposed and document machine has done a notable interest of researchers. According the manner of garage and implication structures the document structures fall into diverse categories; a review of a number of those has been blanketed on this Study consisting of their computing environments, overall performance traits and different parameters.

REFERENCES

- [1] Alessandro, M., L.V. Mancini and S. Jajodia, 2003. Secure dynamic fragment and replica allocation in large-scale distributed file systems. IEEE Transact. Parallel Distribut. Syst., 14(9): 195-200.
- [2] Alexandra, G. and S. Archana, 2002. A load balancing tool based on mining access patterns for distributed file system servers. Proceedings of the 35th Hawaii International Conference on System Sciences, (ICSS' 2002), 7-10 Jan. 2002, IBM Almaden Res. Center, San Jose, CA, USA, pp: 1248-1255.
- [3] Aniruddha, B., S. Smaldone and L. Iftode, 2007. FRAC: Implementing role-based access control for network file systems. Proceedings of the Sixth IEEE International Symposium on Network Computing and Applications, (SNCA' 2007), 12- 14 July 2007, Rutgers Univ., Piscataway, pp: 95-104.
- [4] Azzedine, B. and R. Al-Shaikh, 2006. Servers reintegration in disconnection-resilient file systems for mobile clients. Proceedings of the 2006 International Conference on Parallel Processing Workshops, (PPW' 2006), SITE, Ottawa Univ., Ont., pp: 6-114.

[5] Burns, R.C. R.M. Rees and D.D.E. Long, 2000. Semi-Preemptible Locks for a Distributed File System. Proceeding of the IEEE International Conference Performance, Computing and Communications Conference, 2000, (IPCCC '00), Dept. of Comput. Sci., IBM Almaden Res. Center, San Jose, CA, pp: 397-404.

[6] Benjamin, C.R., M.A. Smith and D. Diklic, 2002. Security considerations when designing a distributed file system using object storage devices. Proceedings of the First International IEEE Security in Storage Workshop, (SSW' 2002), 11 Dec. 2002, IBM Almaden Res. Center, San Jose, CA, USA, pp: 24-34.

[6] Chul, L., S.H. Baek, A.H. Park, 2008. A hybrid flash file system based on NOR and NAND flash memories for embedded devices. IEEE Transact. Comput., 57(7): 1002-1008.

Cuneyt, A. and M. Sarit, 2003. A scalable bandwidth guaranteed distributed continuous media filesystem using network attached autonomous disks. IEEE Transact. Multimedia, 5(1): 71-96.

[7] Douceur, J.R., A. Adya, W.J. Bolosky, D.R. Simon and M. Theimer, 2002. Reclaiming space from duplicate files in a serverless distributed file system. Proceedings of the 22nd International Conference on Distributed Computing Systems, Microsoft Research, pp: 14.

[8] Dalibor, P., B. Thomas, V.H. Fabio, H. David and S. Burkhard, 2009. The design and evaluation of a distributed reliable file system. Proceedings of the International Conference on Parallel and Distributed Computing, Applications and Technologies, (PDCAT' 2009), 8-11 Dec. 2009, Dept. of Inf. IFI, Univ. of Zurich, Zurich, Switzerland, pp: 348-353.

A STUDY ON CYBER SECURITY IN MODERN INTERNET OF THINGS***Ms.N.Raghavi¹****Assistant Professor**Department of Computer Science
MKJC,Tamilnadu,India****Ms.M.Ponnarasi²****Assistant Professor**Dept of Computer Science
MKJC,Tamilnadu,India****Ms.K.Suganya³****Assistant Professor**Department of Computer Science
MKJC,Tamilnadu,India*

Abstract:

Cyber Security plays a very prominent and important role in present scenario. As the world has becoming digitized all the data's are sent and received through the network and there comes a problem in security. Even though we have facilities to safe guard our own data there is an opportunity to get attacked by the Malwares or by the malicious activity performed by the hackers in many forms. The fast development of technology not only saves the time and also at sometimes it leads to the malicious activities that are happening in the network while transmission of data's. People's lifetime changes as per the technologies have developed and has greater involvement in IoTs (Internet of Things) in everyone's life. In some cases while transferring of data's there are lot of attacks occurring through the network while the data is sent from the client to the server. The most prominent attack occurs often is the firewalls are DDoS attack. The Distributed Denial of attack will be processed as user request the firewall fails to identify the attack occurring in that network. This paper discusses about the risk occurring in the network due to DDoS. It gives the awareness of the DDoS attack and its vulnerabilities occurring in the network and for secured transmission in the network.

Key Words: IoT, DDoS attack, CPS.

I.INTRODUCTION:

Nowadays data's are sent and received in any form of data it might be through e-mail or an audio or video by selecting the content and clicking the icon we are sending the data we never know weather the data are sent to the right person correctly and without leaking our information to the other third parties. Current environment is based on the upcoming technologies that change the mankind. Due to this development we forget to safe guard our information in a right way. As per the data's are sent through digital devices we also have increase incyber crimes occurring in networks. So cyber security has become a latest issue. The main motive of cyber security is not only limited to securing the information in IT industry but also to various other fields like cyber space etc. Developing cyber security and protecting information infrastructures are essential to each nation's security and economic wellbeing.

II CYBER-PHYSICAL SYSTEM

CPS this system connected to the network. It has higher performance and good efficiency in control-computing hybrids. This creates anew generation systems with physical and integrated facilities to connect with physical world. This has great advantages in certain application domains like smart grids, intelligent healthcare, industrial control systems and many other systems. Mainly in industrial control systems that are explicitly applied to smart manufacturing, are the most common CPS. They will behave as a proactive against threats to network security and also protects the data's in particular organizations but certain business cannot avoid IoT attacks from ongoing places. Implements certain block chain platforms will give security to the IoT devices. The system consists of interconnected computing devices, mechanical and digital machines, objects, animals or people that are given distinctive identifiers (UIDs).

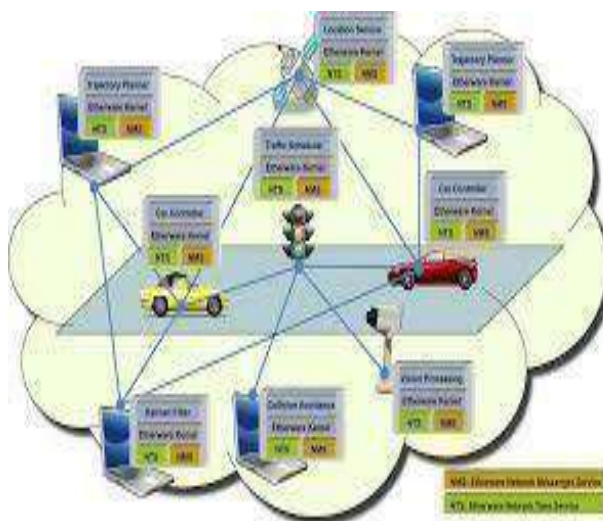


Fig:1.1. Cyber Physical System

III. DDoS ATTACK:

This Distributed Denial-of-Service (DDoS) attack happens once when more system data's are flooded with information or at a targeted system. It sends the duplicate packets and affects the bandwidth of a particular channel and that place are held up with traffic. After creating the traffic the malicious activities take place in that particular bandwidth even the firewall fails to identify the malicious activities occurred in that particular network. It creates a mass of demonstration in a website, and this process comes under the illegal activity of the Federal Computer Fraud and Abuse Act. While not an enormous share, the report suggests more than 400,000 DDoS attacks are recorded every month worldwide. A thirteen drop across six months equates to tens-of-thousands fewer attacks this year.

There are many common forms of DDoS attacks, like volume based mostly, protocol and application layer.

7 Best Practices for Preventing DDoS attacks

4. Control and block the malicious activities.
5. Safeguard your Network Infrastructure.
6. Avoid redundancies while transmissions.
7. Bandwidth size should be incremented.
8. Understand the Warning Signs.
9. Consider DDoS-as-a-Service.



Fig:2.1 DDoS Attack

IV RELATED WORK

DDoS attacks are the preplanned attacks that aims to stop the original users while accessing the system and tries to fill the bandwidth of a particular channel that are currently available at a particular point of time. It affects the whole system. The hacker does not attack to steal or access any information instead they terminate the performance of the network and the system. The application layer of this attack is very difficult to identify because they create traffic in that busy area. They have used coefficient to check the traffic and patter for being attack traffic or legitimate traffic and a six-step technique and it is designed to secure networks from DDoS attack traffic. In their article they have completely projected the DDoS attack and the reduction of attacks.

In another paper in order to prevent the attack, they have created a database with A set „R“ that contains the list of detected malicious nodes, their activities based on the attack that describes when further the same attack has been created or not and a type of attack, interaction category (Active-Passive), by identifying the attack and it can able to easily generate an APD (Active Profile Database) this database stores all the attacks and its type and also efficiency of an attack and analyzes the occurrence and behavior of the attacker nodes. This process will help us to avoid the future attacks by knowing the details of the previously attacked nodes. The existing solution has an organized list of nodes ordered as per their malicious magnitude called a blacklist table. The preventive threshold is a given integer and is represented by Ψ and represents the malicious node will be a blacklist table consisting of highest value.

V SYSTEM MODEL

Network Model:

The 100 nodes and the area are confined to 800m² for simulation that includes the early distribution of randomly and uniformly nodes with 500 m transmission range. For routing we chose AODV protocol and for MAC we chose IEEE802.11 protocol. To illustrate the propagation two ray reflection model. The network with 10 and 20 attackers has been used for simulation of ten Count Bit Rate. To simulate the traffic background different Packet sizes with different time of source activity are used. The experiments are integrated with repeated transformation of data.

7. Attacker Model

The unauthorized activities are considered as an imperfection in the security system. It will be considered as an unauthorized persons when the third party enters into our personalized data. Internet of Things most probably leads to the lack of central control, threatens the resources, bandwidth is limited when large amount of data is transmitted, wireless devices, mobility of a node, scalability etc. The most dangerous attacks are spoofing, eavesdropping, replay, and other attacks. The existing nodes within the network move freely in any direction and therefore the new nodes join the network; a number of the nodes maybe compromised by an adversary to perform some malicious behavior in the network. Every contributing element with in the IoT networks is vulnerable to internal also as external threats. As, a result the IoTs require robust security scheme to make sure the network security.

The current IDS architecture of IoT consists of three taxonomies; (i) stand-alone, (ii) co- operative, and (iii) hierarchical. In stand- alone design each node is answerable for its own security with none collaboration with the rest of the nodes within the network. On the opposite hand, in co-operative based architecture, the nodes have their own IDS systems. Whereas, in third type of architecture which is hierarchical based, the network is divided into clusters and particular nodes are chosen based upon a certain criterion as CHs cluster to perform the ID by performing the roles and responsibilities. The primary advantage of this kind of design is that the adequate utilization of the resources, however includes a disadvantage of choosing a node as a CH that is impractical in ad-hoc networks, where the nodes move freely in all directions.

B. Trust-Based Intrusion Detection

To efficiently identify malicious nodes via trust management, an important step is to select a proper trust threshold in advance. if k (the number of normal packets) increases its value, then t -value can become larger. Since k should be always smaller than N (the total number of incoming packets), t - value would fall into the interval of $[0,1]$. Thus, in the best scenario, t -value can be infinitely close to 1, indicating that a node is more credible by sending a good packets. It is worth emphasizing that a node can be regarded as malicious by only sending One malicious packet, but this may case many false positives. Trust-based intrusion detection can provide more flexibility for reducing false rates and recovering a false detected node by adjusting the threshold.

In an IOT based WSN, each node could have two major functions: sensing and relaying. Sensor nodes collect data and then send the collected data to the corresponding cluster head directly in one hop or by relaying via a multi-hop path. In this case, a CH can receive the info from different sensor nodes, then can process and compute trust values. In the end, CHs deliver the info to the bottom station. The trust computation is typically supported a period of time of t , which consists of several time units. The sensor nodes in a cluster can record the traffic including the total number of outgoing packets. After several time units elapse, the time window slides to the proper (e.g., one-time unit), and therefore the sensor nodes can drop the info collected during the earliest unit with the purpose of reducing storage consumption. Further, the cluster head can periodically request the status of a target node and thus can establish a map of trust values. In the mechanism, dangerous behavior of a node will reduce its trust price greatly. For a sensor node, its trust value are often computed by its cluster head, while for a cluster head, its trustvalue are often computed by the bottom station.

VI. CONCLUSION

The Magnitude of DDoS and therefore harm as escalated with the inclusion of various different attack sources and therefore creating suitable environment for harming the security and performance of the IoT technology. This article stresses in the possible security technique and proposed a prevention scheme that is favorable to be applied in IoT networks that are vulnerable to DDoS attacks. Based on the essential structure and functions of existing IDS, we have sued results in the proposed algorithm in a manner pertaining to time. Proposed prevention algorithm may be a multi way adjustable administratively and technically for varied security desires and is additionally adjustable according to the present info at the same time updatable blacklist table. Following this will cause generate recommendation for reaction module and thus approaching to assure the network performance, security and survivability at the time of attack occurrence.

VII. REFERENCE

- [1] T. A. Ahanger and A. Aljumah, "Internet of Things: A Comprehensive Study of Security Issues and Defense Mechanisms," in IEEE Access, doi:10.1109/ACCESS.2018.2876939 URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8519613&isnumber=6514899>
- [2] Ahamad Ahanger, Tariq. (2018). Defense Scheme to Protect IoT from Cyber Attacks using AI Principles. International Journal of Computers Communications & Control.
- [3] 13.915-926.10.15837/ijccc.2018.6.3356.
- [4] J. Zhou, Z. Cao, X. Dong, and A. V. Vasilakos, "Security and Privacy for Cloud-Based IoT: Challenges," IEEE Commun. Mag., vol. 55, no. 1, pp. 26–33, Jan. 2017.
- [5] A. V. Dastjerdi and R. Buyya, "Fog Computing: Helping the Internet of Things Realize Its Potential," IEEE, vol. 49, no. 8, pp. 112–116, Aug. 2016.
- [6] H. Lin and N. W. Bergmann, "IoT Privacy and Security Challenges for Smart Home Environments," MDPI, 2016.
- [6] K. Rose, S. Eldridge, and L. Chapin, "THE INTERNET OF THINGS: AN OVERVIEW, Understanding the Issues and Challenges of a More Connected World," 2015.
- [7] IEEE, "Towards a definition of the Internet of Things (IoT)," 2015.
- [8] L. Catarinucci et al., "An IoT-Aware Architecture for Smart Healthcare Systems," IEEE, 2015.
- [9] R. H. Weber, "Internet of Things – New security and privacy challenges," Comput. law Secur. Rev., vol. 26, pp. 23–30, 2010.
- [10] E. Oriwoh, H. M. al-Khateeb, and M. Conrad, "Responsibility and Non-repudiation in resource-constrained Internet of Things scenarios," in Conference: International Conference on Computing and Technology Innovation, 2015.

A STUDY ON ANALYSIS OF E-COMMERCE STRATEGIES FOR TRADITIONAL ORGANISATIONS MOVING ON TO THE INTERNET

SHEMA¹

*Assistant Professor,
KJC,Tamil Nadu, India*

Dr. G. DEEPALAKSHMI²

*Head & Assistant Professor,
MKJC,Tamil Nadu, India.*

A.ATHIYAVEN³

*Assistant Professor,
MKJC,Tamil Nadu, India.*

ABSTRACT OF THE STUDY

Electronic commerce is a type of business model, or segment of a larger business model, that enables a firm or individual to conduct business over an electronic network, typically the internet. E-commerce is the buying and selling of goods and services, or the transmitting of funds or data, over an electronic network, primarily the Internet. These business transactions are business-to-business, business-to-consumer, consumer-to-consumer or consumer-to-business. The term entail is used in reference to transactional processes around online retail. E-commerce is conducted using a variety of applications, such as email, fax, online catalogs and shopping carts, Electronic Data Interchange (EDI), File Transfer Protocol, and Web services. It can be thought of as a more advanced form of mail-order purchasing through a catalog. E commerce is the movement of business onto the World Wide Web.

Key words: Business to Business, Business to customer, Customer to customer, Internet, Electronic Data Interchange.

INTRODUCTION OF THE STUDY

E-commerce is a means of conducting business, where the buying or selling of goods and services or the transmitting of funds or data, occur via electronic medium. There are no physical market places and the entire process of marketing and selling of goods, takes place on-line or electronically. This means, the buyer and the seller do not often meet face to face. It is a replica of a physical market place in the virtual world.

E-commerce, also called e-trading, operates in all four major market segments – Business to Business, Business to Consumer, Consumer to Consumer and Consumer to Business. Examples of E-commerce include on-line shopping, electronic payments, on-line auctions, internet banking, on-line ticketing etc.

The term electronic commerce or e-commerce refers to any sort of business transaction that involves the transfer of information through the internet. By definition it covers a variety of business activities which use internet as a platform for either information exchange or monetary transaction or both at times.

Features of E commerce

1. Ubiquitous in nature

The customer limit of a real business is limited to the city where it is located, whereas the reach of an e commerce business is global. Ecommerce encourages customers to buy things from any corner of the world. Because of the ubiquitous nature of the E Commerce business, the working hours and geographical boundaries of catering increase. People can connect to an E Commerce business in the availability of the internet.

2. Large customer reach

Ecommerce businesses have reached to the customers living all around the world. Internet reaches national boundaries. Ecommerce provides an opportunity to businessmen to cater to the customers seamlessly without being limited by cultural and national boundaries.

3. Universal standard

One of the most essential features of an eCommerce business is the universal acceptability of the website. An eCommerce website must work upon the universally accepted systems and methods. Maintaining global standards helps the users of an eCommerce website to use the website efficiently.

4. Interactive platform

Another important feature of an eCommerce business is its interactivity. Interactivity means the two-way communication between the users of the eCommerce website and businesses. Customers feel easy to purchase from the websites where they can interact before, after, and even during the purchase. For example, when a student wants to buy an online course, he would need the assistance of the experts of the website to help him to choose a suitable level of course for him. In addition to this, businesses can know about the problem faced by the users of their eCommerce platform and can fix them in order to reduce the losses.

5. Rich in content and information

The richness of content and information means the content and information can be provided in the form of audio, video, and images on the website. For example, a customer will be impressed more if there is a video available on the website containing the information about the features of the products and information about how to use it or a hyperlink to the blog post explaining the uses of the product. Educated and conscious customers always seek such information before making a purchase decision. In addition to this, such content makes the eCommerce website attractive to the consumers.

6. Information density

With the help of technology, the cost of collecting information has reduced drastically. It has become very easy to process the information about a customer and to pass it to the destined department to use it in the future. This not only has become easy but also has become quite inexpensive to do so. Besides this, the accuracy and timeliness of the information have also improved. Using the information density, a merchant can sell the same product to a different group of people at different prices.

7. Easy to use the checkout

Internet connectivity has made the ECommerce platform easy for people to access anytime and anywhere in the world. 80% of the times people browse through your website either when they have nothing to do or just to look for something. To convert this type of users into buyers, you need to have easy to use ECommerce platform and even more easy to use checkout portal.

8. Advertisement report

If it requires too much hassle to go through the final checkout, then they might lose interest in buying the product, and you will lose business. Therefore, one of the important features of an ECommerce business is multiple payment options such as payment through credit cards, debit cards, PayPal, and cash on delivery, etc.

9. Reporting tool

The reporting tool is not an essential feature of an eCommerce business, but you can certainly use it to enhance the user experience on the eCommerce platform. The reporting tool will help you to know about the problems faced by your users immediately, and you can fix them earliest before it causes any more problems to the other users of your eCommerce platform. Reporting tool also increases the reliability of your users when they know that they will get help in case something goes wrong with their transaction.

10. Promotional and discount code tool

Another less common feature of an eCommerce platform is the promotional and discount code tool. Take the example of Zomato, Swiggy, and UberEATS. They always provide a discount and promotional code on the food ordered by the people.

11. Integrated Blog and articles section

Nowadays, people have become aware, and they make an intelligent purchase only. Integrating blogs and articles section in the eCommerce platform is another method to attract such customers. The image of your business will improve when they can learn great things from you.

Types of Ecommerce Models**1. Business to Consumer (B2C):**

When a business sells a good or service to an individual consumer (e.g. You buy a pair of shoes from an online retailer).

2. Business to Business (B2B):

When a business sells a good or service to another business (e.g. A business sells software-as-a-service for other businesses to use)

3. Consumer to Consumer (C2C):

When a consumer sells a good or service to another consumer (e.g. You sell your old furniture on eBay to another consumer).

4. Consumer to Business (C2B):

When a consumer sells their own products or services to a business or organization (e.g. An influencer offers exposure to their online audience in exchange for a fee, or a photographer licenses their photo for a business to use).

Advantages of E-Commerce

- E-commerce provides the sellers with a global reach. They remove the barrier of place (geography). Now sellers and buyers can meet in the virtual world, without the hindrance of location.
- Electronic commerce will substantially lower the transaction cost. It eliminates many fixed costs of maintaining brick and mortar shops. This allows the companies to enjoy a much higher margin of profit.
- It provides quick delivery of goods with very little effort on part of the customer. Customer complaints are also addressed quickly. It also saves time, energy and effort for both the consumers and the company.
- One other great advantage is the convenience it offers. A customer can shop 24×7. The website is functional at all times, it does not have working hours like a shop.
- Electronic commerce also allows the customer and the business to be in touch directly, without any intermediaries. This allows for quick communication and transactions. It also gives a valuable personal touch.

Disadvantages of E-Commerce

- The start-up costs of the e-commerce portal are very high. The setup of the hardware and the software, the training cost of employees, the constant maintenance and upkeep are all quite expensive.
- Although it may seem like a sure thing, the e-commerce industry has a high risk of failure. Many companies riding the dot-com wave of the 2000s have failed miserably. The high risk of failure remains even today.
- At times, e-commerce can feel impersonal. So it lacks the warmth of an interpersonal relationship which is important for many brands and products. This lack of a personal touch can be a disadvantage for many types of services and products like interior designing or the jewelry business.

- Security is another area of concern. Only recently, we have witnessed many security breaches where the information of the customers was stolen. Credit card theft, identity theft etc. remain big concerns with the customers.
- Then there are also fulfillment problems. Even after the order is placed there can be problems with shipping, delivery, mix-ups etc. This leaves the customers unhappy and dissatisfied.

Examples of E-Commerce

- Amazon
- Flipkart
- eBay
- Quikr
- Olx

Benefits of E commerce

- Increased comfort - transactions can be made 24 hours a day, without requiring the physical interaction with the business organization
- Time saving- Customer can buy or sell any product at any time with the help of internet.
- Quick and continuous access to information-Customer will have easier to access information check on different websites at the click of a button.
- Convenience-All the purchases and sales can be performed from the comfort sitting a home or working place or from the place a customer wants to.
- Switch to others companies-Customer can easily change the company at any time if the service of a company is not satisfactory.
- Customer can buy a product which is not available in the local or national market, which gives customer a wider range of access to product than before.
- A customer can put review comments about a product and can see what others are buying or see the review comments of other customers before making a final buy.

Challenges in E Commerce

- Private and public corporation is not involved jointly to grow the business of e-commerce. Private and public joint initiative is needed to develop the ecommerce business. Joint initiatives bring credibility inside people, which is needed for flourishing the ecommerce business.
- There is a lack of system security, reliability, standards, and some communication protocol.
- Customer loses their money if the website of ecommerce site is hacked. Most common problem of e-commerce website is not having enough cyber security.
- Financial institutions and intermediaries: Thus far, financial institutions and banks in developing countries are hesitant to take an active role in promoting e-commerce. However, merchants need the involvement of banks to broaden the reach and appeal of ecommerce and to help prevent fraud and potential losses attributable to credit card fraud.
- In developing countries there is a culture of buying product by negotiating price with seller, which is not easily possible in case of e-commerce in developing countries because of lack of infrastructure facility.
- One of the biggest challenges is the cutting down the price of internet. Authorities are trying to keep low the price of bandwidth low. But the high cost of spreading networks and operating expenses hinder to keep price low for internet.
- Trust is the most important factor for the use of the electronic settlements. Traditional paper based rules and regulations may create uncertainties the validity and legality of e-commerce transactions. Modern laws adopted and impartiality implemented in the electronic transactions form the basis of trust in the developed world. Where legal and judicial systems are not developed ecommerce based transactions are at a disadvantage because of lack of security whether real or perceived.

- New methods for conducting transactions, new instruments, and new service providers will require legal definition, recognition, and permission. For example, it will be essential to define an electronic signature and give it the same legal status as the handwritten signature.

OBJECTIVES OF THE STUDY

- To understand the present status and trends of E-Commerce
- To reveal the key variables influencing the increased usage of E-Commerce.
- To examine the main products dealt on e-commerce
- To make out the length of supply chain.
- To understand the benefits and Challenges of e-commerce.

REVIEW OF LITERATURE

Mitra Abhijit (2013) suggests E-Commerce has unleashed yet another revolution, which is changing the way businesses buy and sell products and services. New methodologies have evolved. The role of geographic distances in forming business relationships is reduced. E-Commerce is the future of shopping. With the deployment of 3G and 4G wireless communication technologies, the internet economy will continue to grow robustly. In the next 3 to 5 years, India will have 30 to 70 million internet users which will equal, if not surpass, many of the developed countries.

Chanana Nisha and Goele Sangeeta (2012) propose that the future of E-Commerce is difficult to predict. There are various segments that would grow in the future like: Travel and Tourism, electronic appliances, hardware products and apparel. There are also some essential factors which will significantly contribute to the boom of the E-Commerce industry in India i.e. replacement guarantee,

M-Commerce services, location based services, multiple payment option, right content, shipment option, legal requirement of generating invoices for online transactions, quick Service, T & C should be clear & realistic, the product quality should be same as shown on the portal, dedicated 24/7 customer care centre should be there.

Awais Muhammad and Samin Tanzila (2012) indicate that use of internet has made the world a global village. The use of Internet has reduced the distances and brought the people together. A nation's back bone is commerce and it will be strengthened if backed by electronic tools in which e-commerce plays a vital role. The important feature in ecommerce is privacy which not only increases competitive advantage but confidence level also.

Dutta and Dutta, (2009) found tangibles have the highest impact on overall customer satisfaction. The largest discrepancy between the customer expectations and perceptions is in terms of empathy which includes Bank locations and ATM machines in convenient places and telebanking and internet banking facility. The study regards this major source of concern for Indian banking industry as a huge service quality gap exists for all the banks in this category.

FINDINGS OF THE STUDY

- Almost 90% of the people are using Cashless digital transactions
- Increased investments of big retail brand in logistics, wholesale retail and market share expansion.
- Increased B2B e-commerce.
- Internet content available in local languages with the help of artificial intelligence.
- Growth of consumer demand in non-metro cities

SUGGESTIONS

- Research the e-commerce space and find your niche
- Select your business name and choose a legal structure
- Obtain business permits and licenses
- Choose your E commerce Platforms
- Market your business in a unique way

CONCLUSION

The e-commerce industry will be a leader with popularity in electronic business world in the upcoming years. The e-commerce revolution has fundamentally changed the business of transaction by giving new opportunities and breaking borders easily. In India, it has strongly impacted the traditional business system and changing the life of people by making it easier. While it gives benefits to customer and seller, e-commerce gives challenges to traditional business for competitive position.

Developing countries face many obstacles that affect the successful implementation of e-commerce with the help of comparing with developed country. When the internet cost will be low then the e-commerce will flourish easily and will make many of traditional business to run out of their business. Convenience is one of the benefits that customer gets from the e-commerce and thus increasing customer satisfaction. This is due to customer can place a purchase an order from anywhere with internet connection.

E-commerce business provider should give importance on every customer by giving smooth service and many options for payment and have more functions available online. Other benefits are expanded product offerings and expanded geographic reach. But e-commerce business faces a lot of challenges in flourishing their business.

REFERENCES:

1. <https://www.google.com/search?q=TIPS+TO+START+A+E+COMMERCE+BUSINESS&oq=TIPS+TO+START+A+E+COMMERCE+BUSINESS&aqs=chrome..69i57j0i22i30l4j0i10i22i30j0i22i30l4.16084j0j15&sourceid=chrome&ie=UTF-8>
2. <https://www.adb.org/sites/default/files/publication/675186/ewp-632-technology-adoption-b2c-e-commerce-asia.pdf>
3. Text Book: Building a story Brand- Donald Millar

AN OVERVIEW OF CYBER SECURITY ON SOCIAL MEDIA: ISSUES, CHALLENGES AND SOLUTIONS

Dr. K. Priya¹,

*¹Department of Computer Science,²
Marudhar Kesari Jain College for Women,
India.*

P. Vishnu Priya²

*Department of Computer Applications,
Marudhar Kesari Jain College for Women,
India.*

ABSTRACT

Cyber security is an important role in the Information Technology(IT) protecting the information has become an enormous issue in the day-to-day life. Social networking sites are not only to interact or communicate with other people globally, but also one important way for business process. **Many government and companies are taking several measures in order to rectify these cyber crimes.** Cyber security is the practice of ensuring security to the user information and networks from unauthorized people and access. The security concerns like issues, challenges and solutions are on social media like identity misuse, malware, phishing attacks and third party application threats have also been discussed. This paper presents an overview of cyber security and also discusses about the security issues in Social Media.

Key Words: Cyber crimes, Cyber security, Social media

INTRODUCTION

Today internet is rapidly and growing technology in the daily life of individuals. In the technological era many social companies have leaked the privacy of their users. They have been critically selling off private information to third parties. Then **the latest technologies like, online banking, mobile computing, E-commerce,** could have high level of security. Cyber security becomes the main role. The cyber security has become the most important field under research. Cyber security enrolls preserving **Confidentiality, Integrity and Availability (CIA)** of information in the cyberspace. The security gives classification of the domains. This relationship between various domains is given in figure 1.

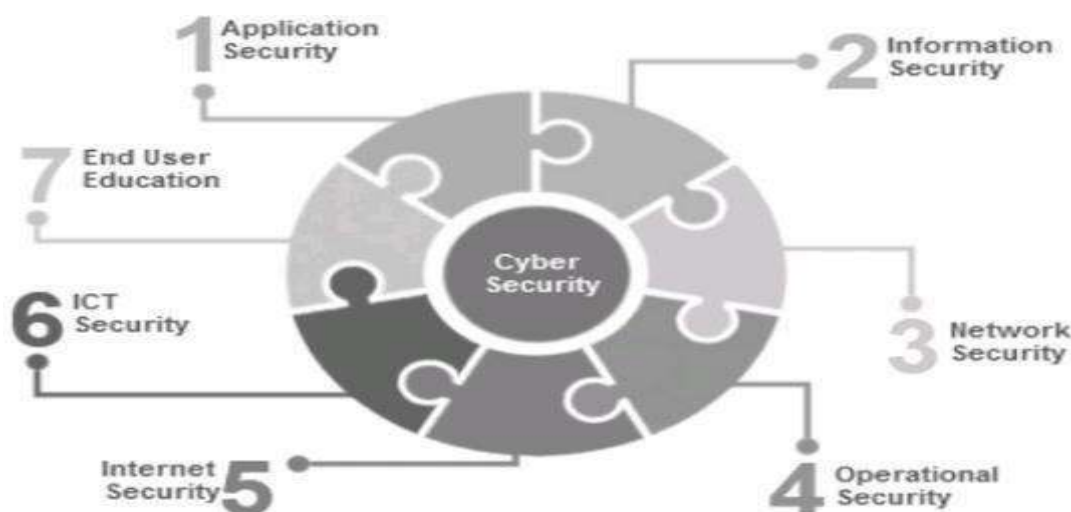


Fig.1: Cyber Security in Various Domains

Application security: It takes various measures to improve the security application. This often done by developing, adding and testing to preventing security vulnerabilities. **Information security:** It practices to main integrity, confidentiality, and business data, to prevent the information to unauthorized individuals or system. **Network security:** It is a simplest term and it always includes both software and hardware. **Operation security (OPSEC):** It is a process of identifying historical and real time security events. And it prevents sensitive information. **Internet security:** It includes various processes like Browser security, web site security, operation system and network security. **ICT Security:** **ICT stands for information communication technology.** Everyday usage of digital technology includes tablet, mobile phone, email, browse etc. **End user education:** The lack of end user knowledge about cyber security is risk reason for 50% of the cyber attack and almost 90% of cyber attack caused by human behaviors.

II CYBER CRIME

A Cyber crime defines illegal activity where computer or social media is either a tool or target or both. And it involves the use of electronic communication or information system. Usage of computer and other main technology in daily life is growing fastest. The oxford dictionary defined the term cyber crime as “**criminal activities carried out by means of computer or the internet**”.

Cyber Crimes Types

Cyber Staking

It is a new form of internet in our society. Stalking means behavior of harassing. It doesn't physically follow is victim which is committed over the online with the use of information technology. The definition of cyber stalking is universally acceptable as it migrate according

to professor Lambert defined stalking crime as follow:”Any person is guilty of the stalking crime who: Maliciously, trace another person, willfully with the intent of placing that person is reasonable fear of death, Sexual assault, or great bodily injury to that person, any member of that persons family, or anyone with whom that person has intimate relationship”. Cyber Stalking now spread its wings to social networking with the use of social media. It may includes false accusation, slander, libel, and mental. The fact that cyber stalking does not involve physical contact. May create the misbehavior that is more than physical stalking. There are three ways of cyber stalking Email, internet and computer.

Cyber Terrorism

Everyone has suffered due to violent act of terrorism. It took the form of attacks on network, telecommunication, and computer system. It includes function like instigation, planning, and spying following types reports terrorist group daily recourse to the internet. Terrorism in cyberspace is mentioned, Penetration of computer networks, Governments’ network and social media.

Dorothy Denning, a professor of computer science, has put forward definition before the house armed service committee in May 2000; “cyber terrorism is the convergence of cyber and terrorism. It refers to unlawful attacks and threats of attacks against computer, networks and the information stored there in when done in intimidate or coerce a government or its people in furtherance of political or social objectives. Further to cyber terrorism, an attack should result in violence against persons or property, or at least cause enough harm to generous fear, attacks that leads to death or bodily injury, explosion, or serve economic loss would be example serious attacks against critical infrastructures could be acts of cyber terrorism, depending on their impact. Attacks that disrupt nonessential service or that are mainly a costly nuisance would not”.

Hackers

A person is likely to cause wrongful loss or damage to the public or any personal information deletes or destroy or alter any information residing in a computer source. The jargon dictionary said the term “**CRACKERS**” is used to “**BENIGN**” hackers from hackers who maliciously cause damage to targeted computer. “**CRACKERS**” is defined as hackers who still information located on computer.

Different Types of Hackers

Black hat hackers (**bad guys**): They are the types of hackers who may still or modify or delete data. Black hats are individuals who use excellent computing skills for illegal purposes. This category of hackers mostly involved in illegal activities. They are also known as crackers.

White hat hackers (**Ethical hackers**): As opposed to the black hat these types of hackers used their hacking skills for good reason. White hat hackers are the cyber security who helps to government and by including penetration testing and identifying loopholes in their cyber security.

Red hat hackers: Almost like white hat hackers and the main objectives for red hat hackers is to find black hat hackers and destroy all the schemes.

Green hat hackers: There the set of individual's .they want to observe and learn about the world of hacking. To watch videos, tutorial about hacking.

Gray hat hackers (**good and bad**): These types of hackers fall in between white hat and black hat hackers. Gray hat hacker have both good and bad intentions. And doesn't use their skill for personal gain, also he is not legal to authorized are hack the organization. Either he is not a white hat hacker.

Blue hat hackers (**revenge**): These types of hackers are very similar to script kiddies main agenda is to take revenge on anyone who makes them angry. Blue hat hackers use simple cyber attacks like flooding your IP address with overloaded packets.

Social Engineering hackers: These types of hackers use psychological manipulation to make people to private contents. It is most popular complex scheme.

Suicide hackers (**Cause**): This type of hackers does not worry about jail term or any other kind of punishment. Suicide hackers are sacrificing their life for an attack.

Malicious insider: A malicious insider or a whistle blower could also be an employee with a grudge these types of hackers main role within the corporate to hack the system.

Elite hackers (**geniuses**): These types of hackers are the "**cutting –edge geniuses**". They are the real experts in the hacking.

Protect Against Yourself from Hackers

Make sure while your security software is up-to-date. Your computer will notify you when a software update is available. Use strong password for all devices. Don't open email from anonymous sending .learn about your internet privacy and secure your wireless network. Use multifactor authentication on your accounts. Uninstall unnecessary software, avoid being scammed, Call the right person for help.

III CYBER SECURITY

The term or word cyber security is used to refer to the security given on on-line services to protect your information. It is also known as **Information technology security or network information security**. Our world today is ruled by technology where all information is maintained in a

digital or a cyber form. Cyber security is necessary it helps in securing data from hacking and various other attacks.

Need of Cyber Security

Cyber security is needed because it protects information and system from major threads. It includes **personally identifiable information (PII)**, data, and malicious attack, Intellectual property. Its main role is to defend those assets against all attack throughout the entire life cycle of a cyber crime.

Newly Emerged Cyber Security

Phishing

It is a common form of cyber threat to steal sensitive data illegally like credit card details, login information and email attachments, websites etc. and it is specially carried out by spoofing. Another definition of phishing is a method of stealing personal data where by an authentic looking email is made to appear as if it is common from a real company or institution, the idea is to trick the recipient into sending secret information such as all inform or login data to the scammer”.

Ransom Ware

It is a form of malware that encrypts an employee's document. And also spread through phishing. Ransom ware will search for valuable data and send copies to the criminals another specialty of ransom ware can spread and access different wireless networks. **Example: Crypto locker.**

Automatic Hacking

This is typically done by cloning the signals that a key can communicate with each other. In the modern technology vehicles are packed with automated tools and creating seamless connectivity for drivers. Self driving vehicles use critical mechanism that requires cyber a

IoT with 5G Wireless Communications

5G Architecture is new to fields and more researchers to find loopholes to make the computer system secure from criminal threads. The emerging **IoT-5G(Internet of things) 5G** scenario extends sensor belong IOT capabilities to robots. Daily usage of Google chrome, browser was found to have flaws.

Data Breaches

A small industry or company may suffer a data breaches. Safeguarding an information system is enormous roll now- a days. Where the data is taken from a computer system without any permission. **Recording key strokes, (DDOS) are the Distributed denial of services types of data breaches.**

A Zero-Day Attack (Vendor or Developer)

A skilled person use code to exploit zero days. Sometimes it is discovered by vendor because it behaves suspiciously it is also knew as software patch. But it is not illegal to sell zero days but it would be careful.

IV CYBER SECURITY TOOLS**Firewall**

It plays an important role in detecting malware .a firewall is a software program that helps screen out hackers, viruses that reach on your computer over the internet. All messengers sending or receiving the internet passes through the firewall present. And it monitors into traffic network it decided to allow or block specific traffic based on a set of security rules.

Antivirus Software

Antivirus software is software basic necessity for every system and it monitors all the activity on your computer system. Most antivirus includes an auto update .it used to prevent, detect, and remove from the computer. Example: McAfee, Norton and kapersky.

Penetration Testing

It is a piece of software main goal to identify security weakness in a network then it is a job for good people with the ability to do bad things.

V SOCIAL MEDIA

Social media is a system based technology. And it plays crucial role in cyber security these social websites have had positive and negative thoughts more people spend and waste time on using their websites which result in losing their colleges or jobs and families! Individuals make many mistakes when using social media services such as network access, transfer sensitive information etc. increase amount of time in social media can leads to social anxiety, depression, and cyber bullying. Social media are face book, Youtube and instagram.

Problems On Social Media

The more time spend on social media can leads to depression, anxiety. it is quite easy to enter on someone's life through a simple message on the mobile phone, email, computer and not checking any facts wasting time on social media always dependence on online relationship.

Tips On Using Social Media Platform Safely

Be very careful about putting too much information into your status update. Choose strong passwords of eight characters by using a combination of symbols, words and figures, should be kept for online activities like online banking. Avoid using your email id, login name, last name, date of birth, month of birth or any such personal information as your passwords that can be traced easily.

VI CONCLUSION

Technology has become an important of our modern life. Cyber attacks also increasing over. Social media become one of the most communication machines among adults and teenagers. Social media has become a prime target for cyber-crime. They also have to be touch with the latest updates and finding more about information and cyber security. Many social media people can adapt themselves to the situation. And gain knowledge on cyber security protects them. Cyberspace is becoming a significant area for crimes, so there is a need for comprehensive collaboration among nations to work together and combat these social network security and social media cyber attacks, which is a continuously growing threat. This article is to know about crimes and how to protect our self against criminal attacks. In our digital technology internet had a massive impact on nearly every human life such as psychology, sociability, and economic condition. Individuals have to be more aware of the cyber security events and news.

REFERENCE

1. Das, Rituparna, and Mayank Patel. "Cyber Security for Social Networking Sites: Issues, Challenges And Solutions." International Journal for Research in Applied Science & Engineering Technology (IJRASET) 5.4,833-838 (2017).
2. Ajay Sarangam, "Social Media Cyber Security: An Overview in 3 Easy Points" jigsawacademy (Feb 2021)
3. Dr. R. Hemalatha & K. Devipriya, "Cyber Security in Social Networking Site: A Short Survey", Journal Of Critical Reviews Issn- 2394-5125 Vol 7, Issue 14, 2020
4. <http://www2.trustware.com/rs/815-RFM693/images/2015/trustwaveglobalsecurityreport.pdf>
5. <http://cybercrime.org.za/defination>
6. <http://www.oxforddictionaries.com/defination/english/cybercrimes>
7. <http://www.pandasecurity.com>
8. <http://www.sociosite.org/cyberstalking.en.php>
9. <http://www.cisco.com>
10. <http://www.info-savvy.com>
11. <http://www.simplylearn.com>
12. <http://www.stanventures.com>
13. <http://study.com>
14. <http://onliledegree.und.edu>
15. <http://forcepoint.com>
16. Anti-phishing working group (sep,2013),phishing activity trends
17. <http://antivirus.comodo.com>

DIGITAL MARKETING IN ECOMMERCE IN INDIA

Ms. P.Priyanka¹.

*Assistant Professor, Dept of Computer Applications.
Marudhar Kesari Jain College for Women, Tamil Nadu, India.*



ABSTRACT

Digital marketing is a avenue of electronic communication which is used by marketers to endorse the goods service towards the marketplace. The supreme purpose of the digital marketing was concerned with consumers and allows the customers to intermingle with the product by virtue of digital media. Digital marketing includes affiliate marketing, including search engine optimization, article marketing, blog marketing, pay-per-click

search engine advertising, and e-mail marketing. Digital marketing extends beyond internet marketing including channels that does not require the use of Internet. It includes mobile phones (both SMS and MMS), social media marketing, display advertising, search engine marketing and many other forms of digital media In future, the scope of the digital-marketing was very wide and it's going to be the life blood of business.

Keywords: Digital Marketing, Digital media, email marketing, CustomerReach

INTRODUCTION

Digital marketing is often referred to as 'online marketing', 'internet marketing' or 'web marketing'. The term digital marketing has grown in popularity over time, particularly in certain countries. In the USA online marketing is still prevalent, in Italy is referred as web marketing but in the UK and worldwide, digital marketing has become the most common term, especially in the year 2013.

Digital is the process of marketing a brand using the any form of electronic device with or without the Internet. It includes both direct response marketing and indirect marketing elements and uses a range of technologies to help connect businesses to their customers. Advertising mediums that might be used as part of digital marketing strategy of a business could include promotional efforts made via Internet, social media, mobile phones, electronic billboards, as well as via digital television and radio channels. Digital marketing is a sub branch of traditional marketing and uses modern digital channel for the placement of productising. downloadable music, primarily for communicating with stakeholders e.g. customers and investors about brand, products and business progress.



Objective

- The main objective of this paper was to recognize the usefulness of Digital Marketing in the Competitive marketing environment.
- To study the impact of digital marketing on consumers purchase

Methodology Applied

- 1) **Primary Data:** Therese arches one through observation and collection of data through questionnaires.
- 2) **Secondary Data:** Secondary data is collected from journals, books and magazines to develop the theory.
- 3) **Sample Size:** The sample size is determined as 100 respondent's opinion from the customers who presently purchasing product switch a help of digital marketing.

Difference between Traditional Marketing and digital marketing:

Table 1: The following table lists a few points that differentiate between Traditional marketing vs Digital marketing

TRADITIONAL MARKETING	DIGITAL MARKETING
Communication is unidirectional. Means, a business communicates about its product subservices with a group of people.	Communication is bidirectional. The customer also can ask queries or make suggestions about the business products and services.
Medium of communication is generally phone calls, letters, and Emails	Medium of communication is Mostly through social media websites, chat, and Email.
Campaigning takes more time for designing, preparing, and launching.	There is always a fast way to develop an online campaign and carryout changes along its Development. With digital tools, campaigning is easier.
It is carried out for a specific audience throughout from generating campaign ideas up to selling a product or a service	The content is available for general public .It is then made to reach the specific audience by employing search engine techniques.
It is conventional way of marketing; best for reaching local audience	It is best for each global audience.
It is difficult to measure the effectiveness of a campaign.	It is easier to measure the effectiveness of a campaign through analytics.

Advantages of Digital Marketing to Consumers and Analysis

Digital marketing technologies permit the customer to keep on with the company information rationalized. These days a lot of customers can way in internet at any place which ever time and companies are constantly updating information regarding their goods or services. Customers know how to visit company's website, examine with reference to the products and make online purchase and afford feedback. Consumers get complete information related to the products or services. They can make comparison with other related products. Digital marketing allows 24 hours of service to make purchase for the consumers. Prices are transparent in the digital marketing.

Various elements of digital marketing



2 Low cost:

Marketing and advertising cost is one of the biggest financial burdens that businesses have to bear. While big businesses may not have so much trouble doling out millions for marketing and advertisement, for small businesses, this may be impossible or with just a few bucks you can subscribe to an email marketing provider and send transactional or direct emails to thousands of customers on your mailing list.

3 Huge return on investment:

Nothing matters more to a business than the return on the investment it makes. Digital marketing offers a substantial return on small investments. Email marketing or running advertising campaigns on social media platforms cost little when compared to traditional marketing techniques.

4 Easy to measure:

The success or otherwise of a digital campaign can easily be ascertained. Compared to traditional methods where you have to wait weeks or months to evaluate the veracity of a campaign, with a digital campaign you can know almost immediately how an ad is performing.

5 Easy to adjust:

The knowledge of the performance of an ad will inform a business on how to proceed. For an ad campaign that is performing well, it is easy to invest more in it with just a click. But for an ad that is not delivering as expected, it can be adjusted accordingly or stopped altogether with ease.

- 6** In traditional marketing, it is more rigid as a mountain of paperwork must be signed and in most cases, these agreements are on a long or medium-term basis. So even if a campaign is not performing as hoped, subscribers will still have to wait for the expiration of the agreed period.

7 Brand development:

Businesses can use their digital platforms to build their company's brand and reputation. A well-developed website, a blog featuring quality and useful articles, a social media channel that is highly interactive are some of the ways by which a business can build its brand.

8 Easy to share:

- 9** Most digital marketing channels feature sharing capabilities that allow campaigns and articles to be shared with multiple followers. This helps to create a multiplier effect and has the capacity to tremendously improve sales results.

10 Precise targeting:

The traditional means of marketing uses the spray and pray method, where an ad runs over a platform with substantial reach with the hope that a few people who love what they see, hear, or read would make a positive approach. Marketing over digital platforms, on the other hand, allows for targeted campaigning where ads are presented to customers based on their preferences or initial action.

11 Global:

The world, they say, has turned into a global village. This has been made possible via digitization. Digital marketing allows ad campaigns to be visible in any part of the world. This gives small start-ups the rare opportunity to go global via the immense exposure provided. The internet abounds with stories of start-ups that became hugely successful over a small period due to the opportunities created thanks to the global nature of marketing via digital platforms.

12 Segmentation:

Not only does marketing over digital platforms allow campaigns to be targeted at specific customers, it also allows for customer segmentation. Segmentation is the process in which large customer groups are further broken down into smaller groups of customers according to a particular classification. Segmentation increases the chances of sales, as well as cut down on cost.

- 13** For segmentation to work effectively, adequate information must be collected from subscribers.

OUTCOMES OF DIGITAL MARKETING**A. Stay updated with products or services**

Digital marketing technologies allow the consumers to stay with the company information updated. Nowadays a lot of consumer can access internet any place anytime and companies are continuously updating information about their products or services.

B. Greater engagement

With digital marketing, consumers can engage with the company's various activities. Consumers can visit company's website, read information about the products or services and make purchases online and provide feedback.

C. Clear information about the products or services

Through digital marketing, consumers get clear information about the products or services. There is a little chance of misinterpretation of the information taken from sales person in a retail store. Easy comparison with others Since many companies are trying to promote their products or services using digital marketing, it is becoming the greatest advantage for the customer in terms that customers can make comparison among products or services by different suppliers in cost and time friendly way. Customers don't need to visit a number of different retail outlets in order to gain knowledge about the products or services. Share content of the products or services

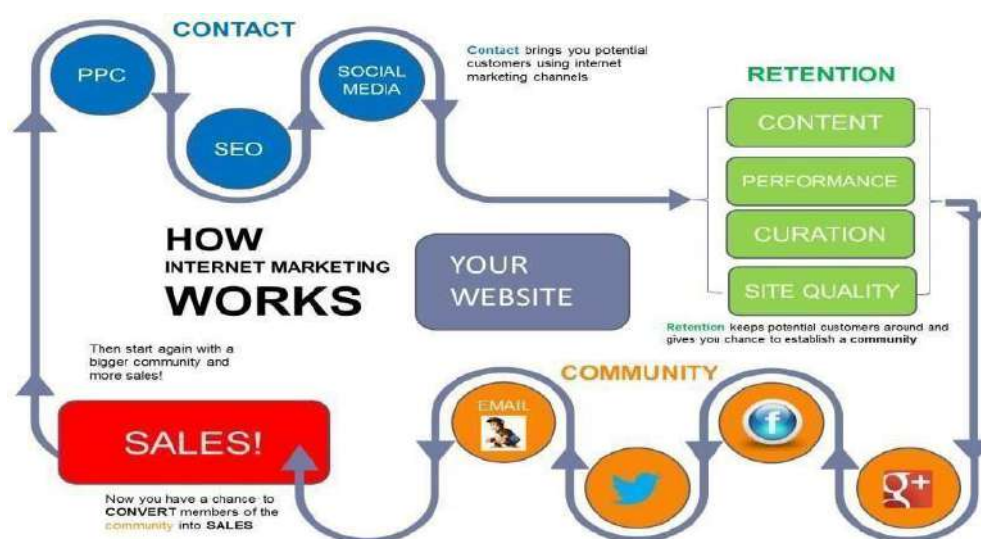
D. Apparent Pricing

Company shows the prices of products or services through digital marketing channel and this makes prices very clear and transparent for the customers. Company may regularly changes the prices or gives special offers on their products or services and customers are always in advantages by getting informed instantly by just looking at anyone mean of digital marketing.

E. Enables instant purchase

With traditional marketing, customers first watch the advertisement and then find relevant physical store to purchase the products or services. However, with digital marketing, customers can purchase the products or services instantly.

In this following diagram shows how internet digital marketing works



WHAT IS YOUR ANNUAL BUDGET FOR INFLUENCER MARKETING IN 2019?

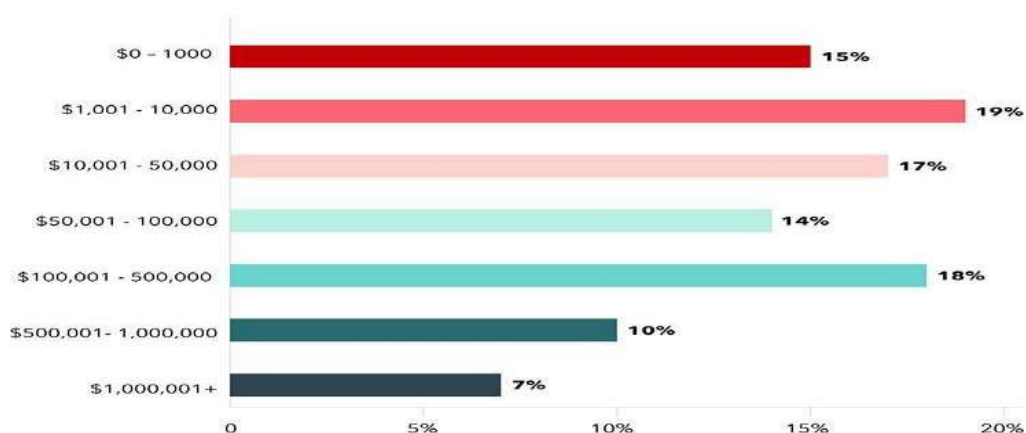


Table2:Profile of the Online Buyers

	Category	Number of Respondents	Percentage of Respondents
Gender	Male	70	70%
	Female	30	30%
	Total	100	100%
Age	Below18 Years	17	17%
	19-30years	25	25%
	31-45years	33	33%
	Above45years	25	25%
	Total	100	100%
Profession	House Wife	11	11%
	Employee	50	50%
	Business	23	23%
	Students	8	8%
	Any other	8	8%
	Total	100	100%
Monthly Family Income(in Rs.)	Below10000	21	21%
	10001-20000	49	49%
	20001-40000	25	25%
	Above40000		5%
	total	00	00%

Table 3.Awarenessof Online Shoppers

Particulars	Number of Respondents	Percentage of Respondents
Having knowledge about Online shopping	00	00%
Not having knowledge about Online shopping	-	-
Total	00	00%

Table4: Availability of Online Information about Product

Particulars	Number of Respondents	Percentage of Respondents
Excellent	54	54%
Good	38	38%
Average	7	7%
Poor		%
Total	100	100%

Table5:Reasons for Choosing Online Shopping

Particulars	Number of Respondents	Percentage of Respondents
Wide variety of Products	23	23%
Easy buying Procedures	38	38%
Lower Prices	19	19%
Various Modes of Payments	14	14%
others	6	6%
Total	100	100%

Table6:FrequencyofOnlinePurchasing

Particulars	Number of Respondents	Percentage of Respondents
Purchase once Annually	12	12%
2 -5PurchasesAnnually	46	46%
6-10PurchasesAnnually	26	26%
11Purchasesand above Annually	16	16%
Total	100	100%

Findings

- Digital marketing have a greater future in the present market.
- Consumers are satisfied through purchasing digital marketing.
- People find it safe mode of online purchase.
- Ratio of male customers is very high in online shopping that is 70%.
- Awareness about online shopping is 100% among the respondents.
- Income of respondents mainly falls in the range of Rs. 10,001 to Rs. 20,000 that is 49%.
- Employees of various companies are purchasing more than others through online shopping that is 50%.
- Most numbers of respondents that is 38% feels that online shopping have simple buying procedures; others feel that they can have a broad variety of products, products with lower price, a variety mode of payments etc.
- 54% of respondents feel that availability of online information about Product & Services is outstanding.
- 46% of the respondents purchase the products 2 to 5 times annually.

CONCLUSION

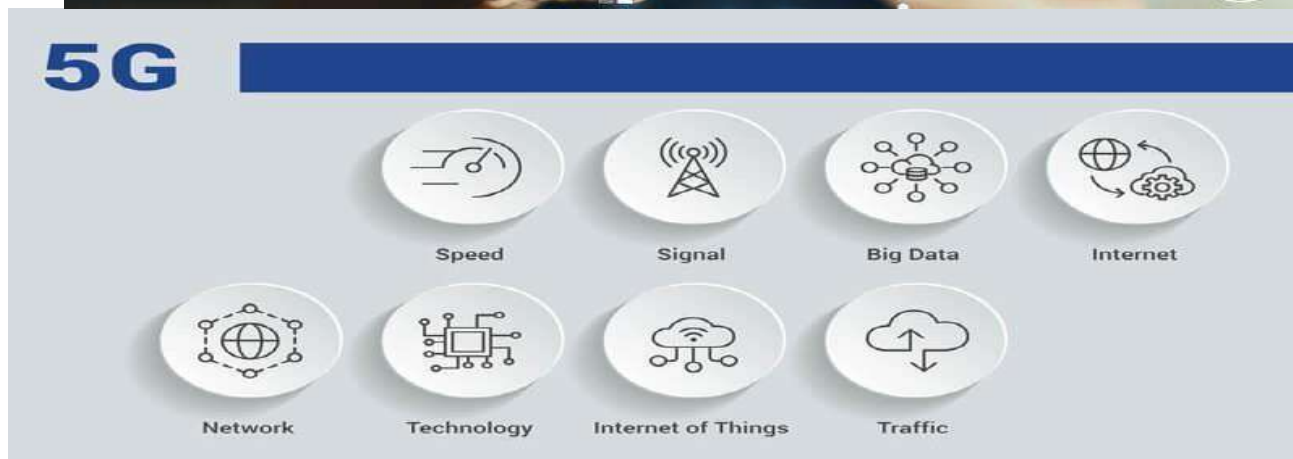
Digital marketing has turned out to be a crucial part of the approach of many companies. At the present time, still for tiny business proprietors at hand have an extremely expensive and competent method by using digital marketing to market their products or services in the society. It has no restrictions. Company can utilize any devices such as tablets, smart phones, TV, laptops, media, social media, E-mail and lot other to support company and its products and services. Digital marketing may achieve something more if it considers consumer desires as a peak priority. Digital marketing requires a dynamic approach in order to be successful in the competitive world of the digital market place and this requires constant innovations in technologies as well as developing competitive strategies to maintain and increase return on investment.

REFERENCES

1. Chaffey D & Smith P, E-Marketing Excellence: Planning and Optimizing Your Digital Marketing, Rout ledge. FourthEdition,2008,580-593
2. <https://prowly.com/magazine/advantages-of-digital-marketing-over-traditional-marketing/>
3. Wag mare GT, E-Commerce, A Business Review and Future Prospects in Indian Business. Internet, Marketing in India. Indian Streams Research Journal, 2(5), 2012, 1-4.
4. G.T.Waghmare,2012).E-commerce;ABusinessReviewandFutureProspectsinIndian Business. Internet Marketing in India. Indian Streams Research Journal, vol. 2, no. IV,(pp. 1-4.

5G NETWORK SECURITY TECHNOLOGY AND IT'S CYBER ATTACKS

*Ms.T.Subhashini,
Assistant Professor,
Dept of Computer Applications.
Marudhar Kesari Jain College for Women, Tamil Nadu, India.*



ABSTRACT

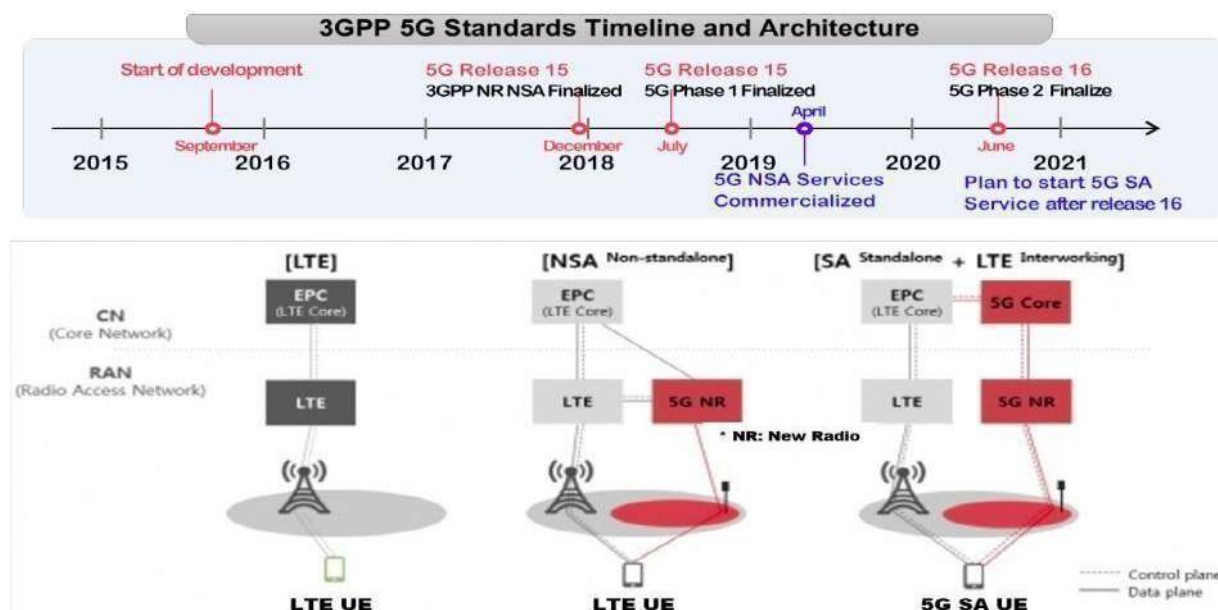
Fifth-generation technology [5G] services commercialized in 2019 have not only provided voice and data communication but also undergone significant structural changes in mobile networks to accommodate Internet-of-things devices, sensitive to latency and reliability, by adopting the latest ICT technologies, such as software-defined networking/network function virtualization, multi-access edge computing, and network slicing. However, this technological

evolution poses new security challenges, such as creation of new access paths, owing to its complex inter-operation structures, security downgrading, and limitations in security visibility. To address these issues, research on 5G security threats and security architecture has been actively underway at international standards organizations, communication carriers, and universities. However, security researchers find it difficult to conduct studies on 5G security technology design and application methods owing to the relatively unknown nature of the mobile carrier network. Therefore, in this paper, we analyzed five new security issues for each 5G section, relative to 5G technical advantages, by reviewing previous studies. In addition, we classified cyber attacks against nine network protocols primarily used in the 5G core network. The result of this study is expected to be used as basic data for modeling 5G security threats.

Keywords: 5G security, Internet of things, 5G security technology, Cyber attacks, 5G threats

INTRODUCTION

The fifth-generation [5G] mobile network is a wireless communication standard technology, established by 3GPP, and its official name defined by the International Telecommunication Union (ITU) is IMT-2020. To respond to the drastic increase in mobile traffic and Internet-of-things (IoT) devices, 3GPP promoted standardization of 5G technology from 2010, as shown in Figure 1, and completed the first stage 5G Release 15 standards in 2018. However, with Release 15, non-standalone (NSA) commercial services (i.e., a structure where user equipment and base stations are based on 5G technology, yet the core network is connected to the 4G core network) have been launched, which provides enhanced mobile broadband [EMBB] services by applying 4G technology and a part of 5G technology. In addition, the second standardization (Release 16) of the standalone (SA) structure reflecting ultra-reliable and low latency communications [URLLC] and massive machine type communications [MMTC] service requirements is underway, with the aim of establishing standards in the first half of 2020[28]. Until 4G mobile communication, mobile networks have been developed to improve the performance of data transmission speed and the capacity of smart phone devices. However, 5G technology is evolving with the consideration of various new IT technologies to build a mobile network environment that can accommodate the features and service requirements of various IOT devices, ultimately realizing a hyper connected society, which utilizes AI and autonomous vehicles.



By providing a faster “mobile communication environment,” where there is no difference between wired and wireless networks in terms of data transmission and reception capacity and speed, an “IOT communication environment” can be implemented. Such an environment can offer realistic multimedia content, such as 4K/8K and AR/VR, through new devices, such as AR, VR, drones, and smart phones, while simultaneously ensuring low power consumption in IOT devices and stability of services, even in environments where many devices are connected[27]. ITU-R classified the three major services of 5G mobile network technology into ultrahigh speed and large capacity (EMBB), URLLC, and MMTC, according to each service requirement, such as speed, bandwidth, and latency. The technology aims to provide up to 20 times faster speed, 10 times more IOT device connections, and 10 times shorter low-latency services than 4G mobile communication technology[11]. We divided the technical features of the 5G network by component (i.e., user device, wireless access network, core network, externally interoperable applications) and summarized them in Table 1(reconstruction based on resources). As the main features, 5G techniques selected the software- based architecture, such as cloud-based virtualization technology, network slicing, multi-access edge computing (MEC) support, and service-based interface, by adopting the latest ICT technology to achieve the performance goal of 5G services and provide a

flexible and scalable mobile network, according to the business-to-business environment. 5G security threat modeling must precede the design of the 5G core network security technology, and in this study, we intend to classify the method so as to 5G network attacks for security threat modeling. While previous studies mainly analyzed the relationship between protected assets and security threats, this paper is meaningful in that it analyzed security issues related to inter-operated 5G technology characteristics in detail and classified attack types from the perspective of network protocols. Chapter II describes security threats, according to the evolution of 5G network technologies, by analyzing the previous studies of 5G security threats. Chapter III classifies the security vulnerability issues of mobile networks and the types of network-based cyber attacks that are likely to occur in a 5G network and describes new protocol security issues of the SA-based 5G core network. Lastly, Chapter IV concludes this paper with a summary of this study and future research direction.

2 NEW SECURITY THREATS OF THE 5G NETWORK

RELATED WORKS

Major countries, such as EU, USA, Korea, and China are highly interested in 5G security issues and engage in more fierce competition for the commercialization of 5G services. Therefore, research on 5G security architecture has been underway at the security working group (WG) of ITU-T SG17, 5G PPP (participation by the European Commission, manufacturers, carriers, service providers, and research institutes), as well as 3GPP, an international standardization organization. 5G Working Group of Next Generation Mobile Networks (NGMN), led by mobile communication carriers, handles network slicing and MEC security requirements. The European Telecommunications Standards Institute, network function virtualization security (NFV SEC) WG, mainly handles the security specifications of the NFV platform. With regard to the 5G security standards as a responsibility of 3GPP SA3, discussions on standards for security architecture, authentication, network slicing security, and subscriber information protection started in 2016. Security standards (TS 33.501) were announced in August 2018 at 5G Release 15. The Release 15 security standards have further strengthened security to address various security issues discovered in the previous generations. The following was introduced for the improved security function: the International Mobile Subscriber Identity (IMSI) information encryption function to protect subscriber information (IMSI user identifier stored in SIM card, etc.), Security Edge Protection

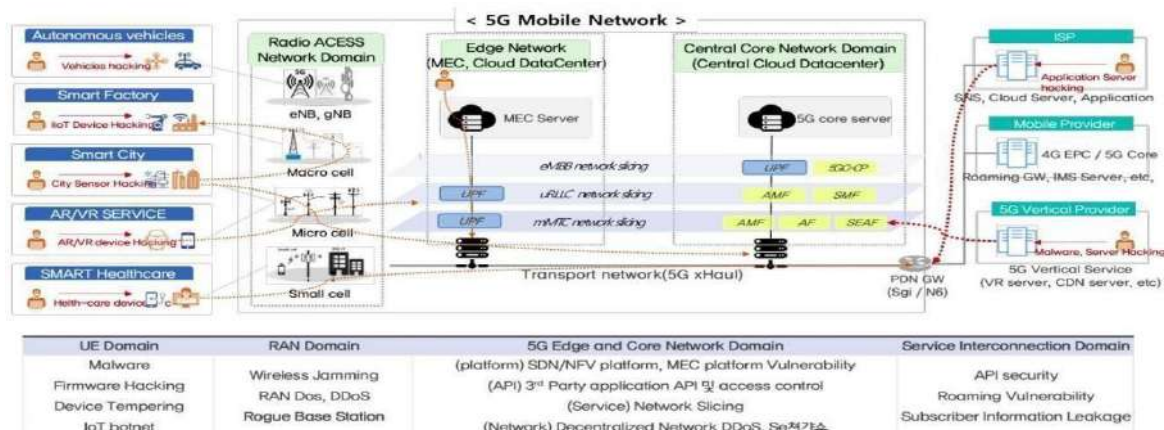
Proxy (SEPP) to solve the Signaling System No.7 (SS7) security issues between roaming domains and to implement the application layer security function between different carriers (i.e., public land mobile network) and the integrated authentication framework (i.e., security anchor function (SEAF)), which can use the same authentication method for 3GPP access and non-3GPP access. SEAF enables devices to be re-authenticated without executing a full authentication method (e.g., AKA authentication), even when they move between different access networks or between different service networks. The European Union Agency for Cyber security (ENISA) classified 5G network threat types into seven categories, in addition to conducting a long-term evolution (LTE) security threat analysis and then analyzed them as a threat landscape according to the CIA criteria. USA's 5G America analyzed the potential security threats by classifying them into UE/Device, radio access network (RAN) section, edge, core, S-Gi, and interoperation section, through 5G threat surface research.

Global IT companies are also actively conducting research on the development of security technologies, such as a 5G firewall. Positive Technology, which is a global firm that develops a signaling firewall for mobile networks, has been conducting research on security vulnerabilities and developing mobile network security technologies for various protocols, such as the GPRS tunneling protocol (GTP), signaling system no. 7 (SS7), and Diameter. Cisco[8] has analyzed 5G security threats that could occur when building a 5G network and conducted research on the development of response technology products.

SECURITY ISSUES OF 5G NETWORK

Summarizes the typical security threats for each section centered around the 5G mobile network. In general, over the mobile traffic path, user equipment is connected to application servers of Internet protocol (IP) service networks (Internet service providers, roaming interoperation between countries, etc.) through a radio access network (RAN) and a core network, providing mobile network functions for mobility management, authentication, billing, etc. Here, because the 5G network is connected to vertical industrial networks, such as automobiles, medicine, factories, and IoT devices, as well as the existing legacy networks (2G, 3G, 4G) and Internet service networks (e.g., SNS, cloud server, etc.), it will create a network connection structure consisting of complex heterogeneous networks, centered around the 5G networks. The complexity of these networks can lead to weak linkages arising from the interconnection of networks and devices with different

security requirements and different levels of security technology applied. Therefore, the biggest security threat is that it can downgrade 5G security. In this section, we describe these security issues by classifying them into five categories.



IOT DEVICESECURITY

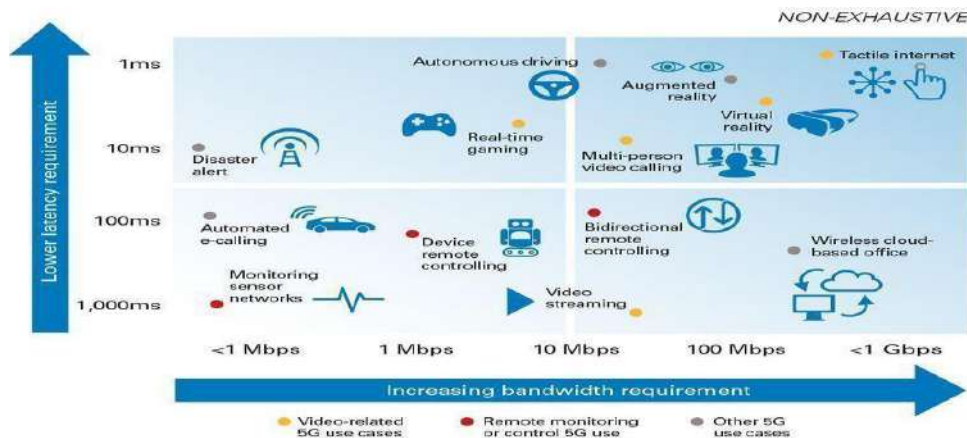


Figure 3: Advanced applications enabled by 5G

5G technology is expected to handle high-speed large-capacity traffic, 20 times faster than LTE, and the number of IOT devices connected to 5G networks is expected to grow by 10 times (1 million per unit area). The advantage of 5G is that it can establish a hyper-connected environment that provides MMTC services by allowing a large number of IOT devices to access 5G networks. However, if IOT devices that are vulnerable to security management are infected with malicious

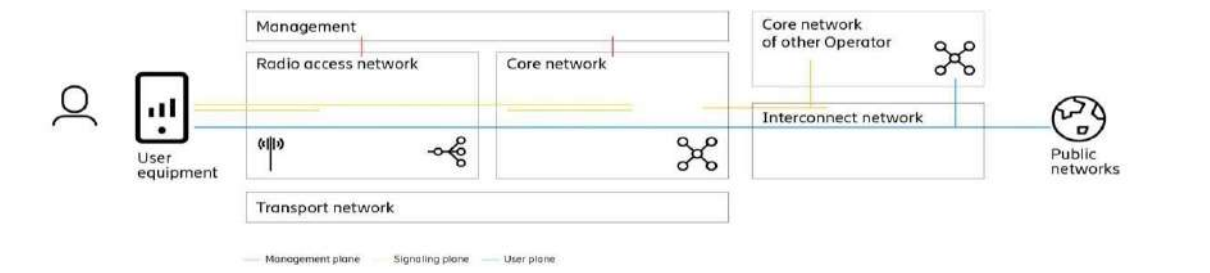
code, which triggers large-scale distributed denial-of-service (DDOS) traffic, they may have a direct impact on the 5G network. According to the European ENISA Threat Landscape Report 2018 [26], the size and intensity of DDOS attacks are growing at an alarming pace. In 2016, the first case of an IOT devices-related DDOS attack occurred, and in 2018, a terabyte DDOS attack (1.35 TBps) was targeted against the Git Hub server; the level of the DDOS attack has been gradually increasing up to 1.7 terabytes.

Unlike smart phones, designing common security standards and architecture for IOT devices is not easy because industry-specific device types (e.g., smart factory devices, smart city sensors, and CCTV), applications, and supply chain ecosystems are diverse and complex. In addition, as the installation of a high-level security function for low-end IoT devices is difficult, they are vulnerable to tampering because they have weak passwords and old security protection. Therefore, they are more likely to be exposed to vulnerable environments, such as improper access by malicious applications and leakage of subscriber information (IMSI) by man-in-the-middle attacks. Hackers may have access to vulnerable IOT devices and infect them with malicious code to construct a large quantity of IOT bot nets and then remotely control them through a C&C server to use IOT devices as a means of attack.

.5G RANSECURITY

The RAN section is composed of various types of base station equipment (macro cell, microcell, femto cell, etc.). The RAN base station equipment is connected to UE through a wireless communication interface (i.e., Air Interface) and acts as relay equipment, which connects to the 5G core equipment through a wired transmission network. 5G RAN technology has the advantage of allowing various types of wireless access technologies to gain access to the 5G network by accommodating not only 3GPP wireless access technologies (i.e., 2G, 3G, 4G) but also non-3GPP access technologies, such as Wi-Fi and wired Internet. However, as various heterogeneous wireless access and mass IOT device access are allowed, protection of the RAN section is critical. To connect mobile communication services, control signals (movement, authentication, billing, etc.) are exchanged between the UE and the base station (eNB, gNB) equipment of the RAN section and the communication equipment (MME) of the core network. Here, when abnormal control traffic, due to millions of user devices connected to the RAN base stations, are transmitted and received,

resilience issues for failures and strengthening the security of small cells for home and business, which are easily accessible, are crucial factors.



The RAN security threat includes a RAN DDOS attack that requests excessive access to wireless resources by a huge number of IOT bot nets, infected with malicious code, and a jamming attack on wireless signal channels. The base stations transmit and receive abnormal data, owing to RAN DDoS and radio interference jamming attacks, thus resulting in the exhausting of radio interface resources in the RAN section, which ultimately leads to availability issues, preventing normal data reception. Rogue base station issues allow attackers to launch various types of attacks, such as the interception of user location information, tampering of transmission information, and DDOS attacks between the mobile user and the network through a man-in-the-middle attack, between the mobile UE and the 5G network, using a false base station. Rogue base station issues were continuously raised for 2G, 3G, and 4G legacy networks and various improvements were applied to the 5G security standards. However, if the distribution of small scale femtocells is accelerated to resolve the shadow area of wireless mobile communications, rogue base station issues can still be raised in cases where attempts by hackers targeting small cells, for which security management is relatively neglected, compared to macro base stations, have increased, such as the security threats of low-end wireless local area networks with the wide spread use of wireless APs.

DECENTRALIZED 5D CORE ARCHITECTURE SECURITY

The 5G core network adopts a decentralized core network structure. Until the 4G network, a centralized network structure was adopted, in which traffic signals and data transmission paths were centralized in a central office, up to the 4G core network. However, for the local redeployment of the communication function, the 5G network was decentralized by dividing it into a core network and an edge network. Because the path of control plane and user plane traffic is physically separated, the central office mainly processes control traffic, and user data is processed by cloud-based edge communication centers at local offices. This decentralized core network structure is an efficient structure for providing ultra-low latency services and network slicing services. However, because the protection targets are widely and locally distributed, there are security visibility issues which are tightly interconnected between legacy network (2G, 3G, 4G networks) devices. Therefore, a decentralized response to cyber attacks is inevitable, resulting in a decentralized response to them from a comprehensive perspective.

SOFTWARE-BASED INFRASTRUCTURE SECURITY

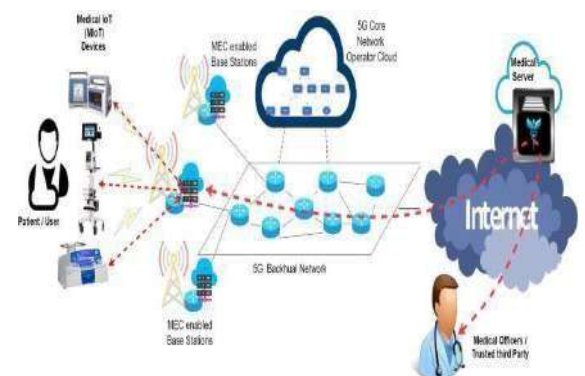
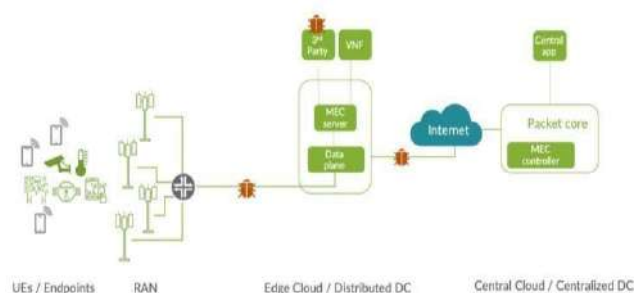
The 5G network has shifted from hardware-dependent infrastructure to software-based infrastructure [4]. The software-based infrastructure is implemented using 5G communication servers, network equipment, and network slicing services, through SDN and NFV virtualization technology. It provides an independent network slicing service for each URLLC, MMTC, and EMBB application service by separating a single physical network into multiple virtual networks. Here, instead of dedicated equipment, the network communication function is implemented on a general-purpose x86 server, in the form of a virtualized SW (i.e., Virtual Machine). However, although virtualization technology, which is the core of 5G equipment and service implementation, has advantages in terms of resource efficiency, flexibility, and availability, by sharing physical networks and HW server resources (e.g., CPU, memory), it can be relatively vulnerable to load attacks on physically shared HW resources, unauthorized access to network slicing and shared resources, malicious code distribution through shared resources, and configuration errors for virtualization management SW[8][24][10][16]. SDN/NFV security: SDN technology controls the network delivery function, which has been processed in hardware, by separating the network control function (SDN controller) and traffic delivery function (SDN switch). Traffic bypass attacks that exploit control protocol vulnerabilities between SDN controllers and switches, unauthorized access

between switches and controllers, and resource depletion of SDN systems by DOS attacks can paralyze services. For example, a saturation attack can occur, which exhausts the SDN switch flow table by attacking SDN controllers. In addition, NFV technology implemented on a general-purpose server has a high possibility of security issues, unless hypervisor security, malicious VM migration issues, changes or authentication of applications running on virtualized network functionality, and authorization for networking functions are properly controlled. If there is no protection mechanism in place for authentication and authorization of applications, malicious third party applications can obtain network information from SDN controllers.

(1) Network slicing security: Network slicing is a new technology introduced to the 5G network. It is a virtual network transmission technology that logically separates traffic for each EMBB, URLLC, and MMTC service, while using the same physical network. Here, if the network slicing for each service is not properly separated, there is the possibility of attack from one slice to another. For example, an attacker may launch a network slicing resource depletion attack by maliciously overstretching traffic capacity in a network slice dedicated to a specific service, and subsequently, affect other network slices or simultaneously activate specific applications. Without proper encryption applied to the network slice, an attacker could eavesdrop or tamper with data belonging to other slices.

MECSECURITY

MEC / Edge Cloud Attack Surfaces



MEC refers to the concept of providing services by constructing an application server inside a mobile network, close to the user's device, using a method that goes through the existing mobile communication core internal network and connects to the application server of the Internet service.

The combination of the 5G network and the concept of edge computing in MEC has the advantage of providing IOT applications and services, such as telemedicine, autonomous vehicles, factory automation, and IOT sensor data information processing, in real-time without delay. However, because the edge computing server is forward deployed inside the 5G edge network (connected with UPF equipment), a new connection path can be created, which leads to security issues. In a mobile network, MEC is implemented through cloud and virtualization technology and is expected to operate in an open system running third party applications. Therefore, MEC systems built into the internal network of mobile networks can be the main targets of hackers[10]. For example, MEC can be built as a virtualization platform and MEC applications can run on the same platform as some virtual network functions. If a MEC application is a third party application that is difficult for mobile carriers 7 5G core network security issues and attack classification Hwankuk K. to control, it can consume virtualized network resources or obtain access to unauthorized sensitive information with inappropriate application programming interface (API) permissions. Furthermore, there is a risk of providing a new attack path for an attacker, who can attempt to launch an attack on edge network functions, such as UPF, which is distributed 5G network internal equipment, by inserting malicious applications.

3 CLASSIFICATION OF 5G CORE NETWORK PROTOCOLATTACKS

SECURITY VULNERABILITY ISSUES OF THE 5GNETWORK

To protect 5G networks and services, security technologies that are different from those of the previous generations must be designed, developed, and operated. To establish and operate a safe 5G network, Ericsson[6] derived the security requirements in the stages of standardization, equipment development, network construction, and service operation. Security requirements and security vulnerability issues for each stage have been summarized in Table x. Because the mobile communication network is composed of a complex ecosystem, it is relatively difficult to solve the associated security vulnerability issues.

Table 2: Security requirements and vulnerability issues for each stage to service from standardization

Stage	Security requirements	Vulnerability issues
Standardization	Design of a secure communication protocol for network inter-operation	Protocol Vulnerability (Definition of basic security requirements & specifications)
Development	Development of equipment that meets security levels required by standards	Equipment implementation vulnerability (Different implementations of common function, SW errors, etc.)
Deployment	Design and construction of secure networks & services	Network construction vulnerability (Configuration error, Open API, 3rd SW)
Operation	Detection and monitoring for cyber attacks, incident response management	Operational vulnerability (Vulnerability response, supply chain security)

First, in the standardization stage, communication protocols and interfaces must be designed safely for the interoperation of networks and systems between countries. Research on the standards of 5G basic security requirements and architecture has been active by international standards bodies and de-facto standard organizations. 3GPP security standards for authentication and key management for mutual authentication between users and networks, signaling messages of the control plane, and data protection of the user plane have been developed to continuously enhance the security of the mobile network. However, because standardization defines the minimum basic security requirements and specifications, there are concerns that vulnerabilities in standard protocols can occur at all times.

Second, manufacturers should develop equipment that meets the security standards and target levels required by standards. For example, security vulnerability issues continuously occur in the stage of equipment implementation because each equipment manufacturer implements different security functions; equipment implemented with SW may contain SW errors or unknown vulnerabilities at the time of equipment implementation can be found over time. However, when the equipment construction is completed, it may take a long time to patch and verify the SW, which leads to supply-chain security issues.

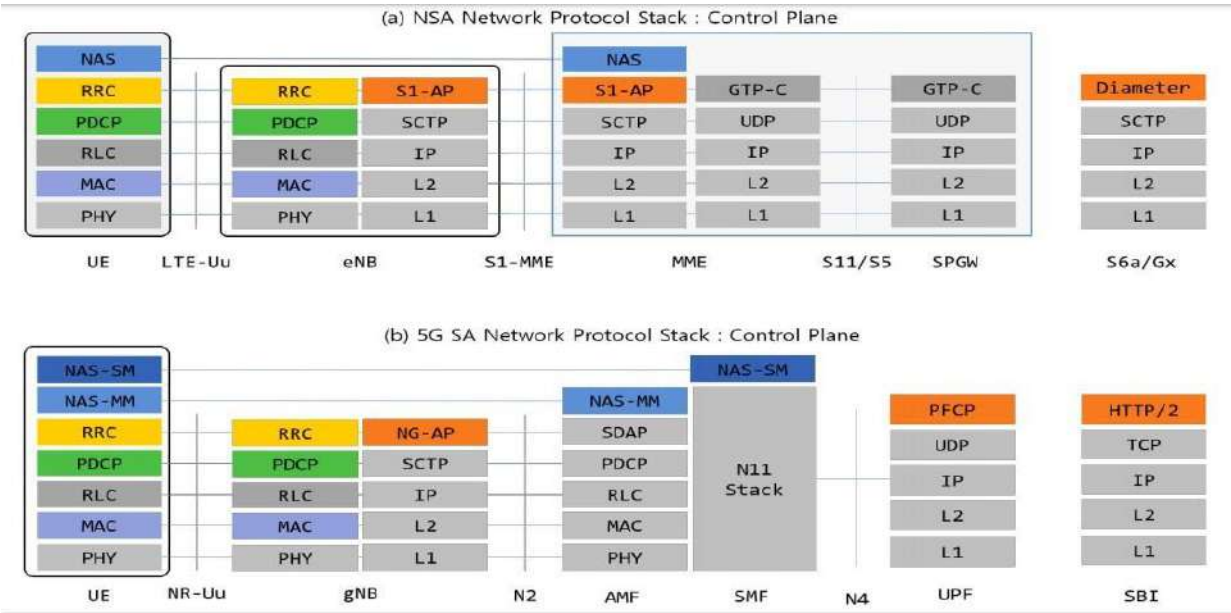
Third, communication carriers must design and build secure networks and services by verifying supply-chain products, to ensure that equipment manufacturers' communication equipment and service applications are implemented to meet security requirements. Nonetheless, configuration setting errors can be present in the process of building networks and services and security issues by third party applications, not by communication carriers, have been continuously raised.

Lastly, in the final stage of service operation, it is crucial to remove security vulnerabilities against highly advanced and intelligent cyber attacks and to restore resilience after cyber breaches. Moreover, the response time required to resolve security issues at each stage can also be a barrier.

In other words, because it takes several years to reflect security vulnerabilities of standard protocols in the standards and it takes approximately 6 months or more for SW patches to provide safety verification to address equipment implementation vulnerabilities, it is critical to close the security gap between each stage.

CLASSIFICATION OF NETWORK ATTACKS ON 5G NSANETWORK

A number of complex standard protocols for mobile networks are used for signal processing, such as user call and data routing, paging, mutual authentication, roaming, and billing. Research on security vulnerabilities in mobile network protocols and cyber-attacks that exploit these vulnerabilities has been continuously conducted since the introduction of the 2G mobile network.



As described in Section 3.1, because the protocols defined in the standards define only basic security functions, security is not sufficiently considered when designing these protocols. Therefore, although various protocols are interoperated and interacted in mobile communication carrier networks, there are many security vulnerability issues because they are applied differently, according to unique settings and technologies for each communication equipment manufacturer or service area. Because future 5G standards are to be modified by the use of new protocols in the 5G core network, as shown in Figure 3, the greater complexity of interoperation between heterogeneous networks and protocols may require a significantly more complex design and implementation of security technology. Therefore, in this section, we first investigate the attacks on the NSA 5G core

network protocols, based on the 4GC EPC core, and then describe the issues for new 5G SA protocols.

Table 3 shows the classification of attack types related to network protocols on 5G core network.

5G security threats classified by ENISA are SPAM, identifier spoofing, location tracking, DOS, subscriber fraud, message intercept, call routing attacks, and infiltration attacks. The most common 5G threats are the following:

- Eavesdropping/interception/hijacking: this threat is intercepting into mobile traffic gaining valuable and confidential subscriber information
- Fraud: Attackers can use services at the expense of the operator or another subscriber using invalid or hijacked IMSI
- Injection of malicious messages: Disrupting sessions and creating DDOS
- Subscriber denial of service: this threat is spoofing subscriber IDs to generate malicious messages that cause service disruption for an individual subscriber
- Message Modification: this prevent message delivery or allow malicious content delivery, disrupting service
- Network DDOS: Malicious, malformed or invalid signalling packets are sent that overwhelm network elements or cause vulnerable elements to fail !

Table 3: Classification of 5G Threat based network protocol

Network Domain 5G NSA Components 5G SA Components		UE Devices	Access Network		SGW AMF	MME AMF(SMF)	PCRF PCF	HSS AUSF	IMS AF	PGW UPF	External Network ISP, Roaming Vertical Service	Ref.
Control Plane	RRC	DoS	O									[12]
		Spoofing	O	O								
		Location Tracking	O	O								
		Routing Attack	O	O								
	NAS	DoS	O	O		O						
		Spoofing	O			O						
		SPAM	O			O						
	GTP-C	DoS			O	O	O			O		[19]
		Fraud			O	O	O			O		[18]
		Routing Attack		ü	O	O	O			O		
	Diameter	DoS					O	O		O	O	[4]
		Routing Attack					O	O		O	O	[20]
		Information Disclosure					O	O		O	O	etc
	SS7	Fraud					O	O		O	O	
		DoS						O		O	O	[4]
		Faud						O		O	O	[20]
User Plane	GTP-U	Location Tracking					O	O		O	O	etc.
		DoS(GTP-in-GTP)			O	O	O	O				[19]
		Fraud			O	O	O	O				[18]
	Voice over 5G	Sniffing			O	O	O	O				etc.
		SIP Signaling DoS	O		O				O		O	[13]
		SIP Replay Attack	O		O				O		O	
		Location Tracking	O		O				O		O	
	IoT over 5G	IoT Aplication DDoS	O	O	O				O		O	[8][11]
		IP based attack	O	O	O				O		O	

(1) RRC protocol-based attack: The RRC protocol is a control protocol that manages a radio bearer of the L3 layer, related to radio resource establishment, reconfiguration, and release between the UE and radio access network. A study by KAIST announced that, through the RRC protocol attack, attackers are able to launch various attacks, such as subscriber ID tampering, DOS attacks against base stations, and authentication bypass due to vulnerabilities of the baseband chipset of the UE and base station equipment implementation.

(2) NAS protocol-based attack: The NAS protocol consists of control protocol messages that manage UE authentication, mobility, and the location between the UE and mobile communication core network equipment (MME, AMF/SMF). A study by KAIST announced that, through the NAS protocol attack, attackers are able to launch attacks such as DOS against MME equipment, subscriber identification information leakage, and man-in-the-middle attacks due to vulnerabilities in standard protocol specification, authentication bypass, and error handling processing errors for communication messages, when implementing the SW of mobile communication equipment.

(3) GTP protocol-based attack: GTP messages are a control protocol related to creating and releasing tunneling inside the core network for IP data transmission. It operates on the user datagram protocol (UDP) and there are GTP-C messages for tunneling session establishment and control in the core network section, GTP-U messages related to data transmission, and messages related to billing. With regard to the section where the GTP is used, control information (GTP-C) messages for GTP tunnel creation, maintenance, and deletion are exchanged between the MME and S-GW sections and the S-GW and P-GW sections. User packets are transmitted through GTP-U tunneling between the base station and base station sections, the base station and S-GW sections, and the S-GW and P-GW sections. GTP prime messages are used to transmit billing information (CDR) in the P-GW and OFCS sections. Because the GTP protocol was introduced with the aim of using it only within the mobile network section, without taking into account the security of encryption and authentication from the initial stage of designing the standards, research results on its vulnerability to hacking attacks have been continuously published. According to studies conducted by Positive Technology[19], GSMA[1], and KISA[18], it is possible to conduct man-in-the-middle attacks and DoS attacks against EPC or 5GC core equipment through the forgery of the field values of GTP messages, malformed GTP messages, and spoofing.

(4) Diameter protocol-based attack: The diameter protocol is an IP-based protocol for Authentication Authorization and Accounting and is used to control quality of service policies, such as MME and P-GW equipment, policy server (policy and charging rules function (PCRF)), and subscriber information management (home subscriber server (HSS)). In the 5G NSA core network, it is used in the MME and HSS sections for user authentication, the P-GW and PCRF sections, and the PGW and online charging system (OCS) sections for billing. Attacks using the diameter protocol include connection hijacking and replay attacks.

(5) SS7 protocol-based attack: The SS7 protocol is primarily used for 2G and 3G; however, this protocol attack is still a threat because, currently, roaming between countries is connected to legacy communication networks. Possible attacks to address the SS7 issues in the 5G standards include SPAM, spoofing, location tracking, subscriber fraud, intercept, DOS, and routing attacks.

5G SA NETWORK PROTOCOL-BASED ATTACK ISSUES

The SA 5G core network pursues a service-based infrastructure. Until 4G core network, internal communication functions, service management, and applications were developed under the control of carriers, and interoperation between equipment was possible through the P2P interface. However, for 5G, the interoperation between equipment was unified with HTTP-based web interfaces, and the open API facilitated internal communication service functions and data access for service providers of vertical industry, such as IOT and factory automation. The mobile network section is relatively closed, compared with the IP network, which has served as a high barrier against hackers. However, because the Internet web technology adopted in the 5G SBI architecture is well known to attackers and web application services still have many security vulnerabilities, it can be exploited as attackers' preferred attack methods. In addition, open API security can be a problem by providing API functions, such as SCEF and NEF, to the outside. It is expected that vulnerability management of well-known existing web applications and access control to open APIs will be essential.

The protocols used for 5G signaling and data transmission will also change significantly. Among the network protocols on the control plane, typically, the SS7 and diameter protocols are expected to use HTTP/2, JSON, and REST API. In addition, the GTP-C protocol is expected to be

modified to an HTTP-based interface. However, the GTP protocol is likely to be continuously used to interconnect data on the control plane and user plane paths.

In 5G security standards, security functions, such as mutual authentication and encryption of signaling protocols on the control plane, have been improved to address security threats occurring in 2G, 3G, and 4G networks. However, the response standard for protocol-based attacks on the user plane is relatively inadequate. Communication carriers operating 5G networks are concerned whether they can detect the connection of IP traffic transmitted from the user plane path, various types of IOT DDOS traffic passing through 5G networks, DDOS attacks through virtualized network slicing, and abnormal traffic in numerous edge networks.

4 .CONCLUSIONS

The 5G network introduced technological advantages by adopting a software-defined infrastructure to accommodate the connection of IOT devices. While 5G Security is advanced step forward, the risks inherent interconnection prior network continue to grow against a much larger volume of traffic and applications. IOT traffic, with its high complexity and large number of interconnect partners and hubs, can be an especially vulnerable and attractive target for attackers.

This paper summarizes five security issues arising from 5G technical advantages: 1) security issues such as the response to DDOS attacks caused by security vulnerability of IOT devices, 2) RAN failure and small cell security management owing to heterogeneous wireless network access and coverage expansion, 3) visibility for security monitoring and expansion of the protection target owing to the decentralized mobile network structure, 4) dynamic security management and access control caused by the sharing of physical HW equipment regarding virtualization platform and network slicing technology, and 5) security issues related to third-party applications and API reliability and connection paths to internal mobile communication networks, which are caused by applying MEC.

REFERENCES

- [1] Securing the 5g era. <https://www.gsma.com/security/securing-the-5g-era/> [Online; accessed on February 3, 2020],2020.
- [2] I. Ahmad, T. Kumar, M. Liyanage, J. Okwuibe, M. Ylianttila, and A. Gurtov. Overview of 5g security challenges and solutions. IEEE Communications Standards Magazine, 2(1):36–43, March 2018.

- [3] P. Bisson and J. Waryet. 5G PPP Phase1 Security Landscape. Technical report, 5G PPP, June 2017.
- [4] ENISA. Signalling Security in Telecom SS7/Diameter/5G. Technical report, ENISA, March 2018.
- [5] Ericsson. 5g security - scenarios and solutions. Technical report, Ericsson, 2017.
- [6] Errison. A guide to 5G network security. Technical report, Errison, December 2018.
- [7] M. A. Ferrag, L. Maglaras, A. Argyriou, D. Kosmanos, and H. Janicke. Security for 4g and 5g cellular networks: A survey of existing authentication and privacy-preserving schemes. Journal of Network and Computer Applications, 101:55–82, January 2018.

SMART HOME AUTOMATION SYSTEM USING IOT

Dr.Priya¹,

B.Vinisha², C.Kowdelya³, S.Banu Priya⁴

Dept of Computer Application

Marudhar Kesari Jain College for Women, TamilNadu, India.

ABSTRACT

Due to vigorous development in the field of the automation sector. Human survival is becoming more and fast advanced and being better in all wishes and aspects. In olden days, with this furious growth in the number of gaining consumers with the use of Internet, this has been playing a vital role in our day –today life. And the IOT is the novus and emerging network technology feature in our present life. Internet of things feeds an major vital role in our present life as well as in the educational field because they have the feature and ability to provide information and they can complete the given tasks while the people are engaged with other work. This is the best methodology for an energy management system. The process to the whole method is given by its admin only to various users. This internet of things is even expandable for organizing. Controlling various appliances used at home and also for our privacy, security and safe measure purpose of the home via sensors as long as. It exists on wifi internet coverage

Magical word: Internet of things; Home automation; smart homes; sensor

INTRODUCTION:

The internet of things (IOT) is a complex that allows devices to be attached and remotely accessed across the internet. Technology becomes more necessary in human life people need communication and information is one exposure that control by current technology maturing, which admit us to make a device and system that can aid all activities and set on problems by themselves. One of the arrangements man to ease the work of the arrangement is the smart home. It is an electronic system for an extent or a lodge, appartment ease task management. In 2015, controlling motor driven apparatus of homey devices such as electric potential different current and after words calculated the contraptions power consumed is accomplished by suryadevara al al. The management of energy in a home can be improved by this methodology, which makes conventional utilization of electricity. The fact finding conducted by Mbarek al al put forward a safe data collection scheme using compressed sensing(15). The authenticatic process is then

performed for each sensor mode and the necessity of the hash algorithm to certify data integrity is the smart home system.

LITERATURE REVIEW

- **BLUETOOTH BASED HOME AUTOMATION SYSTEM USING IMMOBILE**

In Bluetooth home automation system the home appliances are connected to the Arduino BT board at input output ports using relay. One circuit is designed and implemented for receiving the feedback from the phone, which indicate the status of the device.

The password protection is provided so only authorized user is allowed to access the appliances. The Bluetooth connection is established between Arduino BT board and phone for wireless communication. The programs of Arduino BT board is based on high level interactive C language of micro controllers; the connection is made via Bluetooth.

- **GSM BASED HOME AUTOMATION SYSTEM USING IMMOBILE**

GMS is the SMS based home automation. The Mobile Phone and GSM technology, the GSM based home automation is lure to research. The sensors of system convert the physical qualities like sound, temperature and humidity into some other quantity like voltage. GPRS base home automation and dual for multi frequency (DTMF) based home automation, these options are considered mainly for communication in GSM. It shows how the home sensor and devices interact with the home network and communication through GSM and SIM (subscriber identity module).

The microcontroller analysis all signal and convert them into command to understand by GSM module.

- **WI-FI BASED HOME AUTOMATION SYSTEM USING CELLPHONES**

WI-FI based home automation system mainly consist three modules, the server, the hardware interface module, and the software package. The server is connected to the internet so remote users can access server web based application through the internet using compatible web browser. Arduino software is culpable for gathering events from connected sensors, then applies action to actuators and pre-programed in the server DB. WI-FI technology is used by server, and hardware interface module to communicate with each other. Server use database to keep log of home automation system components, we choose to use XML files to save system log.

- **HOME AUTOMATION USING ANDROIDADK**

The microcontroller board (Arduino ADK) is based on the ATmega 2560. It has a USB host connection to associate with android based phone and that is based on the MAX3421e IC. It has audio output that is from the android device to the component and it also support for the component server as one or more human interface devices(HID) to the android device. The device of home are associated to the ADK and the connection is established between the android device and ADK. This paper depends upon android and arduino platform in which both are FOSS (free open source software). Including motion sensor for safety system will detect an unauthorized and it will automatically notice the user through cell phone or the security system.

METHODOLOGY

Remote home automation using IOT (INTERNET OF THINGS). There are various methodology to control home machines, for example, IOT based home computerization under wi-fi through android application from any smartphone. Arduino based home robotization home computerization by android. In this framework, the basically situated keen automation framework which assess the development of a low- cost security frame model using motion sensor and IOT the man power growth is defined utilizing the motion sensor. This very responsive approach has low arithmetical necessity. All of the items with a temperature above supreme zero discharge warm vitality as radiation. Typically this radiation is not well defined by our maked eye since it infrared wave lengths, yet in this infrared can be identified by electronic contrivance intended for

identifying the growth development of human. The distance sensor has a scope of around 20feet (6 meters).

The sensor is intended to explain the gently changing terms and conditions that would cause the basic change as the everyday up comes and the ecological condition varying, yet it reacts by rolling out its flex when unexpected improvements occur, for example when there is a movement. This device is composed mainly for indoor use. Operation outside or in high temperatures may cause security unfavorable. Because of the high tenderness of distance sensor device, it is not prescribed to use the some condition like quick ecological changes and solid or vibration and futhermere in not accessing in coordinate daylight or dright flow from a warmer or moderate and household \ office security frameworks have improved is prominence as of late a home\office holder's search for approaches to ensure their own space and upgrade their home estimations.

CONCLUSION

In this exploration security framework motion sensor has been utilized which is less power and low-minimal effort. Ithasan immense focal point go ,and is anything but laborious to communication with arduino. This security framework can be implemented in places like home, office, shop and so on . The pliancy run for distinguishing movement of this framework can be provide with grass break finders to upgrade the level of security. Utilization of multi- sensor information combination and complex calculation can be utilized to build the successful FOV (field of view) for more spaces with a unique end goal to develop the ava exactness and to improve the schema for preparing the MOTION sensor flag, using more propelled procedures, for example probabilistic hypothesis.

FUTURE SCOPE

In this project , we suggest a easy solution for home automation based on ESP8266 chips and raspberry pi boards. These two choices are easy to work, small, cost effective . The contribution of the paper require the survey of the last 10 years that published in the open source home automation system. As compared to other papers like literature, this paper requires the details of the implementation of the solution for both hardware and software.

Most of the smart home appliance systems presented in the literature have been made with fewer functionalities, using different technologies , controllers , type of communication user interface ,etc.Q key words with a selected list of devices imposing a unitary APT, firmware,

etc. We supply the open source firmware. The definition of open source firmware is that no hacks and no 3rd party centre or clouds are required. All the devices speak the same language (API) these are controlled in same way. Usually, a basic package for automating a small house exceeds 1000 dollars. Q key tries to do a very good job at providing a curated list of supported devices. Q key provides a PWA(progressive web app). The future will lead in increase of period of sensor products, as well as devices. It leads every aspect of our home life is automation. The characteristic that will be shortly added to q key is monitoring the air humidity. Extreme humidity levels can cause die.

Future task will be the combination of video surveillance in q key. The user can manage their video cameras very easily. Motion eye has become very popular in the open source world, with 50 releases on github, and more than 650,000 downloads since 2014.

REFERENCES

- Anandhavalli, D; Mubina, NS; Bharath, P. Smart Home Automation system and GSM . Int.J.Inf Future. Res.2015,2,2547-2552
- Li,ZM; song,M.;Geo, L.Design of smart Home System based on zigbee. Appl. Mech.Mater.2014, 635-637,1086-1089.
- Badabaji, S; Nagaraju,Vs An IOTbased System .Int .J. Pure Appl.Math.2018, 119,4659-4667.
- Fernandez- Caremes, T.M. An Intelligent power outlet System for the Smart Home of the Internet of Things. Int.J.Distrib.sens. Netw.2015,11,214805[Cross Ref][PubMed]
- Shahajan, M ; Islam, G.M.J;Das, S.K. Islam,s. ; Islam,M;Islam,

ABBREVIATIONS

The following abbreviations are used in this manuscript.

IOT= Internet Of Things

CS= Compresed sensing

FOV= Field of view

API= Application Programming Interfac

PWA= Progressive Web App.

A STUDY ON IMPACT OF COMPENSATION AND WORKING CONDITION TOWARDS JOB SATISFACTION OF COLLEGE TEACHERS

Ms.J.JEEVITHA¹, Research Scholar,

PG and Research Department of Commerce,MKJC.

Dr. C. NITHYA², Research Guide & Supervisor,

Head, PG &Research Department of Commerce, MKJC, Vaniyambadi ,TN,India.

Abstract:

This study examine the relationship between the compensations and workplace condition on Job Satisfaction of college teachers. The sample consisted of 200 full time working faculty members from self-financing college in affiliated Thiruvalluvar University of Tirupattur District. The main factor under consideration are Compensation and workplace condition of organization affect the job satisfaction. The eight important factors are measure the impact of compensation towards job satisfaction. The ten important factors are measure the impact of workplace condition towards job satisfaction. The correlation analysis result indicated that moderate positive & significant relationship between compensation and workplace condition ($r = 0.512$); high positive & significant relationship between compensation and job satisfaction ($r = 0.856$); high positive & significant relationship between workplace condition and job satisfaction ($r = 0.882$).

Key words: Compensation, Working Condition, Women Faculty, Job Satisfaction.

1. INTRODUCTION:

Many Institution fail to understand the importance of Compensation and working environment for employee job satisfaction and thus face a lot of difficulties during their work. There are two important policies that have been done to improve the human resource performances such as; compensation policy and creating the work environment support to create the job satisfaction.

Definition: Keith Davis and Newstrom, “Job satisfaction is the set of favourable or unfavorable feelings with which employees view their work”. **Andrew Brin**, “job satisfaction is the amount of pleasure or contentment associated with a job. If you like your intensely you will experience high job satisfaction. If you dislike your job intensely, you will experience job dissatisfaction. **Arnold**, “job satisfaction will be defined as the amount of overall positive affect or feelings that individuals have towards their jobs.

Compensation: Rivai (2011) job itself, payment, promotion, supervision, and co-workers are the main factors influencing of job satisfaction. **Sedarmayanti (2014: 241)** there are six

variable indicators to measure the compensation such as; salary, bonus, incentive, insurance, security, and working holidays.

Work Environment: Lane & Esser et al. (2010) dissimilar factors within the working environment such as wages, working hours, autonomy given to employees, organizational structure and communication between employees & management may affect job satisfaction. Sedarmayanti (2011) lighting, air temperature, noise level, color usage, work ability, and work relationship main factors influencing of job satisfaction.

Job Satisfaction: The favourable or unfavorable feeling of employee about specific job factors. Smith Kendal and Hulin (1969) main factors effecting job satisfaction like Work, Pay, Promotions, Supervision, and Co-workers. Unnamalai .T (2015) made a study on Factors Influencing Job Satisfaction. The result of the study have shown that salary, working environment, promotional opportunity and interpersonal relationship have a main factors influencing of job satisfaction.

2. REVIEW OF LITERATURE:

Ilias Kroupis et al. (2017) working condition, pay, promotion, work itself, supervision, are the main factors influencing of job satisfaction. Senthilkumar, V. & Kannappa, R. (2016) has conducted a study on Employees Job Satisfaction in Collegiate Education with Special Reference to Arts and Science Colleges at Trichy in Tamil Nadu. The result of the study found that, salary, working environment, promotional opportunity and interpersonal relationship are the main factors influencing job satisfaction. Abdul and Raheela (2015) conducted a study on impact of working environment on job satisfaction. The simple random sampling method has been utilized to select a sample of 210 out of 70 selected employees of educational institutes, banking sector, and telecommunication industry operating in the city of Quetta, Pakistan. Self-administered survey method using the data collection. The main purpose of the study was impact of working environment on employee job satisfaction. The study revealed that there is positive & significant relationship between working environment and employee job satisfaction. Arul Edwin Fredrick. P (2015) in their study entitled various factors influencing job satisfaction of MBA Teachers in college with special reference to Madurai District, India". The result of the study shows that job satisfaction depends on promotion, pay, fairness, support and working conditions are the main factors influencing teacher's job satisfaction. Huma Bilal (2012) studied on "Job satisfaction of University Teachers: Impact of Working Conditions and Compensation". The sample consisted of 100 teachers from different public and private universities of Islamabad and

Rawalpindi. The main factor under consideration are working conditions of organization, rewards & benefits which may affect the job satisfaction. The result of the study shows that there is a positive relationship between working condition, compensation practices and job satisfaction. (WC& JS $r = 0.46$, $p < 0.01$, (C & JS $r = 0.35$, $p < 0.01$).

Statement of problem: Job satisfaction is not only for the employees but, the society as a whole. College Teachers are the most important group of professionals for our nation's future. As compared to other levels of educational system in the society, higher education has a much bigger role to play. As Therefore, it is astonishing to know that even today many of the college teachers are dissatisfied with their jobs. To know the job satisfaction level of teachers is very important since it increases productivity and classroom performance in the college. Self -finance colleges are playing very important role in providing higher education in our nation. The number courses and staff employed in self -finance colleges are remarkable. The salary structure and other facilities provided to self- finance staff are different from government college teachers. So it makes us necessary to study the job satisfaction of self -finance college teachers. In Tirupattur District, there are good number of self- finance colleges where women faculty members are working more. So impact of compensation and working condition of job satisfaction towards of women faculty members in self- financing college at Tirupattur district has carried out.

OBJECTIVE OF THE STUDY

1. To study the demographic profile of women faculty members in various self -financing colleges.
2. To determine the impact of Compensation and Workplace Condition on Job Satisfaction
3. To analysis relationship between Compensation and Workplace Condition on Job Satisfaction.

LIMITATION OF THE STUDY

1. The Study was conducted only faculty members in working self- financing Arts and Science Colleges.
2. The Study Covers only Tirupattur District of Tamil Nadu.

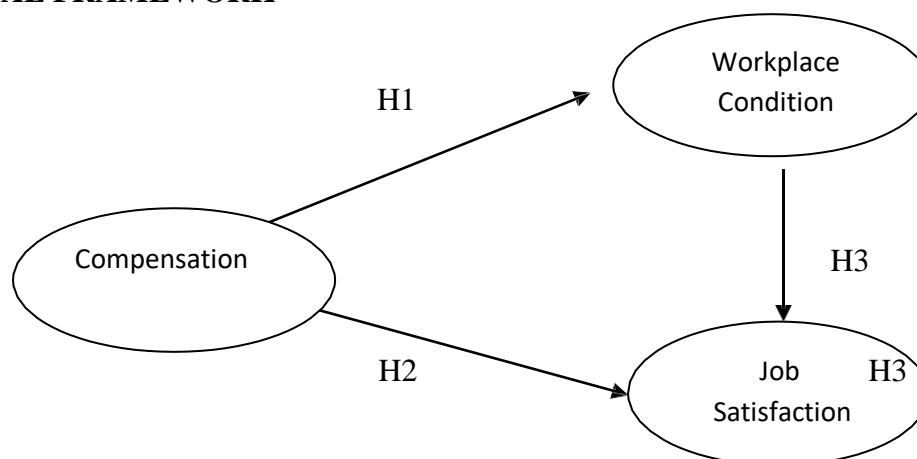
HYPOTHESIS

H1: The Compensation has a significant impact on Workplace Condition.

H2: The Compensation has a significant impact towards the Teachers Job Satisfaction.

H3: The Workplace Condition has a significant impact towards the Teachers Job satisfaction.

CONCEPTUAL FRAMEWORK



3. METHODOLOGY

An Exploratory and Empirical method has been used for the present study. The purpose of the study is the relationship between Compensation and Workplace Condition on Job Satisfaction. The population of the study was faculty members of private self-financing colleges of Tirupattur District of Tamil Nadu. The sample was consisted of faculty members of Assistant Professor only. The data is gathered randomly from the 200 faculty members in self-financing college in Tirupattur District. A 7- point Likert Scale is used to present study. Primary data is collected from the respondents using a pre-structured questionnaire through online and directly meeting the respondents from faculty of various colleges in Tirupattur District. The secondary data were collected from the journals, magazines, books and websites. Percentage analysis is used to find out the demographic data. Descriptive statistics like mean, standard deviation computed. Correlation is used to find out the strength of relationship between variables taken from the study.

4. ANALYSIS AND INTERPRETATION:

DEMOGRAPHIC PROFILE OF THE RESPONDENTS

In the Demographic profile of the respondents, the basic profiles of the faculty respondents like age, marital status, educational qualification and monthly salary were determined with the help of Percentage analysis

Table No: 1**Demographic Profile of the Female Faculty Respondents**

	Particulars	Frequency	Percent
Age	below 30 years	96	48
	31 to 40 years	91	46
	40 years above	13	6
	Total	200	100
Marital Status	Married	147	73
	Unmarried	53	27
	Total	200	100
Educational qualification	PG Only	55	27
	M.Phil	132	66
	Ph.D	13	7
	Total	200	100
Monthly Salary	Below 8000	46	23
	8001-10000	98	49
	10001-15000	27	13
	15001 above	29	15
	Total	200	100.0

From the above table it can be inferred that 48% of the faculty respondents were below the age group 30 years, 46% of the respondents were in the age group of 31-40 years and 6% of the respondents were in the age group of above 40 years. 73% of the respondents were married and 27% were unmarried. 27% of the female faculty respondents have the educational qualification of Post graduate degree, 66 % have completed M. Phil and 7% of the respondents have completed the Ph.D. As far as monthly salary is concerned, 23% of the respondents earn less than Rs.8000, 49% of the respondents earn Rs.8,001-Rs.10,000, 13% earn Rs.10,001-Rs.15,000 and 15% earn above 15,001.

COMPENSATION IMPACT TOWARDS JOB SATISFACTION

In order to find out the Compensation factors influencing job satisfaction among the female faculty members, mean values are found out and the ranks are assigned based on the mean values.

Table No: 2**Compensation impact towards Job Satisfaction**

Compensation factors	Mean	Std. Deviation	Rank based on mean scores
The benefits like loan facilities, free tour facilities, free education for children festival allowances	4.28	2.13	6
Provision of Provident fund is available	4.56	2.01	3
Cash awards or special gifts for academic excellence	4.14	2.19	7
Salary is adequate	3.92	2.07	8
Special awards for publication	4.46	2.19	4
Promotion	4.35	2.09	5
Leave salary available	4.61	2.14	1
There is annual increment	4.59	2.13	2

The above table reveals that Leave salary available is the first and foremost important compensation factors of job satisfaction, followed by annual increment, provident fund and Special awards for publication etc. The least important is Salary.

WORKPLACE CONDITION IMPACT TOWARDS JOB SATISFACTION

In order to find out the Workplace condition factors influencing job satisfaction among the female faculty members, mean values are found out and the ranks are assigned based on the mean values.

Table No: 3

Workplace Condition impact towards Job Satisfaction

Workplace Condition factors	Mean	Std. Deviation	Rank based on mean scores
The management is supportive and encouraging	5.41	1.78	1
There is a delegation and decentralization of work	5.09	1.93	6
Sufficient time is given for lesson plan preparation assignment and paper correction etc.	4.92	1.87	8
Allotment of workload is sufficient	5.07	1.79	7
Relationship with students is good	5.13	1.72	5
Relationship with other department faculty is good	5.16	1.82	4
There is a good treatment	5.40	1.57	2
Cooperation and Coordination are good	5.32	1.53	3
My superior encourage new ideas at all levels	4.80	1.92	9
There is freedom in decision making	4.70	2.04	10

The above table reveals that, The management is supportive and encouraging available is the first and foremost important workplace condition factors for job satisfaction,

followed by There is a good treatment, Cooperation and Coordination are good, Relationship with other department faculty is good etc., The least important is freedom in decision making.

RELATIONSHIP BETWEEN COMPENSATION AND WORKPLACE CONDITION WITH JOB SATISFACTION

In order to test the strength of relationship between Compensation, Workplace Condition and Jobs satisfaction level, correlation analysis is performed.

Table No: 4

	Compensation	Workplace condition	Job Satisfaction
Compensation	1	.512**	.856**
Workplace condition	.512**	1	.882**
Job Satisfaction	.856**	.882**	1

** . Correlation is significant at the 0.01 level (2-tailed).

It can be noticed from the above correlation table that there is a moderate positive & significant relationship between compensation and workplace condition ($r = 0.512$); high positive & significant relationship between compensation and job satisfaction ($r = 0.856$); high positive & significant relationship between workplace condition and job satisfaction ($r = 0.882$).

4. RESULT AND DISCUSSION

From the analysis is concluded that majority (48%) of the faculty respondents were belonging to the age group of below 30 years, 73% of the respondents were married, 66% of the respondents were M.Phil qualification and 49% of the respondents were salary package is 8000-10000. There are most important factors of compensation on job satisfaction is the highest mean rating score of 4.61 if for the item is leave salary, lowest mean rating score of 2.92 is for salary. There are most important factors of workplace condition on job satisfaction is the highest mean rating score of 5.41 if for the item is the management is supportive and encouraging, lowest mean rating score of 4.70 is for the item is freedom in decision making. The correlation analysis result indicated that, (i) there is moderate positive & significant relationship between Compensation and Workplace Condition, (ii) there is High Positive and significant relationship between Compensation and Job Satisfaction by same result of previous research like Shaunak Roy (2019), Sopiah (2013), Yaseen (2013), Muguongo et al.

(2015), Huma Bilal (2012), Salisu et al. (2017) and also Mabaso and Dlamini (2017). (iii) there is High Positive & significant relationship between Workplace Condition and Job Satisfaction by same result of previous research like by Muhammad et al. (2015), Abdul and Raheela (2015) Salunke (2015), Raziq and Maulabakhsh (2015) Agbozo et al. (2017), and Lee and Brand (2005).

5. CONCLUSION

The compensation and working conditions are important factor of their overall job satisfaction. The self-financing college teachers are not satisfied with their current salary and also freedom in decision making. But aided & government teachers are getting higher salary and higher status. The result of study indicated that there is Compensation and Working condition has a positive & significant impact towards Job satisfaction of College Teachers.

SCOPE FOR FURTHER RESEARCH

The following are some of the scope for further research base on the above study.

1. Job Satisfaction of teachers in the Engineering Colleges of Tirupattur District.
2. Data can be procured from regression carry out the studies.

REFERENCES:

1. Abdul Raziq and Raheela Maulabakhsh (2015) "Impact of Working Environment on Job Satisfaction", *Procedia Economics and Finance*, Vol.23, pp.717-725. www.elsevier.com/locate/procedia.
2. Arul Edwin Fredrick, P. (2015) "A study about the various factors influencing job satisfaction of MBA Teachers in college with special reference to Madurai District, India.", *International conference on inter disciplinary research in engineering and technology, (ICIDRET)*. I, article id icidret 21, www.icidret.in, pp.137-142.
3. Agbozo, G., Kafui, Owusu, I., Sakyi, Hoedoafia, Mabel, A. Atakorah, Y., Boateng, (2017) "The Effect of Work Environment on Job Satisfaction: Evidence from the Banking Sector in Ghana", *Journal of Human Resource Management*, Vol. 5(1), pp. 12-18.
4. Huma Bilal (2012) "Job satisfaction of University Teachers: Impact of Working Conditions and Compensation", *Review of Integrative Business & Economics Research*, Vol. 1(1) pp.101-113. Society of Interdisciplinary Business Research (www.sibresearch.org)

5. Ilias Kroupis, Thomas Kourtessis, Olga Kouli, George Tzetzis, Vassiliki Derri, George Mavrommatis (2017 March) "Job Satisfaction and burnout among Greek P.E. teachers. A comparison of educational sectors, level and gender, Cultura CCD Vol.12, No.13, pp.5-14. ISSN: 1696-5043 CCD 34.
6. Komang Oka Permadi, Nengah Landra, Gusti Agung Eka Teja Kumuma, Nengah Sudja (2019) "The Impact of Compensation and Work Environment towards Job Satisfaction of Affect the employee performances", International Journal of Management and Commerce Innovation, Vol.6, Issue No.2, pp.1248-1258. ISSN: 2348-7585(online) www.researchpublish.com.
7. Lee, S. Y., and Brand, J. L. (2005) "Effects of control over office workspace on perceptions of the work environment and work outcomes", Journal of Environmental Psychology, vol.25, pp.323-333.
8. Lane, K., Esser, J., Holte, B., & Anne, M. M. (2010) "A study of nurse faculty job satisfaction in community colleges in Florida", Teaching and Learning in Nursing, Vol. 5(1), pp.16-26.
9. Muguongo, M., Makena, Muguna, Andrew, T., Muriithi, Dennis K. (2015) "Effects of Compensation on Job Satisfaction Among Secondary School Teachers in Maara Sub – County of Tharaka Nithi County, Kenya" *Journal of Human Resource Management*, Vol. 3 (6), pp. 47-59.
10. Muhammad, Ghulam; Rehaman, Shafiq-ur- Ahmed, Nadeem (2015) "Impact of Work Environment on Teachers' Job Satisfaction A Case Study of Private Business Universities of Pakistan", European Journal of Business and Management, Vol. 7, No. 13, pp.55-68
11. Mabaso, C., Mzwenhlanhla, dan Dlamini, Bongani Innocent (2017) "Impact of Compensation and Benefits on Job Satisfaction", Research Journal of Business Management, Vol. 11(2), pp.80-90.
12. Rivai, V., and Sagala, E. J. (2011) "Manajemen sumber daya manusia untuk perusahaan: dari teori ke praktik. Edisi ke-2. Jakarta: PT. Raja Grafindo Persada, pp.118-241.
13. Raziq, Abdul, Maulabakhsh and Raheela. (2015) "Impact of Working Environment on Job Satisfaction", Procedia Economics and Finance, Vol. 23, pp.17-725.
14. Senthil Kumar, V. (2015 July) "A Study on Job Satisfaction of Higher Secondary School Teachers at Trichy", *IOSR Journal of Business and Management (IOSR-JBM)*,

Vol.17, Issue No. 7, ver. III, pp.12- 16 www.iosrjournal.org e-ISSN: 2278-487x, p-ISSN: 2319-7668.

15. Salunke, Ganesh. (2015) "Work Environment and Its Effect on Job Satisfaction in Cooperative Sugar Factories in Maharashtra, India", *Abhinav International Monthly Referred Journal of Research in Management & Technology*, Vol. 4, issue 5, pp.97-115.
16. Sopiah (2013) "The Effect of Compensation toward Job Satisfaction and Job Performance of Outsourcing Employees of Syariah Banks in Malang Indonesia", *International Journal of Learning & Development*, Vol. 3, No. 2, pp.45-58.
17. Salisu, Jamilu, B., Chinyio, Ezekiel, Suresh, Subhasini (2015) "The Impact of Compensation on the Job Satisfaction of Public Sector Construction Worker of Jigawa State of Nigeria", *The Business and Management Review*, Vol. 6, No. 4, pp.16-28.
18. Sedarmayanti. (2014) "Sumber Daya Manusia dan Produktivitas Kerja. Jakarta: Mandar Maju, pp.21-28.
19. Unnamalai, T, (2015), "A study on factors influencing job satisfaction of faculty members (with special reference to Arts and Science Colleges in Tiruchirapalli)", *International Journal of Management, (IJM) www.Iaeme.com* 6(1), 161-170.

AN ANALOGOUS STUDY BETWEEN WATERFALL AND INCREMENTAL SOFTWARE DEVELOPMENT LIFECYCLE MODEL

*Ms. S. Abirami,
Head & Assistant Professor
Department of Computer Application,
Marudhar Kesari Jain College for Women, Tamil Nadu, India.*

ABSTRACT

The purpose of this research is to measure the impact of compensation and workplace condition towards the job satisfaction Software development life cycle or SDLC for brief may be a methodology for planning, building, and maintaining info and industrial systems. There square measure numerous SDLC models wide used for developing software package. SDLC models provides a theoretical guide line concerning development of the software package. software package Development Life Cycle (SDLC) methodologies square measure mechanisms to assure that software package meet established needs. These methodologies impose numerous degrees of discipline to the software package development method with the goal of constructing the method additional economical and sure. SDLC models square measure vital for developing the software package in a very systematic manner such it'll be delivered among the time point and may even have correct quality. Every SDLC has its blessings and downsides in keeping with that we tend to decide that model ought to be enforced underneath that conditions. Within the gift situation all software package systems square measure imperfect as a result of they can not be designed with mathematical or physical certainty. According SDLC every and each model has the advantage and disadvantages. The conception of system lifecycle models came into existence that stressed on the requirement to follow some structured approach towards building new or improved system. For this we want to match SDLC models. during this paper we are going to compare 2 completely different noted life cycle models like-waterfall model & progressive model.

Keywords: Increment model , falls model , software package engineering , software package method model and software package

1. INTRODUCTION

In today's life computer is taken into account a time- saving device and its progress helps in executing complex, long, repeated processes during a very short time with a high speed. additionally to using computer for work, people use it for fun and entertainment. Noticeably, the amount of companies that produce software programs for the aim of facilitating works of offices, administrations, banks, etc, has increased recently. During the last few decades, software has been developed from a tool used for analyzing information or solving a controversy to a product in itself. However, the first programming stages have created variety of problems turning software an obstacle to software development particularly those counting on computers. Software consists of documents and programs that contain a set that has been established to be a component of software engineering procedures. Moreover, the aim of software engineering is to make an acceptable work that constructs programs of prime quality. A software development process, also referred to as a software development life cycle (SDLC), could be a structure imposed on the event of a merchandise. the method of building computer software and data systems has been always dictated by different development methodologies. A software development methodology refers to the framework that's accustomed plan, manage, and control the method of developing an system. Software development life cycle (SDLC) could be a method by which the software are often developed in a very systematic and which is able to increase the probability of completing the software project within the time deadline and maintaining the standard of the wares as per the quality. The System Development Life Cycle framework provides a sequence of activities for system designers and developers to follow for developing software. it's often considered as a subset of system development life cycle. Any software development process is split into several logical stages that allow a software development company to arrange its work efficiently so as to create a software package of the specified functionality within a particular timeframe and budget.

The development models are the assorted processes or methodologies that are being designated for the event of the project betting on the project's aims and goals. There are several development life cycle models that are developed so as to realize completely different needed objectives. The models specify the assorted stages of the method and also the order during which they're administrated. The selection of model

has terribly high impact on the testing that's administrated. it'll outline the what, wherever and once of our planned testing, influence regression checking and mostly determines that test techniques to use. There are varied package development models or methodologies. they're as follows:

1. Waterfall model
2. V model
3. Incremental model
4. RAD model
5. Agile model
6. Iterative model
7. Spiral model

2. SDLC PHASES

Software development life cycle (SDLC) could be a method by which the software will be developed in an exceedingly systematic manner and which can increase the probability of completing the software project within the time deadline and maintaining the standard of the software package as per the quality. Software life cycle models describe phases of the software cycle and also the order during which those phases are executed. Each phase produces deliverables required by the following innovate the life cycle. Requirements are translated into design. Code is produced in line with the look which is named development phase. After coding and development the testing verifies the deliverable of the implementation phase against requirements.

There are following six phases in every Software development life cycle model:

Requirement gathering and analysis: Business requirements are gathered during this phase. This phase is that the main focus of the project managers and stake holders. Meetings with managers, stake holders and users are held so as to see the wants like; Who goes to use the system? How will they use the system? What data should be input into the system? What data should be output by the system? These are general questions that get answered during a requirements gathering phase. After requirement gathering these requirements are analyzed for his or her validity and also the possibility of incorporating the necessities within the system to be development. Finally, a Requirement Specification document is made which serves the aim of guideline for the following phase of the model.

Design: during this phase the system and software design is ready from the need specifications which were studied within the first phase. System Design helps in specifying hardware and system requirements and also helps in defining overall system architecture. The system design specifications function input for the subsequent phase of the model.

Implementation /Coding: On receiver design documents, the work is split in modules/units and actual coding is started. Since, during this phase the code is produced so it's the most focus for the developer. this can be the longest phase of the software development life cycle.

Testing: After the code is developed it's tested against the wants to create sure that the merchandise is truly solving the requirements addressed and gathered during the wants phase. During this phase unit testing, integration testing, system testing, acceptance testing are done.

Deployment: After successful testing the merchandise is delivered / deployed to the customer for his or her use.

Maintenance: Once when the shoppers starts using the developed system then the particular problems comes up and desires to be solved from time to time. This process where the care is taken for the developed product is understood as maintenance.

3. WATERFALLMODEL

Waterfall model was planned by Royce in 1970 that could be a linear ordered package development life cycle (SDLC) model. The falls model is that the classical model of package engineering. This model is one in all the oldest models and is wide employed in government comes and in several major corporations. As this model emphasizes coming up with in early stages, it ensures style flaws before they develop. thence it's not abundantly helpful once the project necessities area unit dynamic in nature The model begins with establishing system necessities and package necessities and continues with beaux arts style, elaborated style, coding, testing, and maintenance. The falls model is a baseline for several different lifecycle models. In some organizations, a modification electrical device maintains the standard of the merchandise by reviewing every modification created within the maintenance stage. think about applying the total falls development cycle model once correcting issues or implementing these sweetening requests. In every stage, documents that designate the

objectives and describe the necessities for that part area unit created. At the tip of every stage, a review to work out whether or not the project will proceed to following stage is control. Many folks believe that this model can't be applied to any or all things. In real-world development, however, one will discover problems throughout the look or secret writing stages that time out errors or gaps within the necessities. The falls methodology doesn't proscribe returning to associate earlier part, for instance, arriving from the look part to the necessities part. However, this involves expensive process. Every completed part needs formal review and intensive documentation development. Thus, oversights created within the necessities part area unit overpriced to correct later. as a result of the particular development comes late within the method, one doesn't see results for a protracted time. This delay may be displeasing to management and customers. many folks conjointly suppose that the quantity of documentation is excessive and inflexible. though the falls model has its weaknesses, it's instructive as a result of it emphasizes vital stages of project development. even though one doesn't apply this model, he should think about every of those stages and its relationship to his own project.

Following is a diagrammatic representation of different phases of waterfall model.

Advantages:

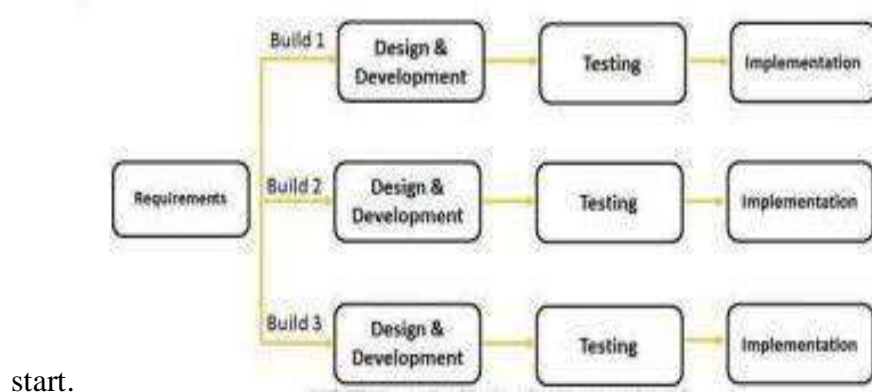
- ☐ Easy to understand and implement.
- ☐ Widely used and known
- ☐ Reinforces good habits
- ☐ Identifies deliverables and milestones
- ☐ Works well on mature products and weak teams.

Disadvantages:

- ☐ Idealized, doesn't match reality well.
- ☐ Doesn't reflect iterative nature of exploratory development. Unrealistic
- ☐ to expect accurate requirements so early in project. Software is
- ☐ delivered late in project, delays discovery of serious errors. Difficult to
- ☐ integrate risk management.
- ☐ Difficult and expensive to make changes to documents.
- ☐ Significant administrative overhead, costly for small teams and projects

4. INCREMENTALMODEL

The progressive approach could be a methodology of package development wherever the model is meant, enforced and tested incrementally (a very little a lot of is additional every time) till the merchandise is finished. It involves each development and maintenance. the merchandise is outlined as finished once it satisfies all of its necessities. The progressive model is Associate in Nursing intuitive approach to the water model. Multiple development cycles happen here, creating the life cycle a “multi-waterfall” cycle. Cycles ar divided into smaller, a lot of simply managed iterations. every iteration passes through the necessities, design, implementation and testing phases. The progressive model divides the merchandise into builds, wherever sections of the project ar created and tested individually. This approach can doubtless realize errors in user necessities quickly, since user feedback is invited for every stage and since code is tested sooner when it's written. In progressive models, as in ordered models, the necessities of the ultimate system or product ar familiar at the beginning of the event. In progressive models but a restricted set of necessities is allotted {to every/to every} increment and with each consecutive (internal) unleash a lot of necessities ar addressed till the ultimate (external) unleash satisfies all necessities. A project that is victimization the progressive model could begin with general objectives, that is additionally outlined as provisions and people necessities ar enforced by following subsequent coming back portion of the objectives, until all objectives ar performed it push tough issues to the long run to prove early success and this model may be helpful for the comes wherever the fundamental package practicality is needed at the



Advantages

- ☐ Generates working software quickly and early during the software lifecycle.
- ☐ More flexible—less costly to change scope and requirements.
- ☐ Easier to test and debug during a small iteration.
- ☐ Easier to manage risk because risky pieces are identified and handled during its iteration.

- ☐ Each iteration is an easily managed milestone.

Disadvantages

- ☐ Each phase fan iteration is rigid and do not overlap each other.
 - Problems may arise pertaining to system architecture because not all requirements are gather edup front for the entire software lifecycle.

Because of the weaknesses shown above, the application of the waterfall model should be limited to situations where the requirements and the implementation of those requirements are very well understood.

5. COMPARISON

The water model consists of parts that area unit completed consecutive before continuing to successive phase. For comparison to alternative models, the salient attributes of the water model area unit that it's

- A formal technique.
- A kind of top-down development.
- Composed of freelance phases to be done consecutive.
- Used in varied ways in which.

Because of the weaknesses shown higher than, the applying of the water model ought to be restricted to things wherever necessities the want [the necessities} and therefore the implementation of these requirements area unit alright understood.It is developed to beat the weaknesses of the water model. The progressive model performs the water in overlapping sections making an attempt to make amends for the length of water model comes by manufacturing usable practicality earlier. this might involve a whole direct set of necessities that area unit enforced in a very series of tiny comes. As an alternate, a project victimization the progressive model could begin with general objectives. Then some portion of those objectives is outlined as necessities and is enforced, followed by successive portion of the objectives till all objectives area unit enforced. But, use of general objectives instead of complete necessities may be uncomfortable for management. as a result of some modules are going to be completed long before others, well-defined interfaces area unit needed. Also, formal reviews and audits area unit tougher to implement on increments than on a whole

system. Finally, there may be a bent to push troublesome issues to the long run to demonstrate early success to manage me.

Comparison of Waterfall model and Incremental Model

Features	WaterfallModel	IncrementalModel
Requirementspecifications	Beginning	Beginning
Cost	Low	Low
Simplicity	Simple	Intermediate
Riskinvolvement	high	Easilymanageable
Expertise	High	High
Flexibilitytochange	Difficult	Easy
Userinvolvement	Onlyatbeginning	Intermediate
Flexibility	Rigid	Lessflexible
Maintenance	Least	Promotesmaintainability
Duration	Long	Verylong

6. CONCLUSIONS

There square measure several SDLC models like, Waterfall, RAD, spiral, progressive, formed etc. utilized in varied organizations relying upon the conditions prevailing there. of these totally different code development models have their own benefits and downsides. within the code business, the hybrid of of these methodologies is employed i.e with some modification.

In this paper, we tend to targeted on the 2 {different|totally totally different|completely different} models of code development life cycle- falls model & progressive model and compare them with different aspects. as an example, if a corporation has expertise in building accounting systems, I/O controllers, or compilers, then building another such product supported the present styles is best managed with the falls model and if it's too risky to develop the complete system quickly, then the progressive development ought to be thought - about choosing the right life cycle model is very necessary during a code business because the code has got to be delivered at intervals the time point in time & ought to even have the specified quality. This study can build the method of choosing the SDLC model easy& thus can convince be terribly effective for code business.

REFERENCES

1. M.Davis,H.Bersoff,E.R.Comer,“A Strategy for Comparing Alternative Software Development Life Cycle Models”, Journal IEEE TransactionsonSoftwareEngineering ,Vol.14,Issue10,1988
2. Jovanovich,D.,Dogs,T.,“Comparisonofsoftware development models,” Proceedingsof the 7th International Conference on, 11-13June2003, ConTEL2003,pp. 587-592.
3. Klopper,R.,Gruner,S.,&Kourie,D.(2007),“Assessmen to faframe work to compare Software development methodologies” Proceedings of the 2007Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries,56-65.doi:10.1145/1292491.1292498
4. Khuranagouravandsgupta(2012)“Study&ComparisonofSoftwareDevelopmentLifeC ycleModels”IJREAS,Vol.2(2),1514-1515.

ANALYSIS OF HOUSING DATA TO PREDICT HOUSE PRICE USING MULTIPLE LINEAR REGRESSION

Dr.G.Dona Rashmi¹, Dr. S.Poongodi²

Assistant Professor, Department of Data Analytics (PG)¹,

PSGR Krishnammal College for Women,Tamilnadu,India

ABSTRACT

A linear regression model extended to include more than one independent variable is called a multiple regression model. It is more accurate than to the simple regression. The purposes of multiple regressions are: planning and control, prediction or forecasting. The principal advantage of multiple regression models is that it gives us more of the information available to us who estimate the dependent variable. It also enables us to fit curves as well as lines [1].“Prediction” refers to the output of an algorithm after it has been trained on a historical dataset and applied to new data when forecasting the likelihood of a particular outcome, The problem that we are going to solve here is that given a set of features that describe a house in Boston, our learning model must predict the house price. We used R package for implementation and visualization of results.

KEYWORDS: Multiple linear model, mean square error, Regression model

1. INTRODUCTION

Linear regression is a statistical analysis which depends on modeling a relationship between two kinds of variables, dependent (response) and independent (predictor). The main purpose of regression is to examine if the independent variables are successful in predicting the outcome variable and which independent variables are significant predictors of the outcome [5]. A set of techniques used to analyze the relationship between two or more independent variables and a dependent variable which defines multiple regressions.. The linear regression equation has the following form when there is just one independent variable:

$$\hat{Y} = a + bX$$

The equation is extended to include the additional variables when there are more than one independent variable.

When more than one explanatory variable is included in a "model," multiple linear regression is utilized as an analysis procedure. That is, when we believe there is more than one explanatory variable that might help “explain” or “predict” the response variable, we’ll

put all of these explanatory variables into the “model” and perform a multiple linear regression analysis[4].

2. MULTIPLE LINEAR REGRESSION:

The multiple linear regression model is a simple linear regression model that has been extended. The explanatory variable is represented by a “x” in simple linear regression. We'll have more than one explanatory variable in multiple linear regressions, thus there will be more than one "x" in the equation. We'll use subscripts next to the "x's" in the equation to distinguish between the explanatory variables.[4]

3. MULTIPLE LINEAR REGRESSION MODEL:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_v x_v + \varepsilon$$

where y is an observed value of the response variable for a particular observation in the population

β_0 = the constant term (equivalent to the “y-intercept” in Simple Linear Regression)

β_j = the coefficient for the jth explanatory variable ($j = 1, 2, \dots, v$)

x_j = a value of the j th explanatory variable for a particular observation ($j = 1, 2, \dots, v$)

ε = the residual for the particular observation in the population

Multiple linear regression can be used in the following situations:

- Trying to forecast an individual's income based on a variety of socioeconomic factors.
- Trying to predict pupils' overall exam performance in ‘A’ levels based on values of a set of exam scores at the age of 16.
- Trying to estimate systolic or diastolic blood pressure based on a variety of socioeconomic and behavioural factors (occupation, drinking smoking, age etc.). [2]

This analysis does not allow us to make causal inferences, as it does with simple linear regression and correlation, but it does allow us to study how a set of explanatory variables is linked with a dependent variable of interest. In terms of a hypothesis test, the null hypothesis, H_0 , in the case of a simple linear regression is that the coefficient connecting the explanatory (x) variable to the dependent (y) variable is 0. To put it another way, the explanatory variable and the dependent variable have no relationship. H_1 is the alternative hypothesis, which states that the coefficient relating the x and y variables is not zero. To look at it another way, there is a relationship between x and y.

4. 'R' TOOL

R packages are a collection of R functions, compiled code and sample data. They are stored under a directory called "library" in the R environment. By default, R installs a set of packages during installation. More packages are added later, when they are needed for some specific purpose [3].

RStudio is an integrated development environment (IDE) for R. It comes with a console editor with syntax highlighting and direct code execution, as well as graphing, history, debugging, and workspace management capabilities.

FEATURES OF R PROGRAMMING

- ✓ Open-source. R is an open-source software environment.
- ✓ Highly Active Community. ...
- ✓ A Wide Selection of Packages. ...
- ✓ Comprehensive Environment. ...
- ✓ Can Perform Complex Statistical Calculations. ...
- ✓ Distributed Computing. ...
- ✓ Running Code Without a Compiler.

5. PERFORMANCE METRICS MEAN SQUARED ERROR

Mean Squared Error, or MSE for short, is a popular error metric for regression problems. It is also an important loss function for algorithms fit or optimized using the least squares framing of a regression problem

MSE is calculated by the sum of square of prediction error which is real output minus predicted output and then divide by the number of data points. It gives you an absolute number on how much your predicted results deviate from the actual number.

Root Mean Square Error(RMSE) is the square root of MSE. It is used more commonly than MSE because firstly sometimes MSE value can be too big to compare easily[1]. Secondly, MSE is calculated by the square of error, and thus square root brings it back to the same level of prediction error and makes it easier for interpretation.

While R Square is a relative measure of how well the model fits dependent variables, Mean Square Error is an absolute measure of the goodness for fit

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

6. IMPLEMENTATION & RESULT DISCUSSION

A Dataset derived from information collected by the U.S. Census Service concerning housing in the area of Boston Mass. This dataset contains information collected by the U.S Census Service concerning housing in the area of Boston Mass. The name for this dataset is simply boston. It has two proto tasks: nox, in which the nitrous oxide level is to be predicted; and price, in which the median value of a home is to be predicted

In each case of the dataset, there are 14 attributes. They are as follows:

1. CRIM - per capita crime rate by town
2. ZN - proportion of residential land zoned for lots over 25,000 sq.ft.
3. INDUS - proportion of non-retail business acres per town.
4. CHAS - Charles River dummy variable (1 if tract bounds river; 0 otherwise)
5. NOX - nitric oxides concentration (parts per 10 million)
6. RM - average number of rooms per dwelling
7. AGE - proportion of owner-occupied units built prior to 1940
8. DIS - weighted distances to five Boston employment centres
9. RAD - index of accessibility to radial highways
10. TAX - full-value property-tax rate per \$10,000
11. PTRATIO - pupil-teacher ratio by town
12. B - $1000(B_k - 0.63)^2$ where B_k is the proportion of blacks by town
13. LSTAT - % lower status of the population
14. MEDV - Median value of owner-occupied homes in \$1000's

Install and load all of the essential packages. Load the housing dataset for Boston. Using the str() function, you can see the attribute details and values. After that, divide the dataset in half for training and testing, 70 percent for training and 30 percent for testing. Build the model and check out the summary. Now it's time to test the new records with the model. Plot the results of the tests and assess them as indicated below.

```

> library(ggplot2)
> library(reshape2)
> library(caTools)
> house<-read.csv("C:/Users/HP/Desktop/R LAB/boston.csv")
> str(house)
'data.frame':   506 obs. of  14 variables:
  crim      : num  0.00632 0.02731 0.02729 0.03237 0.06905 ...
  zn        : num  18 0 0 0 0 12.5 12.5 12.5 12.5 ...
  indus     : num  2.31 7.07 7.07 2.18 2.18 2.18 7.87 7.87 7.87 ...
  chas      : int  0 0 0 0 0 0 0 0 0 ...
  nox       : num  0.538 0.469 0.469 0.458 0.458 0.458 0.524 0.524 0.524 ...
  rm        : num  6.58 6.42 7.18 7 7.15 ...
  age       : num  65.2 78.9 61.1 45.8 54.2 58.7 66.6 96.1 100 85.9 ...
  dis       : num  4.09 4.97 4.97 6.06 6.06 ...
  rad       : int  1 2 2 3 3 3 5 5 5 ...
  tax       : int  296 242 242 222 222 222 311 311 311 ...
  ptratio   : num  15.3 17.8 17.8 18.7 18.7 18.7 15.2 15.2 15.2 ...
  b         : num  397 397 393 395 397 ...
  lstat     : num  4.98 9.14 4.03 2.94 5.33 ...
  medv      : num  24 21.6 34.7 33.4 36.2 28.7 22.9 27.1 16.5 18.9 ...
> split<-sample.split(house,SplitRatio =0.75)
> train<-subset(house,split==TRUE)
> test<-subset(house,split==FALSE)
> model<-lm(medv ~ crim + rm + tax + lstat, data = train)
> test$predicted.medv <- predict(model,test)
> p1<-test %>%
++   ggplot(aes(medv,predicted.medv)) +
++   geom_point(alpha=0.5) +
++   stat_smooth(aes(colour='black')) +
++   xlab('Actual value of medv') +
++   ylab('Predicted value of medv')

```

Figure 1: Prediction Process

```

> error<-test$medv-test$predicted.medv
> rmse<-sqrt(mean(error)^2)
> rmse
[1] 0.08046412
>

```

Figure 2: Mean Square Error Values

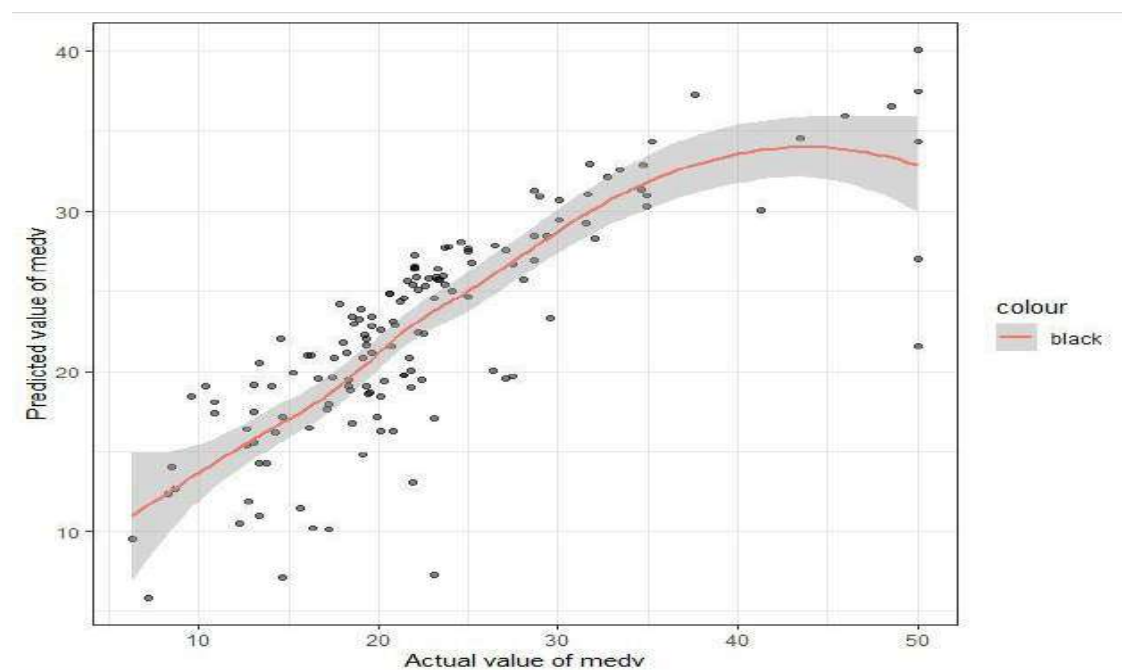


Figure 3. Visualization of Predicted Values

7. CONCLUSION

In this paper, the Multiple Linear Regression model is used to analyze the housing data to predict housing price in the city Boston. From the dataset the median value of house price is the outcome or the dependent variable in our analysis is predicted and Mean Squared Error is calculated to the same level of prediction error and simplifies understanding.

REFERENCES

1. Turóczy Zsuzsannaa, Liviu Mariana, “Multiple regression analysis of performance indicators in the ceramic industry”, Elsevier,2012, under responsibility of Emerging Markets Queries in Finance and Business local organization
2. Guiden Kaya Uyanik, Nese Gule, “Astudy on Multiple Linear Regression Analysis” Procedia- Social and Behavioral Science, Dec 2013
3. <https://academia.stackexchange.com/questions/27921/should-i-cite-all-r-packages-i-used>
4. Rebecca Bevans, “ An Introduction to Multiple Linear Regression”, Scribber, Feb 2020
5. Fuyuki Yoshikane , Yutaka Suzuki, “Multiple Regression Analysis between Citation Frequency of Patents and their Quantitative Characteristics”, The 2nd International Conference on Integrated Information , Procedia - Social and Behavioral Sciences 73 (2013) 217 – 223,Elsevier 2013.

SECURITY IN IoT: OVERVIEW

Ms.S. Uma Mageshwari¹ , Dr.R. Santhi²

Research Scholar¹, Research Supervisor²,

Research & Evaluation Centre,

Govt Arts and Science College Tamil Nadu, India.

ABSTRACT

As far as IoT is concerned, there are lot of challenges and security issues need to be focussed to protect the device from hacking. Since IoT nodes are working based on wireless technology, there can be loopholes in the communication between the devices. To protect the IoT device and to achieve privacy, confidentiality, and authentication the various security concerns need to be identified and resolved with suitable methodology. This paper highlights the Security problems in IoT and discuss the solutions with different approaches to resolve the network security threats in IoT.

Keywords: Security, IoT, Threats.

1.INTRODUCTION

The IoT makes the world to work smartly with the wireless connection among the devices. In this era, the advancement in technology makes the IoT device easily vulnerable by the intruder in unsecured network connection. Therefore, a depth study on IoT security problems must be carried out to identify and resolve the various security threats for each device. The IoT architecture consists of three layers. They are, Perception Layer, Network Layer and Application Layer. The common security requirements in all the layers are Authentication, Confidentiality, Privacy, Encryption, Key management, and Intrusion detection. The IoT attacks can be classified as Physical attack, Network attack, Software attack and Encryption attack. The purpose of IoT attacks are given below,

- **Physical attack** - Grab unauthorised access
- **Network attack** - Extract data remotely
- **Software attack** - Risk in accessing the entire software
- **Encryption attack** - Devices are vulnerable due to lack of encryption feature

2. LITERATURE REVIEW

[1] Mohana.S and Lakshmi Priya T.K.S proposed a three stages communication methodology to achieve device to device Encryption in IoT.

[2] Swarnalatha.K et al., proposed a web-based architecture using GSM for the establishment of smart home.

[3] Sujithra and Dr.G.Padmavathi discussed the security aspects, IOT architecture and challenges in IoT.

[4] Irene Castaneda Lopez et al., proposed a framework which integrate WSN and IoT for smart cities.

3. METHODOLOGY

ENCRYPTION

The [1] Identity Based Encryption (IBE) perform device to device Encryption for IoT device. The Private Key Generator (PKG), use the public keys generated by the Public Key Infrastructure (PKI). The PKI creates the public key using the parameters such as name, email id, IP address etc...After the creation of keys, then PKG communicates to the IoT device and stay connected with another device. In this methodology, the devices are known as users. The IoT communication is carried out in three levels such as Registration, Communication (Registered devices) and Communication (Withdrawn condition). The roles of three levels in the IoT communication are as follows,

Registration

The new device will be registered in PKG. Whereas PKG produce pair of keys such as, public key and secret key for all the new devices involved in IoT communication

Safe Communication

Check the validity of the user. If it is valid, PKG assign public key to both server and client. Otherwise, PKG produce illegal user.

Withdrawn Condition

The PKG runs and check the wrong user by using the algorithm.

GSM Technology

The GSM technology [2] is used to establish communication between the smart home and IoT agent. The Web Users, Web Server and IoT agent are the three important factors in this

approach to establish secured communication. The roles of Web Users, Web Server and IoT agent in the architecture is depicted below,

- Web Users – Any user can control the devices with commands
- Web Server – Maintains the database and act based on the request and response message from the user.
- IoT Agent – Monitors SMS continuously between Web Server and GSM module

The GSM act as an interface which connects the computer with microcontroller through cable. In the Smart home, all the devices are controlled by the microcontroller. For every new SMS, the microcontroller read and execute the message. As well as microcontroller read the device status (ON and OFF State). Then, send the SMS to IOT agent through RS232 protocol.

M2M Protocol for WSN

This approach integrates [4] the Wireless Sensor Networks (WSN) using M2M protocol and FIWARE platform in the framework of IOT. The FIWARE platform provides the flexibility regarding connection and applications to be used in smart cities. The WSN is constructed on Zigbee technology using API protocol. The platform is constructed exactly to do service according to the request. This methodology can be used in various sectors such as transportation, environment, government and so on.

4. CONCLUSION

In the upcoming days, the IOT paves the strong impact in all the places. The IOT makes the users more convenient than before. Such IOT devices can be sniffed by the hackers during communication. In this paper, the three different methodology to secure IOT environment are illustrated briefly. Therefore, suitable technology must be identified and can be used in IOT devices to safeguard and provide the secured way of communication between the devices.

5. REFERENCES

- [1] Mohana. S, Lakshmi Priya T.K.S, “Identity-based Encryption for device-to-device Security in IoT environments”, World Scientific News, WSN 41(2016) 120-130, EISSN: 2392-2192.

- [2] Swarnalatha.K et al., “Web Based Architecture for Internet of Things Using GSM for Implementation of Smart Home Applications”, World Scientific News, WSN 41(2016) 40-50, EISSN: 2392-2192.
- [3] Sujithra, Dr.G.Padmavathi “IOT Security Challenges and Issues – An Overview”, World Scientific News, WSN 41(2016) 232- 238, EISSN: 2392-2192.
- [4] Irene Castaneda et al., “Implementation of M2M protocol for Wireless Sensor Networks Oriented to its use in Smart Cities supported by IoT”, International Journal of Applied Engineering Research, ISSN: 0973-4562, Vol. 13, Number 5(2018) pp. 2624-2629.
- [5] Sattar B.Sadkhan , Zainab Hamza, “ Cryptosystems used in IOT –Current Status and Challenges”, 2017 International Conference on Current Research in Computer Science and Information Technology(ICCIT), Slemani,Iraq.
- [6] Shancang Li et.al., “The Internet of Things: a security point of view”, Internet Research, Vol. 26, DOI: 10.1108/IntR-07-2014-0173, April 2016.
- [7] Singh M, Rajan M A, Shivraj V L and Balamuralidhar P, "Secure MQTT for Internet of Things (IoT)," 2015 Fifth International Conference on Communication Systems and Network Technologies, Gwalior, 2015, pp. 746-751.
- [8] Mussab Alaa et.al., “A review of smart home applications based on Internet of Things”, Journal of Network and Computer Applications, September 2017, <http://dx.doi.org/10.1016/j.jnca.2017.08.017>.
- [9] Nadeem Abbas, “A Mechanism for Securing IOT –enabled Applications at the Fog layer”, Journal of Sensor and Actuator Networks, 2019.
- [10] Mohammed El-hajj et.al., “ A Survey of Internet of Things (IoT) Authentication Schemes”, Sensors 2019, DOI:10.3390/s19051141.

AN ANALYSIS OF OPINION MINING USING DIFFERENT MACHINE LEARNING ALGORITHMS

Mrs. N. Sathya Priya

Assistant Professor,

Department of Computer Science,

Sri GVG Visalakshi College for Women, Tamilnadu.

Now-a-days due to the rapid growth of internet, people are expressing their views and opinions regarding, products, services and policies on the web in large numbers. This huge amount of feedback is very crucial for both organizations as well as individuals. The task of analysing these reviews is done by Opinion Mining (also known as Semantic Analysis). Sentiment analysis is a type of natural language processing for tracking the mood of the public about a particular product or topic. Sentiment analysis, which is also called opinion mining, involves in building a system to collect and examine opinions about the product made in blog posts, comments, reviews or tweets. Sentiment analysis offers the people to choose the right product and also offers the organization to improve the quality of their product. The objective of the proposed system is to classify the opinion into positive and negative with higher performance. Different machine learning algorithms were applied for the classification process.

Key Words: machine learning, Negative, Opinion, Opinion Mining, Positive, Reviews, Sentiment analysis, Techniques.

I. INTRODUCTION

The World Wide Web is growing at an alarming rate not only in size but also in the types of services and contents provided. Each and every user is participating more actively and are generating vast amount of new data. These new Web contents include customer reviews and blogs that express opinions on products and services – which are collectively referred to as customer feedback data on the Web.

Every day, millions of people are using websites like Facebook, Twitter, Orkut, Google+ and many more to share their view in form of status updates, tweets, blogs etc. With the explosive growth of social media (e.g., reviews, forum discussions, blogs, micro-blogs, Twitter, comments, and postings in social network sites) on the Web, individuals and organizations are increasingly using the content in these media for decision making.

Nowadays, if one wants to buy a consumer product, one is no longer limited to asking one's friends and family for opinions because there are many user reviews and discussions in public forums on the Web about the product.

II. OPINION MINING AND SENTIMENT ANALYSIS

Opinion Mining (OM) also called Sentiment analysis (SA) is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes.

The term Opinion Mining first appeared in (Dave, Lawrence and Pennock, 2003). Sentiment Analysis or Opinion Mining represents a large problem space. There are also many names and slightly different tasks, e.g., sentiment analysis, opinion mining, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, review mining, etc. The term Opinion Mining is commonly used in academia and Sentiment Analysis is in industry.

Users express their opinion through the reviews. These reviews are used by the individuals and organizations for decision making. Opinion Mining is a different problem that is to be solved due to the highly unstructured nature of natural language. It is basically the generation of subjective knowledge from source materials. This new trend offers many challenging research problems. The reviews and opinion usefulness is increasing day by day.

Opinion Definition:

An opinion is a quadruple, (g, s, h, t),

where g is the opinion (or sentiment) target,

s is the sentiment about the target,

h is the opinion holder and

t is the time when the opinion was expressed.

Components of Opinion Mining

Opinion Holder – Opinion Holder is the holder of a particular opinion, it may be a person or an organization. Opinion Holders are persons who write these reviews or blogs.

Opinion Object – Opinion Object is an object on which the opinion holder is expressing the opinion.

Opinion Orientation – Orientation of an opinion on an object determines whether the opinion of an opinion holder about an object is positive, negative or neutral.

Example: “According to James, this phone has excellent sound quality.”

Opinion Holder – James

Opinion Object – Sound quality of Phone

Opinion Orientation – Excellent

III. STRUCTURAL DESIGN OF OPINION MINING

The growth of user generated content on the web has offered a rich data source for opinion mining. The large number of diverse review sources challenges the individual users and organizations on how to use the opinion information effectively. Therefore, automated opinion mining and summarization techniques became important.

Sentiment Analysis is performed in three stages,

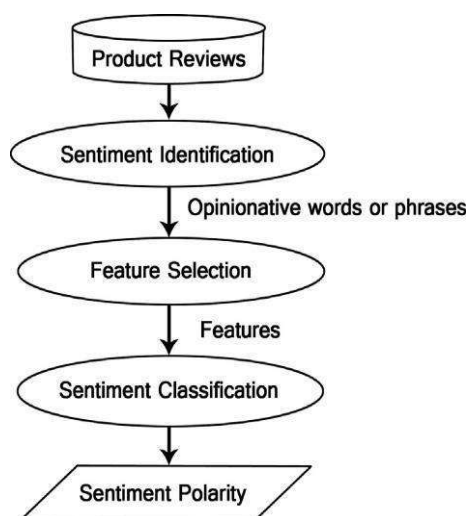


Fig:3.1 Structural Design of Opinion Mining

Data collection and Preprocessing

Data is collected from the web and the opinions are extracted. Preprocessing is done to remove all the unnecessary and irrelevant words. Keywords are extracted from the text that provides accurate classification.

Classification

The key step in opinion analysis is classification of review text. Classification is done to identify the polarity of content. Three classes positive, negative and neutral were used for classification. Review text is classified by calculating and averaging polarity score of individual score in sentences.

Aggregation and Result Presentation

Aggregation is done to determine the general opinion of the text. The results are presented by directly expressing the sentiment textual or using graphics.

IV. LITERATURE REVIEW

Opinion mining is an interesting area of research because of its applications in various fields. Many authors developed and proposed various techniques for opinion mining and sentiment analysis.

Ion Smeureanu, Cristian Bucur, in "Applying Supervised Opinion Mining Techniques on Online User Reviews" applied Supervised opinion mining techniques on online user reviews.

They used Naïve Bayes for opinion mining and produced an accuracy of 79%. This paper presents a method of sentiment analysis, on the review made by users to movies. Classification of reviews in both positive and negative classes is done based on a naive Bayes algorithm. As training data they used a collection of sentences taken from the movie reviews.

Vandana Korde in "Text Classification and Classifiers: A Survey" applied Naïve Bayes and other various classification techniques for text classification. They collected the different types (format) of document like html, pdf, doc, web content etc. and classified the text documents successfully.

Tripathy et al. [5] did a study on classifying movie reviews using different ML algorithms, which are naïve bayes (NB), maximum entropy (ME), stochastic gradient descent (SGD), and support vector machine (SVM). Their results show that support vector machines (SVM) acquired the highest accuracy.

Norlela Samsudin, Mazidah Puteh, Abdul Razak Hamdan, Mohd Zakree Ahmad Nazri [12] presented a paper "Immune based feature selection for Opinion Mining" in the proceedings of world congress. The study used a feature selection technique based on artificial immune system to select the appropriated features for opinion mining. They applied it in Naïve Bayes, KNN and SVM classifier and proved with better performance.

Pravesh Kumar Singh, Mohd Shahid [14] Husain presented a journal "Methodological Study of Opinion Mining and Sentiment Analysis Techniques". They reviewed and evaluated various techniques used for opinion mining and sentiment analysis. They used Naïve bayes,

SVM, Multilayer Perceptron and Clustering techniques individually for opinion mining and produced better results.

Poobana S, Sashi Rekha K[13] proposed an "Opinion Mining technique From Text Reviews Using Machine Learning Algorithm" by applying Naïve Bayes algorithm. a sentence level opinion mining is used and it is done by counting based approach which compare the opinion by count method.

Gagandeep Singh, Kamaljeet Kaur Mangat[8] presented a paper, “ Performance Analysis of Supervised Learning Methodologies for Sentiment Analysis of Tweets” and applied Naïve Bayes and support Vector Machine individually and produced an accuracy of 70 and 80 percentage. They applied the opinion mining technique for tweet datasets.

Preety and Sunny Dahiya[15] presented research article on “Sentiment analysis using SVM and Naïve Bayes algorithm”. They proposed a method using naïve bayes and modified k means clustering and found that it is more accurate than applying the techniques individually.

K.Umamaheswari, S.P.Rajamohana, G.Aiswaryalakshmi[22] presented a paper "Opinion mining using Hybrid Methods". They applied Particle Swarm optimization for feature extraction and applied SVM for classification of opinions. They got a result of 81% accuracy in the classification of opinion.

V. RESEARCH METHODOLOGY

The proposed method uses different machine learning algorithms to find the efficiency and accuracy of opinion classification. Various ML algorithms such as supervised learning methods Support Vector Machine, K-Nearest Neighbour, Maximum Entrophy and Naïve Bayes algorithm. The algorithms when used separately performs well and produces a better accuracy. Here a hybridized approach of Support Vector Machine (SVM) and Naïve Bayes classification algorithm is used for opinion mining. The algorithms when used individually produces a good accuracy in opinion classification whereas when used as hybridized produces better performance than applying individually.

Support Vector Machine (SVM)

SVM is a supervised learning model. This model is associated with a learning algorithm that analyzes the data and identifies the pattern for classification. The concept of SVM algorithm is based on decision plane that defines decision boundaries. A decision plane separates group of instances having different class memberships. Support vector machines are

the most efficient way for document classification. The basic idea behind SVM classification is to find a maximum margin hyper plane that separates the document vector in one class from the other with maximum margin.

Naive Bayes Classifier

Naive Bayes-classifier is a probabilistic classifier based on conditional probability. The probability for each data to be classified into positive or negative is calculated based on the Bayes theorem. Bayesian techniques use mathematical formulae in order to analyze the content of the message.

Bayesian classifier works on the events which are dependent and the probability of an event occurring in the future that can be detected from the previous occurring of the same event. Naive Bayes classifier performs extremely well for problems which are linearly separable and even for problems which are non-linearly separable it performs reasonably well. It's a probabilistic and supervised classifier given by Thomas Bayes.

According to this theorem, if there are two events say, e_1 and e_2 then the conditional probability of occurrence of event e_1 when e_2 has already occurred is given by the following mathematical formula:

$$P(e_1|e_2) = \frac{P(e_1, e_2)}{P(e_2)}$$

Maximum Entropy:

The contextual information of the document (unigrams, bigrams, other characteristics within the text) in order to categorize it to a given class (positive/negative). Following the standard bag-of-words framework that is commonly used in natural language processing and information retrieval, let $\{w_1, \dots, w_m\}$ be the m words that can appear in a document. Then each document is represented by a sparse array with 1s and 0s that indicate whether a particular word w_i exists or not in the context of the document.

The first step is to collect a large number of training data which consists of samples represented on the following format: (x_i, y_i) where the x_i includes the contextual information of the document (the sparse array) and y_i its class. The second step is to summarize the training sample in terms of its empirical probability distribution:

$$\hat{p}(x, y) = \frac{1}{N} \times \text{number of times that } (x, y) \text{ occurs in the sample}$$

Where N is the size of the training dataset.

Logistic Regression:

Logistic regression is a Supervised Learning technique and used for predicting the categorical dependent variable using a given set of independent variables.

Logistic regression predicts the output of a categorical dependent variable. Therefore the outcome must be a categorical or discrete value. It can be either Yes or No, 0 or 1, true or False, etc. it gives the probabilistic values which lie between 0 and 1.

In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1).

But we need range between $-\infty$ to $+\infty$, then take logarithm of the equation it will become:

$$\log \left[\frac{y}{1-y} \right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n$$

In the proposed system, Data is collected from Polarity review data set which consists of 1000 positive reviews and 1000 negative reviews. Totally 2000 review documents are collected from the Bo Pang and Lillian Lee et al. (URL <http://www.cs.cornell.edu/people/pabo/movie-review-data>).

Preprocessing is performed in the polarity data set in the following phases. They are

- Load the data from its subdirectories
- Convert into word vector
- Remove the stop words

Loading the Data :

In this stage, the data is loaded from the polarity review data set using text directory loader.

Word Vector Conversion:

In this second stage, the document is converted into structured data suitable for analysis. The document is expressed as a matrix of columns indicating the line that represents a keyword extracting a set of documents. This matrix is called word-document matrix, it can be expressed in the form of a matrix of $m \times n$ words and documents.

Stop Word Removal :

Removal of stop word is the third stage in the preprocessing. Stop words are the common words that appear in text which carries little meaning and serve only syntactic meaning but do not indicate subject matter; e.g. "the", "a", "and", "that". A list of stop words

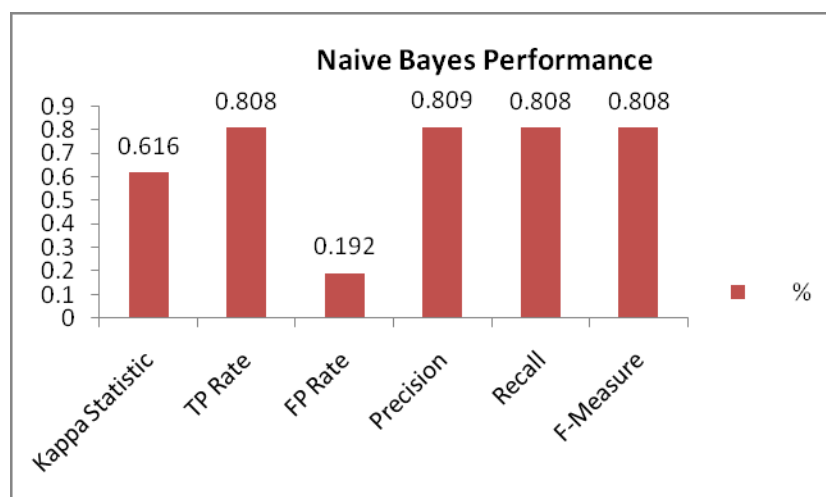
is extracted and after scanning word by word, these words are removed since stop words does not give any information. Only the extracted nouns except the stop words are used. Totally 1166 keywords are extracted from the 2000 instances after stop word removal process.

VI. RESULTS AND DISCUSSION

The review polarity dataset is applied to Support Vector Machine and Naïve Bayes algorithm for classification of reviews into positive and negative. Applying the algorithms individually produces a better performance. But the algorithm when combined for classification of same produces a better result than former. The tool used here is Weka which is popularly used data mining tool. Waikato Environment for Knowledge Analysis (WEKA) data mining software is used to perform the text mining classification process.

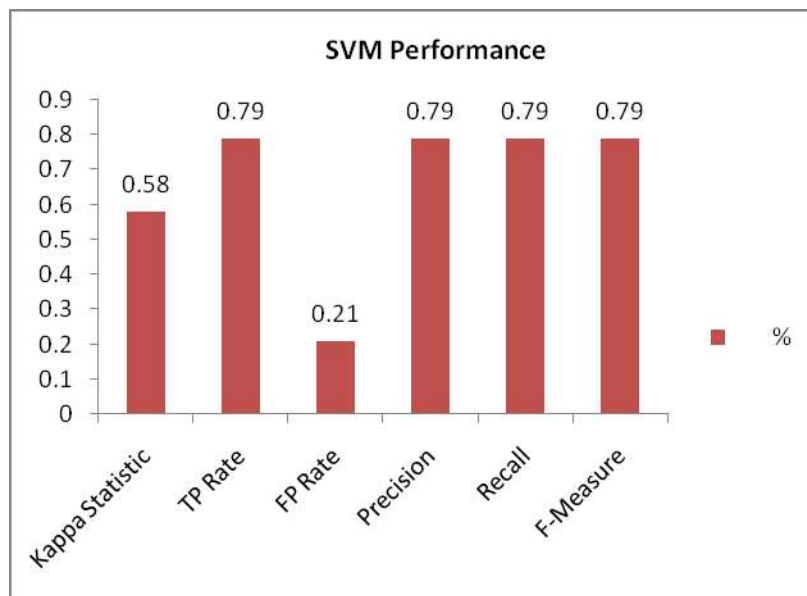
Naïve Bayes Classification

Applying Naïve Bayes algorithm alone for classification process, 1616 (80.8%) of instances are correctly classified and remaining 384 (19.2%) of instances are incorrectly classified out of 2000 instances. The percentage of correctly classified instances is often called accuracy or sample accuracy. So this data set consists of 80.8% accurate instances.



Support Vector Machine Classification

The support vector machine algorithm is applied to classify the instances. In this classification, 1580 (79%) of instances are correctly classified and remaining 420 (21%) of instances are incorrectly classified out of total 2000 instances. This data set consists of 79% accurate instances when SVM classification is used.



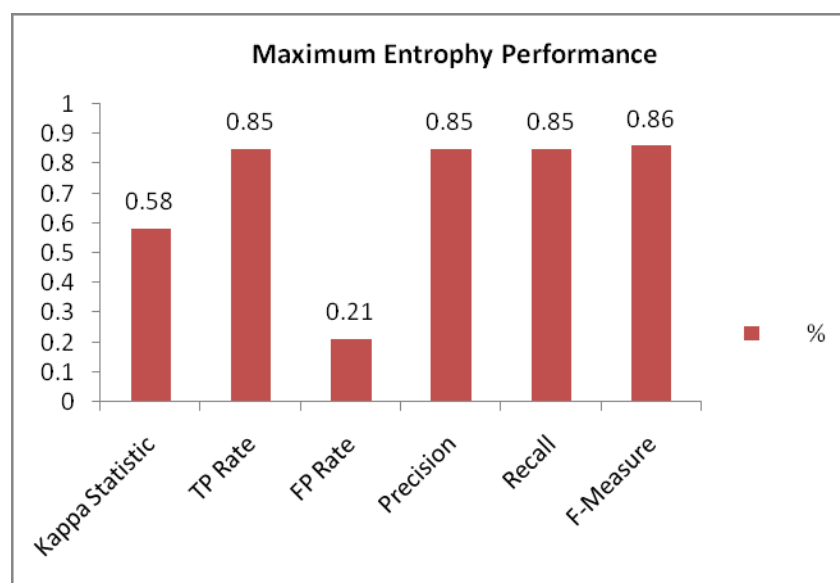
Logistic Regression

The logistic regression technique is applied to classify the instances. In this classification, 1756 (87%) of instances are correctly classified and remaining 244 (12%) of instances are incorrectly classified out of total 2000 instances. This data set consists of 87% accurate instances when Logistic Regression classification is used.



Maximum Entropy:

The Maximum Entropy technique is applied to classify the instances. In this classification, 1700 (85%) of instances are correctly classified and remaining 300 (15%) of instances are incorrectly classified out of total 2000 instances. This data set consists of 85% accurate instances when Maximum Entropy classification is used.



VII. CONCLUSION

As global Internet network grows rapidly, it is commonly used by a vast number of people to exchange information. By information we mean almost anything, from newspaper articles, to video streaming. One of quite new phenomena is an advent of social network websites, discussion boards (forums), price and product comparators and much more, where users can share their opinions in certain areas. This material of social networking sites is used by companies to check the reviews of their product, which further help them for its betterment.

The research is an analysis which uses different machine learning algorithms (NB, SVM, LR and ME) for the review polarity dataset individually which produces a classification accuracy of 80%, 79%, 87% and 85%. The proposed method shows that the Logistic Regression and Maximum Entropy has the highest accuracy. It uses meta algorithms for improving the performance and produced a classification accuracy of 87% and 85% which is better than others. The proposed method produced good classification accuracy for opinion mining for review polarity which consists of 2000 reviews. The method works better and provides a good performance.

VIII. REFERENCES

- [1] Aamera Z.H.Khan, Dr.Mohammed Atique, Dr.V.M.Thakare, "Sentiment Analysis using SVM", International Journal of Advanced Research in Computer Science and Software Engineering, Vol.5, Issue.04, April 2015. (ISSN: 2277 - 128X).

- [2] Ananta Arora, Chinmay Patil, Stevina Correia, "Opinion Mining: An Overview", International Journal of Advanced Research in Computer and Communication Engineering, Vol.4, Issue.11, November 2015. (ISSN: 2278-1021).
- [3] Asmita Dhokrat, Sunil Khilare, C. Namrata Mahender, "Review on Techniques and Tools used for Opinion Mining", International Journal of Computer Applications Technology and Research, Vol.4, Issue.6, 2015. (ISSN:2319 - 8656).
- [4] Bing Liu, "Sentiment Analysis and Opinion Mining", Morgan and Claypool Publishers, 2012.
- [5] Chetashri Bhadane, Hardi Dalal, Heenal Doshi, "Sentiment Analysis: Measuring Opinions", Science Direct, 2015.
- [6] Dola Saha, Prajna Paramita Ray, "Sentiment Analysis on Tweet Dataset using Datamining Techniques", International Journal of Advanced Research in Computer Science and Software Engineering, Vol.5, Issue.08, August 2015.(ISSN: 2277 - 128X).
- [7] A.G.Dongre, Sushmit Dharurkar, Swanand Nagarkar, Rahul Shukla, Vivek Pandita, "A Survey on Aspect based Opinion Mining from Product Reviews", International Journal of Innovative Research in Science, Engineering and Technology, Vol.5, Issue.2, February 2016. (ISSN: 2319 - 8753).
- [8] Gagandeep Singh, Kamaljeet Kaur Mangat, "Performance Analysis of Supervised Learning Methodologies for Sentiment Analysis of Tweets", International Journal of Advanced Research in Computer Science and Software Engineering, Vol.5, Issue.8, August 2015. (ISSN: 2277 - 128X).
- [9] Gaurangi Patel, Varsha Galande, Vedant Kekan, Kalpana Dange, "Sentiment Analysis using SVM", International Journal of Innovative Research in Computer and Communication Engineering, Vol.2, Issue 1, January 2014. (ISSN: 2320 - 9801).
- [10] Gautami Tripathi, Naganna.S, "Feature Selection and Classification approach for Sentiment Analysis", An International Journal, Vol.2, No.2, June 2015.
- [11] Kh airullah Khan, Baharum Baharudin, Aurnagzeb Khan, Ashraf Ullah, "Mining Opinion Components from Unstructured Reviews: A Review", Journal of King Saud University - Computer and Information Sciences, 26, 2014.
- [12] Norlela Samsudin, Mazidah Puteh, Abdul Razah Hamdan, Mind Zakree, Ahmad Nazri, "Immune Based Feature Selection for Opinion Mining", Proceedings of the World Congress on Engineering, Vol.3, July 3-5, 2013. (ISSN:2078 - 0966).

- [13] Poobana.S, Sashi Rekha.K, “Opinion Mining from Text Reviews using Machine Learning Algorithm”, International Journal of Innovative Research in Computer and Communication Engineering, Vol.3, Issue.3, March 2015. (ISSN: 2320 - 9801).
- [14] Pravesh kumàr Singh, Mohd Shahid Husain, “Methodological Study of Opinion Mining and Sentiment Analysis Techniques”, International Journal on Soft Computing, Vol.5, No.1, February 2014.
- [15] Preety, Sunny Dahiya, “Sentiment Analysis using SVM and Naïve Bayes Algorithm”, International Journal of Computer science and Mobile Computing, Vol.4, Issue 9, September 2015. (ISSN: 2320 - 088X) .
- [16] Nidhi R. Sharma, Prof. Vidya D Chitre, “Opinion Mining, Analysis and its Challenges”, International Journal of Innovations and Advancement in Computer Science, vol.3, Issue 1, April 2014. (ISSN: 2347 - 8616).
- [17] Reema Verma, Dr.Kiranjyoti, "Opinion Mining and Analysis of the Techniques for User Generated Content", International Journal of Advanced Research in Computer Science and Software Engineering, Vol.5, Issue 5, May 2015. (ISSN:2277 - 128X).
- [18] Richa Sharma, Shweta Nigam, Rekha Jain, “Mining of Product Reviews at Aspect Level”, International Journal in Foundations of Computer Science & Technology, Vol.4, No.3, May 2014.
- [19] Rushabh Shah, Bhoomit Patel, “Procedure of Opinion Mining and Sentiment Analysis: A Study ”, International Journal of Current Engineering and Technology, Vol.4, No.6, December 2014.
- [20] Md.Safdar, Dr.Md.Jawed Ikbal Khan, Md. Daiyan, “Mining Explicit Features for Opinion mining of Customer Reviews”, International Journal of Emerging Technology and Advanced Engineering, Vol.5, Issue 3, March 2015. (ISSN: 2250 - 2459).
- [21] G. Sneha, CT.Vidhya, "Algorithms for Opinion Mining and Sentiment Analysis: An Overview", "International Journal of Advanced Research in Computer Science and Software Engineering, Vol.6, Issue.02, February 2016. (ISSN: 2277 - 128X)."
- [22] K. Uma Maheswari, S.P. Raja Mohana, G. Aishwarya Lakshmi, “Opinion Mining using Hybrid Methods”, International Journal of Computer Applications, 2015. (ISSN: 0975 - 8887).
- [23] Vaishali Mehta, Prof. Ritesh K Shah, “Approaches of Opinion Mining and Performance Analysis: A Survey”, International Journal of Advanced Research in Computer Science and Software Engineering, Vol.4, Issue.10, October 2014. (ISSN: 2277 - 128X).

[24] Vidisha M. Pradhan, Jay Vala, Prem Balani, “A Survey on Sentiment Analysis Algorithms for Opinion Mining”, International Journal of Computer Applications, Vol.133, No.9, January 2016. (ISSN: 0975 - 8887).

[25] Walaa Medhat, Ahmed Hassan, Hoda Korashy, “Sentiment Analysis: A Survey”, Ain shams Engineering Journal, 5, 2014. (1093 - 1113).

NETWORK SECURITY AND CRYPTOGRAPHY*Ms.S.Brindha¹ , Ms.V.Tamilselvi²**Dept of Computer Science**Marudhar Kesari Jain College for Women,Tamil Nadu, India.*

ABSTRACT

Network security and cryptography is a concept to protect out network and data transmission over wireless networks. Data security is the main aspect of secure data transmission over unreliable network. The rapid development in information technology, the secure transmission of confidential data herewith gets herewith gets a great deal of attention. The generic name for the collection of tools to protect data and to thwart hackers is “computer security”

INTRODUCTION:

The **Network security** that consists of the provisions and policies adopted by a network administrator to prevent and monitor un-authorized access, modification, misuse of a computer network .

An art of using math to encrypt and decrypt the data is known as cryptography. One can save confidential information or transfer it can through various insecure networks that no one can views it. **Cryptography** is a part of Network security.

CRYPTOGRAPHY & NETWORK SECURITY:

It makes use of two types of keys one is public key and another one is private key.

Public key is not highly confidential which guarantees that specific key is known to both the parties.

For the purpose of signing messages or private property to a user cryptography makes use of digital signature

Uses of Public key encryption in Network security that are digital signature , key exchange , encryption or decryption.

Asymmetric-key cryptography is the other name for public –key that can be distinguished from symmetric – key cryptography. One must use algorithm of RSA for asymmetric-key cryptography.

Path, which links up strong encryption with an emergency decryption capability based network security project titles.



The terms are: BENEFITS OF CRYPTO SYSTEM IN NETWORK SECURITY:

Benefits of crypto systems in Network Security:

Under symmetric crypto system if chance's for data to be intercepted the encrypted data can be moved on the link. One can use escrowed encryption for a strong cryptography without following crypto anarchy

It is not a necessary for exchanging of keys in asymmetric or public key, which avoids key distribution problem.

To approve the receivers identity password authentication is used in symmetric cryptosystem.

CRYPTOGRAPHY:

Data security and cryptography are critical aspects of conventional computing. Here we provide basic terminology used in cryptography. In data and telecommunication, cryptography is necessary when communication over any un-trusted medium, which just about any network, particularly the internet.

plaintext: original message is known as plaintext.

ciphertext: coded message is called ciphertext.

Encryption: The process of converting from plaintext to ciphertext.

Decryption: The process of restoring from the ciphertext.

Cryptography is the science and art of transforming message to make them secure and immune to attack.

The two main types of cryptographic schemes:

Public key cryptography or symmetric encryption

Security key cryptography or symmetric encryption

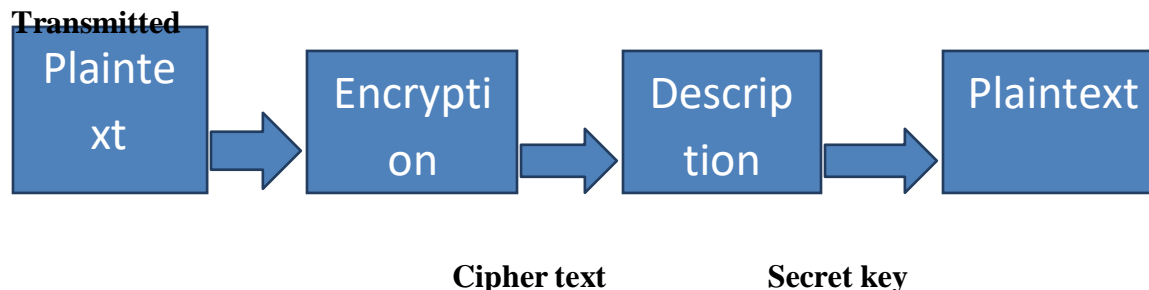
A asymmetric encryption scheme has six ingredients:

plaintext: This is the message or data that is fed into the algorithm as input.

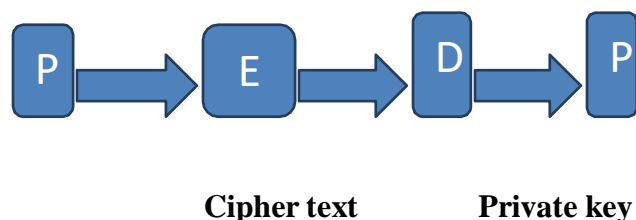
Encryption algorithm: The encryption algorithm performs various transformations on the plaintext.

Symmetric Encryption:

It is also referred as conventional encryption. Symmetric encryption is a form of cryptosystem in which encryption are performed using the same key. In this case an end user on a network, public or private, has a pair of keys one for encryption and one for decryption. The private key cannot be derived from the public key.

**Asymmetric Encryption:**

The asymmetrical cryptography method has been proven to be secure against computationally limited intruders. The security is a mathematical definition based upon the application of said encryption. It is referred as public key encryption. Asymmetric encryption is a form of cryptosystem in which encryption and decryption are performed using the different keys. This encryption can be used for confidentiality, authentication or both.

**GENERATING PUBLIC AND PRIVATE KEYS:**

- Pick two prime numbers, we'll pick $p=3$ and $q=11$.
- Calculate $n=p*q=3*11=33$. calculate $z=(p-1)*(q-1)=(3-1)*(11-1)=20$.
- Choose a prime number k , such that k is co-prime to z , i.e, z is not divisible by k . we have several choices for k : 7, 11, 13, 17, 19. Let's pick $k=7$ so, the number $n=33$ and $k=7$ become the server's public key.

- Now still done in advance of any transmission, the sender has to calculate its secret key. Here is how.
- $K*j=1(\text{mod } z)$
- $7*j=1(\text{mod } 20)$
- 21/20 gives "something" with the remainder of 1. so, $7*j=21$, and $j=3$. this is our secret key.

ENCRYPTION ALGORITHM:

- **Symmetric:**
 - Same key for encryption and decryption.
 - Key distribution problem.
 - It is also called as secret key cryptography.
 - Single key used for both encrypt & decrypt.
 - Key must be known to both the parties.
- **Asymmetric:**
 - key pairs for encryption and decryption.
 - Public and private keys.
 - Private keys are used for decrypting.
 - Public keys are used for encrypting.

WEB SECURITY:

- Web now widely used by business, government, individuals.
- But internet & web are vulnerable
- Have a variety of threats
 - Integrity
 - Confidentiality
 - Denial of service
 - Authentication

Need added security mechanisms

NETWORK SECURITY:

- To protect vital information while still allowing access to those who need it. Trade secrets, medical records, etc.
- Provide authentication and access control for resources. AFS.
- Guarantee availability of resources. **Ex:** 99's (99.999% reliability).

- **Security objects:** IDENTIFICATION, AUTHENTICATION, ACCESS CONTROL.
- **SSL(SECURE SOCKET LAYER):** transport layer security service, original developed by Netscape, used TCP to provide a reliable end-to-end service.
- **Wireless security:** 802.11 security, BLUETOOTH security, WAP2.0 security.
- **10 WAYS security:** back up early or often, use file-level and share-level security, password-protect documents, use EFS Encryption, use Disk Encryption, Make use of a public key infrastructure, Protect data in transit with IP security, secure wireless transmissions, use right management to retain control.

CONCLUSION AND FUTURE WORK

That the explosive growth in the Internet, network and data security have become an inevitable concern of any organization whom internal private network that connected to the Internet. A security for the data has become highly important. The User's data privacy is a central question over cloud. In more mathematical tools, cryptographic schemes are getting more versatile and often involve multiple keys for a single application. In those future, work can be done by key distribution and management in well optimal cryptography algorithm of data security over the clouds.

REFERENCE:

1. Shyam Nandhan kumar, "technique for security of multimedia using neural network," paper id-IJRETM-2014-02-05-020, IJRETM vol:02, issue:05, pp.1-7. sep-2014
2. Bellare, mihir; rogaway, phillip (21 sep 2005). "introduction". introduction to modern cryptography. p .10.
3. Menezes, A. J.; van Oorschot, P. C.; Vanstone, S. "A. Handbook of Applied Cryptography". ISBN 0-8493-8523-7.
4. Davis, R., "The Data Encryption Standard in Perspective," Proceeding of Communication Society magazine, IEEE, Volume 16 No 6, pp. 5-6, Nov. 1978.

A REVIEW ON SECURITY AND PRIVACY ISSUES IN BIG DATA ANALYTICS

Ms.J.Shivani¹

Department of Computer Applications

Marudhar Kesari Jain College for Women,Tamil Nadu, India

ABSTRACT:

Today the Big Data Analytics plays a major role in the world of Information Technology. Securing the valuable data from the interlopers, viruses and worms was a challenge for the past several decades. so many researchers developed few methods and technologies to protect their data but all the methods and technologies were only applicable for the structured data but now in this digital world we require some new security and privacy in structured, semi-structured and unstructured data (Big Data).In this review we will study about security and privacy measures that was proposed by different researchers and analyze the merits and demerits of those methods.

Key words: Big Data, Data Analytics, Security and Privacy issues.

INTRODUCTION:

Big Data has been arising a growing interest in both scientific and industrial field for its potential values . Before employing big data technology into a massive application a basic or a principle topic should be investigated: security and privacy. Big data is stored in MYSQL , SQL and in many more applications like in old days. It is more fast and useful than in previous language .The term “Big Data” is related with manipulating high amount of data exist in digitalized form that is collected by various companies and organizations . As day-to-day data’s are collected from various platforms and applications, social media’s ,network and other sources that big data is emerging. According to IDC(International Data Corporation) report prediction, the global data volume was predicted exponentially from 4.4 zettabytes to 44 zettabytes between 2013 and 2020 . It has be increased massively compared to the amount of data in 2013 . Big data is a collection of massive and complex data sets and data volume that include the huge quantities of data , data management capabilities, social media analytics and real time data. Big data is about data volume and large data sets measured in terms of terabytes or petabytes . This phenomenon is called Big Data. “Big data” is computing and storage power of the market in the era of 1999 one giga byte (1 gb) was considered big data. Today ,it may consist of petabytes(1,024 terabytes) or exabytes(1,024 petabytes) of information, including billions or even trillions of records from millions of people.

Big data handles a petabyte of data or more .It has distributed many data storage. Can leverage parallel task processing, provide data processing (Mapreduce or equivalent) capabilities and has extremely fast data insertion. Has central management and orchestration. Is hardware agnostic. Is extensible where its basic capabilities can be augmented and alerted. Nothing is perfect each and every thing have their own merit and demerit(pros and cons)so big data also have their own. Some of them are given below with their possible solutions.

- a) Storage issues.
- b) Security
- c) Processing issues in big data
- d) Privacy in big data
- e) Redundancy

The analysis of Big data involves many distinct phases which include:

Each of these phases introduces privacy challenges of big data. There are four different aspects of big data security:

Infrastructure security, Data acquisition and recording, Information extraction and cleaning, Data integration, Aggregation and representation, Query processing, Data modeling and analysis, Interpretation, Data privacy, Data management, Integrity and reactive security.

FIVE V'S OF BIG DATA:

- A. Variety-structured , unstructured, multi-factor , probabilistic.
- B. Volume-terabytes , recurs/arch, transactions, tables, files.
- C. Velocity-batch , real/near-time , processes, stream.
- D. Veracity- Truth worthiness, authenticity, origin ,reputations, availability ,accountability.
- E. Value/complexity- statistical , events, correlations , hypothetical.

BIGDATA ANALYTICS TOOLS AND METHODS

With the evolution of technology and the increased multitudes of data flowing in and out of organizations daily , there has become a need for faster and more efficient ways of analyzing such data .Having piles of data on hand is no longer enough to make efficient decisions at the right time. such data sets can no longer be easily analyzed with traditional data management and analysis technique and infrastructures. Therefore ,there arises a need for new tools and methods specialized for big data analytics ,as well as the required architectures for

storing and managing such data. Accordingly, the emergence of big data has an effect on everything from the data itself and its collection, to the processing, to the final extracted decisions. Consequently, proposed the big –data, analysis, and decisions(B-DAD)framework which incorporates the big data analytics tools and methods into the decision making process. the framework maps the different big data storage, management, and processing tools, analytics tools and methods, and visualization and evaluation tools to the different phases of the decision making. Hence, the changes associated with the big data analyses which can be applied for knowledge discovery and informed decision making. each area will be further discussed in this section. however, since big data is still evolving as an important field of research, and new findings and tools are constantly developing, this section is not exhaustive of all possibilities, and focuses on providing a general idea, rather than a list of all potentials opportunities and technologies.

SECURITY WITH PRIVACY

The privacy enhancing techniques have been proposed the last fifteen years, ranging from the cryptographic techniques such as oblivious data structure that hide data access patterns of data anonymization techniques that transform the data to make more difficult link specific data records to specific individuals. the privacy-enhancing symposium(PET)series and journals, such as transaction on data privacy however many such techniques either do not scale to very large data sets and/or do not specifically address the problem of reconciling security with privacy. at the same time, there are a few approaches that focus on efficiency of reconciling security with privacy and they discuss them in what follows.

- f) Privacy –preserving data matching.
- g) Privacy- preserving collaborative data mining.
- h) Privacy-preserving biometric authentication.

The computational, storage and communication costs of given protocols need to be considered. these costs could be especially significant for privacy-preserving protocols that involve cryptography. Given these three dimensions, one can imagine a multi-objective framework where different dimensions could be emphasized:

- i) Maximize utility, given risks and costs constraints.
- j) Minimize privacy risks, given the utility cost constraints.
- k) Minimize cost, given the utility and risk constraints.

Comprehensive solutions to the problem of security with privacy for big data require many research challenges and multidisciplinary approaches . they outline significant directions in what follows:

Data confidentiality: several data confidentiality techniques and mechanism exist – the most notable being access control systems and encryptions . both techniques have been widely investigated .however , for access control systems for big data we need approaches for:

- l)** Merging large number of access control policies
- m)** Automatically administering authorizations for big data and in particular for granting permissions
- n)** Enforcing access control policies on heterogeneous multi-media data
- o)** Enforcing access control policies in big data stores
- p)** Automatically designing , evolving and managing access control policies

ALGORITHM/TECHNIQUES

The current state of the art research in big data includes designing network topology, distributed algorithms , integration of software defined networks (SDN), scheduling, optimizations and load balancing among different commodity computers.

Data –driven information security dates back to bank fraud detection and anomaly –based intrusion detection systems. fraud detection is one of the most visible uses for big data analytics. Credit card companies have conducted fraud detection for decades. however, the custom –built infrastructure to mine big data for fraud detection was not economical to adapt for other fraud detection uses.

- q)** Application software security
- r)** Maintenance, monitoring, and analysis of audit logs.
- s)** Secure configurations for hardware and software
- t)** Account monitoring and control

Privacy-preserving techniques, including privacy-preserving aggregation ,operations over a encrypted data , and de-identification techniques.

Linear regression algorithm is a machine learning algorithm based on supervised learning . regression analysis is used for three types of application: finding out the effect of input variables on target variable. Finding out the change in target variable with respect to one or more input variable. It is the supervised machine learning in which the model finds the best fit linear line between the independent and dependent variable. Finding relationship between target and one or more predictors.

PRIVACY RISKS

While people experience the convenience of the big data they also feel some inconvenience. If the big data is not well protected for the user in the process of use , it will directly threaten the privacy of users and the security of the data. According to different protection content, it can be divided into anonymous identifiers, anonymous protection and privacy protection. In the era of big data , people's data security problems are not only the traditional issues of personal privacy, but more based on the analysis and research of peoples data and the targeted prediction of people's state and behavior . for example retailers can compare parents are more aware of their children's spending habits etc...,and thus post relevant advertising information .another example is the status of content published by users on the internet, and can analyze this person's political information ,like the team and spending habits. At present , many companies believe that after the information is processed anonymously ,the identifiers will be hidden , and then the information will be released . however ,the reality is that the protection of privacy cannot be effectively achieved only through anonymous protection . for example , a company may use some of its search history records in an anonymous manner within 3 months for use by people. Although the identification information contained there in has been carefully handled ,the contents of many of the records contained therein can be accurately defined .positioning . at present , china still lacks rules and regulations for users information management under the era of big data , and it does not have a good supervision system.this coupled with the lack of self-protection awareness among users has caused many losses caused by information leakage.

ADDRESSING OF BIG DATA SECURITY THREATS

Security tools for big data are not new. They simply have more scalability and the ability to secure many data types. The list below explains common security techniques for big data.

ENCRYPTION

Big data encryption tools need to secure data-at-rest and in-transit across large data volumes. Companies also need to encrypt both user and machine –generated data . as a result, encryption tools have to operate an multiple big data storage formats like No SQL databases and distributed file systems like HADOOP.

USER ACCESS CONTROL:

User access control is a basic network security tool. The lack of proper access control measures can be disastrous for big data systems. A robust user control policy has to be based on automated role-based settings and policies. Policy-driven access control protects big data platforms against insider threats by automatically managing complex user control levels, like multiple administrator settings.

INTRUSION DETECTION AND PREVENTION:

The distributed architecture of big data is a plus for intrusion attempts. An intrusion prevention system (IPS) enables security teams to protect big data platforms from vulnerability exploits by examining network traffic. The IPS often sits directly behind the firewall and isolates the intrusion before it does actual damage.

CENTRALIZED KEY MANAGEMENT:

Key management is the process of protecting cryptographic keys from loss or misuse. Centralized key management offers more efficiency as opposed to distributed or application-specific management. Centralized management systems use a single point to secure keys and audit logs and policies. A reliable key management system is essential for companies handling sensitive information.

BIG DATA PRIVACY PROTECTION TECHNOLOGY IS LACKING:

In the span of big data, information is disseminated at an extremely fast pace. At the same time as the transmission of information, due to the weak supervision of data information, lack of technical support, imperfect supervision system, and the vulnerability of information loss, the use of information is not high value and the data is reduced. The value of itself will bring about many negative and negative effects on individuals, business and even the society, resulting in great economic losses.

BIG DATA ANALYTICS: THE BENEFITS:

The benefits of using big data analytics include:

Quickly analyzing large amounts of data from different sources, in many different formats and types. Rapidly making better-informed decisions for effective strategizing, which can benefit and improve the supply chain, operations and other areas of strategic decision

making. Cost savings, which can results from new business process efficiencies and optimizations. A better understanding of customer needs , behavior and sentiment , which can lead to better marketing insights , as well as provide information for product development . improved , better informed risk management strategies that draw from large sample sizes of data . structured and unstructured data can be analysed using big data analytics.

METHODOLOGIES OF BIG DATA ANALYTICS:

‘Big Data’ is the application of specialized techniques and technologies to process very large sets of data . these sets are often so large and complex that it becomes difficult to process using on-hand management tools.

Data analysis techniques: A/B testing, Data fusion and data integration, Data mining ,Machine learning ,Natural language processing(NLP), Statistics, Data collection ,Data storage, Data analyzing, Knowledge creation.

CONCLUSION:

The study of various methodologies by many researchers are making the data secured and provide privacy which makes clear about the various methods , and its merits and demerits and its inabilities for providing security and privacy in big data . by this we conclude that we require some new technologies to strengthen the security and privacy for the data and build the security modifications in the available technologies.

REFERENCE:

1. P. Nandhini “ A Research on Big Data Analytics Security and Privacy in Cloud , Data Mining , Hadoop and Mapreduce “International Journal of Engineering and Research and Applications (IJERA), vol . 8, no. 4, 2018, pp 65-78.
2. Dongpo zhang “Big Data Security and Privacy Protection “ 8th International Conference on Management and Computer Science(ICMCS-2018) Advance in Computer Science Research , volume-77.
3. Fan Yan “Big Data Security and Privacy Protection[j]. Electronic Technology and Software Engineering , 2016(1):227
4. 2014.Bhavani Thuraisingham, “Big Data –Security with Privacy “ , NSF Workshop, September 16-17 ,

A STUDY ON MACHINE LEARNING FOR WEB VULNERABILITY DETECTION: THE CASE OF CROSS-SITE REQUEST FORGERY

*Ms.M Pragathi¹, Ms.K.Suganya²,
Assistant Professor,
Dept of Computer Science
Marudhar Kesari Jain College for Women,,Tamil Nadu, India.*

ABSTRACT

In this article, we propose a strategy to leverage Machine Learning (ML) for the detection of web application vulnerabilities. Web applications are particularly challenging to analyse, thanks to their diversity and therefore the widespread adoption of custom programming practices. ML is thus very helpful for web application security: it can cash in of manually labeled data to bring the human understanding of the online application semantics into automated analysis tools. We use our methodology within the design of Mitch, the primary ML solution for the black-box detection of Cross-Site Request Forgery (CSRF) vulnerabilities. Mitch allowed us to spot 35 new CSRFs on 20 major websites and 3 new CSRFs on production software.

Index Terms—Machine learning, cross-site request forgery,

I.INTRODUCTION

Web applications are the most common interface to security sensitive data and functionality available nowadays. They are routinely used to file tax incomes, access the results of medical screenings, perform financial transactions, and share opinions with our circle of fri to mention a few popular use cases. On the downside, this suggests that web applications are appealing targets to malicious users (attackers) who are determined to force economic losses, unduly access confidential data or create embarrassment to their victims. Securing web applications is documented to be hard. There are several reasons for this, starting from the heterogeneity and complexity of the online platform to the adoption of undisciplined scripting languages offering dubious security guarantees and not amenable for static analysis. In such a setting, blackbox vulnerability detection methods are particularly popular. As against white-box techniques which require access to the online application ASCII text file , black-box methods operate at the extent of HTTP traffic, i.e., HTTP requests and responses. Though this limited perspective might miss important

insights, it's the key advantage of offering a language-agnostic vulnerability detection approach, which abstracts from the complexity of scripting languages and offers a consistent interface to the widest possible range of web applications. This sounds appealing, yet previous work showed that such an analysis is way from trivial. One of the most challenges there's the way to expose to automated tools a critical ingredient of effective vulnerability detection, i.e., an understanding of the online application semantics. Example: Cross-Site Request Forgery (CSRF) Cross-Site Request Forgery (CSRF) may be a well-known web attack that forces a user into submitting unwanted, attacker controlled HTTP requests towards a vulnerable web application in which she is currently authenticated. The key concept of CSRF is that the malicious requests are routed to the online application through the user's browser, hence they could be indistinguishable from intended benign requests which were actually authorized by the user.



II. VULNEARABILITIES FOUND IN WEB APPLICATION

In this section we'll discuss the varied vulnerabilities mostly independent of the underlying platform and which are commonly found in commercial sites albeit these vulnerabilities are publicly known for some time now. Commonly web servers, application servers, and web application environments are susceptible to following types of vulnerabilities.

UnvalidatedInput- Web application request inputs from user to work out the way to respond in accordance to supply web service. The user enters input values to web application request. Attackers may pass harmful information to the online application which tries to bypass the website's security mechanisms.

Broken Access Control- In a web applications the users are categorized in the different level of privileges. Access control determines how the online application allows access to functions to some users and not others, also called authorization. But attacker could also be access higher level of authority.

Broken Authentication and Sessions Management-Web application creates a session when the user logged in, which specifies the period of time that a unique user interacts with a web application. Using session maintains state by providing the client with a singular id. This id is stored in a cookie that is used between the user browser and web server. If this session's details are not protected correctly, an attacker can steal it and misuse it.

Improper Error Handling-In the web application when a mistake occurs under normal transactions and if not handled by a correct error message to the user then there are chances of getting clues on flaws about the online application to the attackers and should disturb the traditional user. **Parameter Modification** Parameter modification is the problem where the attackers do not fill the form but rather pass the parameters from the URL itself, bypassing the form validations. Therefore it may lead to ambiguous effects on the form data and the overall site data.

Insecure Configuration Management-The web server that hosts the web application consists of web configuration files and directories which should not be accessed or viewed by someone unauthorized. So must be protected against the attackers. **Cookie Modification** Cookie stores the information in the text format which is used for state management. The web application uses cookies; the server sends cookie and stores at the client browser. The browser then returns the cookie to the server subsequent time the page is requested. The attackers can easily connect with the server to modify the contents of the user's cookie. **Directory Traversal** Web server consists of directories where files, information, application functions are stored to provide the services which do not have access to users. But attackers obtain these unauthorized directories, by traversing the directory within the address area of the online browser and should misuse it.

Cross Site Request Forgery (CSRF)-Cross-Site Request Forgery (CSRF) is an attack that forces an user to execute unwanted actions on an internet application during which they're currently authenticated. With a touch help from social engineering (such as sending a link via email or chat), an attacker may trick the users of an internet application into executing actions of the attacker's choosing. If the victim may be a normal user, a successful CSRF attack can force the user to perform state-changing requests like transferring funds, changing their email address, then forth. If the victim is an administrative account, CSRF can compromise the whole web application.

III.CSRF ATTACK CLASSIFICATION

CSRF vulnerabilities are generally divided into two types: off-site and on-site. Vulnerabilities outside the CSRF site are essentially external data submission issues within the traditional sense. Usually, programmers will consider adding a watermark to some message or comment forms to prevent SPAM problems (here, SPAM can be simply understood as spam, spam, or malicious replies with off-site links), but sometimes to improve User experience may not impose any restrictions on some operations, so attackers can predict and set request parameters in advance, write scripts in off-site Web pages to forge file requests, or use them with automatically submitted forms to implement GET, a POST request, when the user clicks on a link to access an offsite Web page in the session state, the client is forced to initiate a request. The type of vulnerabilities in the CSRF site is to some extent caused by programmers abusing \$_REQUEST class variables. In some sensitive operations (such as changing passwords, adding users, etc.), the user was originally required to initiate a POST request from the form submission to pass parameters to the program, but due to the use of variables such as \$_REQUEST, the program not only supports receiving the parameters passed by the POST request but also Support receiving parameters passed by getting requests, which will create conditions for attackers to use CSRF attacks. Generally, an attacker only must put the anticipated request parameters during a picture link of a post or message on the location , and therefore the victim are going to be forced to initiate these requests after browsing such a page.

IV.LITERATURE SURVEY- i. State Of The Art:

Automated Black-Box Web Application Vulnerability Testing

AUTHORS: Jason Bau, Elie Bursztein, Divij Gupta, and John C. Mitchell Black-box web application vulnerability scanners are automated tools that probe web applications for security vulnerabilities. In order to assess the present state of the art, we obtained access to eight leading tools and administered a study of (i) the category of vulnerabilities tested by these scanners, (ii) their effectiveness against target vulnerabilities, and (iii) the relevance of the target vulnerabilities to vulnerabilities found in the wild. To conduct our study we used a custom web application susceptible to known and projected vulnerabilities, and former versions of widely used web applications containing known vulnerabilities. Our results show the promise and effectiveness of automated tools, as a gaggle , and also some limitations. In particular, "stored" sorts of Cross-Site Scripting (XSS) and SQL Injection (SQLI) vulnerabilities aren't currently found by many tools. Because our goal is to assess the potential of future research, not to

evaluate specific vendors, we do not report comparative data or make any recommendations about the purchase of specific tools.

ii. Mitch:-A Machine Learning Approach To The Blackbox Detection Of CsrF

Vulnerabilities

AUTHORS: Stefano Calzavara, Mauro Conti, Riccardo Focardi, Alvise Rabitti, and Gabriele Tolomei Cross-Site Request Forgery (CSRF) is one of the oldest and simplest attacks on the Web, yet it is still effective on many websites and it can lead to severe consequences, such as economic losses and account takeovers. Unfortunately, tools and techniques proposed so far to identify CSRF vulnerabilities either need manual reviewing by human experts or assume the availability of the source code of the web application. In this paper, we present Mitch, the primary machine learning solution for the black-box detection of CSRF vulnerabilities. At the core of Mitch, there's an automatic detector of sensitive HTTP requests, i.e., requests which require protection against CSRF for security reasons. We trained the detector using supervised learning techniques on a dataset of 5,828 HTTP requests collected on popular websites, which we make available to other security researchers. Our solution outperforms existing detection heuristics proposed within the literature, allowing us to spot 35 new CSRF vulnerabilities on 20 major websites and three previously undetected CSRF vulnerabilities on production software already analyzed using a state-of-the-art tool.

V.EXISTING SYSTEM:

In the existing system Securing web applications is well known to be hard. There are several reasons for this, starting from the heterogeneity and complexity of the online platform to the adoption of undisciplined scripting languages offering dubious security guarantees and not amenable for static analysis. Though this limited perspective might miss important insights, it's the key advantage of offering a language-agnostic vulnerability detection approach, which abstracts from the complexity of scripting languages and offers a consistent interface to the widest possible range of web applications.

i. DISADVANTAGES OF EXISTING SYSTEM:

In white-box techniques which require access to the web application source code.

Black-box methods operate at the extent of HTTP traffic, i.e., HTTP requests and responses.

Algorithm: Burp and ZAP tools

VI. PROPOSED SYSTEM:

Cross-Site Request Forgery (CSRF) is a well-known web attack that forces a user into submitting unwanted, attacker-controlled HTTP requests towards a vulnerable web application in which she is currently authenticated. The key concept of CSRF is that the malicious requests are routed to the online application through the user's browser, hence they could be indistinguishable from intended benign requests which were actually authorized by the user. The CSRF does not require the attacker to intercept or modify a user's requests and responses: it suffices that the victim visits the attacker's website, from which the attack is launched. Thus, CSRF vulnerabilities are exploitable by any malicious website on the Web.

i. ADVANTAGES OF PROPOSED SYSTEM:

The value of ordinary HTTP request headers like Referrer and Origin, indicating the page originating the request.

The presence of custom HTTP request headers like X-Requested-With, which cannot be set from a cross-site position.

The presence of unpredictable anti-CSRF tokens, set by the server into sensitive forms. Algorithm: RandomForestClassifier.

VII. EXPERIMENTAL EVALUATION

User:-The User can register first. While registering he required a valid user email and mobile for further communications. Once the user register then the admin can activate the customer. Once the admin activated the customer then the user can log in to our system. Users can do the data preprocessing. First required running website name. By using that website the user can test the Csrf. With the help of the bolt tool, the user can fetch related all Csrf and generated algorithm names. The result will be stored in JSON files. Later the user can get the results of Mitch dataset. The mitch dataset tested for POST method also GET method. The result will be displayed on the browser.

Admin:-Admin can login with his credentials. Once he login he can activate the users. The activated user only login in to our applications. The admin can set the training and testing data for the project of the Mitch Dataset. The user search all URLs related Csrftoken admin can view on his page. The admin can also check the POST method performed data from the dataset and GET method related data also.

False Positives and False Negatives:-Mitch produces a false positive when it returns a candidate CSRF that can't be actually exploited. This is something relatively easy to detect by manual testing, though this process is tedious and time-consuming. In general, it's impossible to reliably identify when Mitch produces a false negative, because this is able to require knowing all the CSRF vulnerabilities on the tested websites. To estimate this important aspect, we keep track of all the sensitive requests returned by the ML classifier embedded into Mitch, and that we focus our manual testing on those cases. This is a reasonable choice to make the analysis tractable because we first showed that the classifier performs well using standard validity measures.

Machine Learning Classifier:-The ML classifier used by Mitch was trained from a dataset of around 6, 000 HTTP requests from existing websites, collected and labeled by two human experts. The feature space X of the classifier has 49 dimensions, each one capturing a specific property of HTTP requests. Those can be organized into 3 categories: Structural, Textual, and Functional.

1) Structural: This category of features describes the structural properties of an HTTP request. More precisely, we define the following set of numerical features:

- numOfParams: the entire number of parameters;
- numOfBools: the amount of request parameters sure to a boolean value;
- numOfIds: the number of request parameters bound to an identifier, i.e., a hexadecimal string, whose usage was empirically observed to be common in our dataset;
- numOfBlobs: the amount of request parameters sure to a blob, i.e., any string which isn't an identifier;

VIII. CONCLUSION

Web applications are particularly challenging to analyze, due to their diversity and the widespread adoption of custom programming practices. ML is thus very helpful in the web setting because it can take advantage of manually labeled data to expose the human understanding of the web application semantics to automated analysis tools. We validated this claim by designing Mitch, the first ML solution for the blackbox detection of CSRF vulnerabilities, and by experimentally assessing its effectiveness. We hope other researchers might take advantage of our methodology for the detection of other classes of web application vulnerabilities.

IX. REFERENCES:

- [1] Stefano Calzavara, Riccardo Focardi, Marco Squarcina, and Mauro Tempesta. Surviving the web: A journey into web session security. *ACM Comput. Surv.*, 50(1):13:1–13:34, 2017.
- [2] Avinash Sudhodanan, Roberto Carbone, Luca Compagna, Nicolas Dolgin, Alessandro Armando, and Umberto Morelli. Large-scale analysis & detection of authentication cross-site request forgeries. In *2017 IEEE European Symposium on Security and Privacy, EuroS&P 2017, Paris, France, April 26-28, 2017*, pages 350–365, 2017.
- [3] Stefano Calzavara, Alvisio Rabitti, Alessio Ragazzo, and Michele Bugliesi. Testing for integrity flaws in web sessions. In *Computer Security - 24rd European Symposium on Research in Computer Security, ESORICS 2019, Luxembourg, Luxembourg, September 23-27, 2019*, pages 606– 624, 2019.
- [4] OWASP. OWASP Testing Guide. [https://www.owasp.org/index.php/ OWASP Testing Guide v4 Table of Contents](https://www.owasp.org/index.php/OWASP_Testing_Guide_v4_Table_of_Contents), 2016.
- [5] Jason Bau, Elie Bursztein, Divij Gupta, and John C. Mitchell. State of the art: Automated black-box web application vulnerability testing. In *31st IEEE Symposium on Security and Privacy, S&P 2010, 16- 19 May 2010, Berkeley/Oakland, California, USA*, pages 332–345, 2010.
- [6] Adam Doupé, Marco Cova, and Giovanni Vigna. Why johnny can't pentest: An analysis of black-box web vulnerability scanners. In *Detection of Intrusions and Malware, and Vulnerability Assessment, 7th International Conference, DIMVA 2010, Bonn, Germany, July8-9, 2010. Proceedings*, pages 111– 131, 2010.
- [7] Adam Barth, Collin Jackson, and John C. Mitchell. Robust defenses for cross-site requestforgery. In *Proceedings of the 2008 ACM Conference on Computer and CommunicationsSecurity, CCS 2008, Alexandria, Virginia, USA, October 27-31, 2008*, pages 75–88, 2008.

IMPRESSION OF COVID IN E-COMMERCE PLATFORM FOR THE DURATION OF EPIDEMIC

Mrs. Abarna Sri R¹.

*Asst. Professor Department of Computer Science
Sri GVG Visalakshi College for Women, Tamil Nadu, India*

Mrs. B.Sasikala².

*Asst Professor Department of Computer Science
Sri GVG Visalakshi College for Women, Tamil Nadu, India*

ABSTRACT

The COVID-19 crisis accelerated an expansion of e-commerce towards new firms, customers, and types of products. It has provided customers with access to a significant variety of products from the convenience and safety of their homes and has enabled firms to continue operation despite contact restrictions and other confinement measures. Some of these changes in the e-commerce landscape will likely be of a long-term nature, considering the possibility of new waves of the epidemic, the convenience of the new purchasing habits, learning costs and the incentive for firms to capitalise on investments in new sales channels. Governments also need to address the particular need of SMEs, including by ensuring a fair playing field in the context of intermediated services (e.g. online platforms). Ensuring sufficient competition in the retail sector and a well-functioning enabling environment for e-commerce, including communication services, logistics or trade, is also crucial. Suggestions for administrators and hypothesis are likewise talked about.

1. Introduction

Late many years have seen an acceleration of worldwide pandemics, e. g., SARS in 2003, H5N1 in 2006, and H1N1 in 2009, every one of which hampered business exercises and financial development (Chung, 2015). For model, the Covid-19 infection caused a 13.5% drop in China's industry creation and a 20.5% decrease in retail deals in the initial two months of 2020 while the U. S's financial exchange lost trillions of dollars, prompting a negative abundance impact and lower GDP (Pesek, 2020). As researchers have since a long time ago cautioned, irresistible illnesses can drive business exercises into another reality that seriously impacts activities and one where supervisors have hazy direction about how to viably react (Hudecheck et al., 2020). Because of the Covid-19 effect, for example, customers are progressively going to online buys; in this way, supervisors

should be inventive in looking for elective types of provisions which raise the interest in the help among firms and purchasers.

The effect of the COVID-19 crisis on e-commerce is not uniform across product categories or sellers. In the United States, for example, a surge in demand was observed for items related to personal protection (e.g. disposable gloves), home activities, groceries or ICT equipment, while demand dropped for items related to travel, sports or formal clothing (e.g. suitcases, bridal clothing, gym bags, etc.) (OECD, 2020[2]).³ Shifts towards e-commerce have been observed in several countries, in particular along the food supply chain, including farmers who started using digital technologies to sell their produce directly to consumers or restaurants that switched to providing food or grocery delivery services (OECD, 2020[3]). In Germany, online sales grew significantly for medicines and groceries, historically laggard sectors in terms of e-commerce, while overall online sales contracted by around 18 percent in March 2020 in comparison to the previous year (OECD, 2019[4]). In Korea, where official statistics are available, the e-commerce transaction value rose by 15.8% between July 2019 and July 2020. Significant increases were observed for food services (66.3%), household goods (48%) and food and beverages (46.7%), whereas online transactions involving culture and leisure services or travel arrangement and transportation services declined significantly, by 67.8% and 51.6% respectively. In China, food products were the single biggest winner in e-commerce, with an increase in accumulated sales from January to April 2020 of 36%, relative to the previous year. In contrast, total online sales over January to April 2020 remained almost constant compared to the same period in 2019 (+1.7%), after having grown significantly over 2018-19 (17.8%). Accumulated sales of clothing products contracted by 16% compared to 2019, after significant growth from 2018-19 (23.7%).

2. Conceptual foundation and speculation improvement

- Uses and delight hypothesis

Utilizations and satisfaction hypothesis (UGT) alludes to a persuasive socio-intelligent worldview that clarifies which social and brain science needs spur shoppers to choose specific stages (Li et al., 2018). Advances in Internet advances have empowered web based business stages to change retail and coordination tasks that make more monetary advantages as decreases of expenses and postponements, setting off the parts of UGT in clarifying use inspirations (Luo et al., 2011). For instance, buyers use web based business

generally assists with analysing which job purchaser conduct and inspiration plays in buyers' online associations with firms (Huang et al., 2014). In the event that organizations can successfully satisfy purchaser needs with items through online stages, shoppers will be more able to keep on cooperating with firms (Li et al., 2018).

The current examination contends that UGT assists with catching how buyers embrace and use online business stages to fulfil their buy needs. As far as ease of use and usefulness, UGT tends to the examples and inspirations of online stage's application in looking for data, communicating with substance and communities, and supporting buy choices for explicit circumstances (Korhan and Ersoy, 2016). For instance, the Covid-19 pandemic expanded buyers buy goal towards web based business stages because of the medium's apparent wellbeing and security benefits as opposed to those of customary retailers. UGT assists with clarifying why, considering pandemic dread, customers progressively occupied with online stages to make pertinent buy contemplations. Also, UGT empowers managers to distinguish the explanations for buyer decision and item and administration evaluations (Ray et al., 2019). UGT additionally furnishes supervisors with significance of separated substance system dependent on conditional information and data trade to foster substance methodology all the more successfully (Lim and Kumar, 2019). Hence, this investigation utilizes UGT to investigate the intuitive linkages of PEEP, customers' apparent financial advantages, pandemic dread, and reasonable utilization.

3. Economic benefits

In e-business writing, the ordinary importance of monetary advantage generally alludes to the buyer insight that internet business stages offer value limits, advancements, or other special activities (Liu et al., 2019). Liu et al. (2019) demonstrate that monetary advantage got from internet business stages can produce positive passionate reactions, which thus prompts online buy expectation. In accordance with this rationale, Wang and Herrando (2019) propose that internet business plat-structures improve cooperation's among vender and buyers and among purchasers, as these connections sway expanding web based shopping conduct and create financial advantages. In view of UGT, monetary advantages additionally trigger buyer goal to take part in feasible consumption during unsure circumstances. Advantages, for example, may incorporate coupons, cash-backs, and limits (Ray et al., 2019). UGT likewise proposes that if business contributions don't coordinate with purchasers' assumptions, the business remunerations relationship might

be set off by customers (Simon, 2017). Furthermore, this investigation adds that financial advantages relate to customers' apparent monetary advantages (e.g., cost reserve funds and limits) for buying from internet business stages during pandemic periods.

3.1. Pandemicfear

A pandemic insinuates another affliction that a large number individuals don't have an invulnerability to and one that spreads all throughout the planet (WHO, 2010), and dish demits are getting likely the most serious risk for the current reality (Harvard Global Health Institute, 2020). According to the Harvard Global Health Institute (2020), an overpowering affliction could rapidly cause a considerable number passing's around the world, destabilize governments, and limit trade and travel. The new Covid-19 disease, for instance, has defiled 422,945 people all throughout the planet, achieving 18,907 passing's at the hour of creating (March 25, 2020), and addressing 3.4% of death in assessment with intermittent flu (1.0%) (Worldometer, 2020). This fear of infection becomes an emerging issue, so much that it is essential to sort out how pandemic fear affects client spending and purchase lead (Khan and Huremovi'c, 2019).

The customer lead composing demonstrates that fear implies the deplorable aftereffects of a specific event that can incite changes in client direct and manner (Solomon, 2017). In such way, the Covid-19 pandemic has changed client purchasing conduct as buyers fear infection (Laato et al., 2020; Prentice et al., 2020). For example, a concise report by Nielsen (2020) shows that 45% of Vietnamese customers purchased things proposed for limit, and 25% purchased these things on the web, while Taiwanese purchasers purchased additional second noodles. Moreover, Vietnam firms have changed their assistance programs (e.g., home transport and covers and sanitizers at stable expenses), the outcome of which is a wonderful improvement in bargains (Vietnam News, 2020). All things considered, this assessment conceptualizes pandemic fear as customer infection fear, a conviction which impacts the way where purchasers utilize online business stages to purchase things.

4. Response efforts and recovery plans

Near portion of the study members said governments had not focused on the online business area adequately in their COVID-19 reaction endeavours and recuperation plans.

The pandemic has additionally uncovered holes in arrangement regions vital to improving computerized preparation in non-industrial nations, for example, frail web based business administrative structures and bottlenecks in financing advanced business people and new companies.

In any case, the studied organizations recognized that a few measures taken by the general population and private areas have helped lower obstacles for organizations and customers to utilize internet business administrations.

Expanded public mindfulness crusades on the advantages of web based business, more advanced abilities preparing openings, and decreased e-installment exchange costs were referred to as the most effective.

The report features a few fruitful drives:

In Cambodia, the public authority has passed an internet business law to facilitate the enrolment of web based business organizations.

In Kiribati, computerized arrangements carried out to battle the spread of COVID-19 have given new driving force to endeavours to bridge the advancement gains of ICT and online business.

In Myanmar, internet business and the advanced economy are included in the public authority's COVID-19 monetary alleviation plan

In Rwanda, the Central Bank suspended versatile cash expenses for 90 days and deferred charges on push and pull administrations between ledgers and portable wallets

In Senegal, joint effort pointed toward onboarding new organizations and supporting nearby produce suppliers has demonstrated fruitful (models incorporate the e-KomKoman and the online business Sénégal stages)

In Uganda, internet business stages have been quick to advance contactless instalments and coordination's and sustain new associations with improvement accomplices.

In Tunisia, the "mark de confiance" drive has assisted form with trusting across the online business environment

REFERENCES

- [1] McNulty, E.J., 2020. Leading through Covid-19. MIT Sloan Management Review. <https://sloanreview.mit.edu/article/leading-through-covid-19/>. accessed 23 March 2020.
- Nielsen, 2020a.
- [2] Nielsen Investigation: “Pandemic Pantries” Pressure Supply Chain amid Covid-19 Fears. <https://www.nielsen.com/us/en/insights/article/2020/nielsen-investigation-pandemic-pantries-pressure-supply-chain-amidst-covid-19-fears/>. accessed 22 March 2020.
- [3] Pesek, W., 2020. China’s Recession Will Hurt as Coronavirus Crisis Spreads to U.S. And Europe, Forbes. <https://www.forbes.com/sites/forbes-personal-shopper/2020/03/16/disney-amazon-prime-video-more-best-tv-shows-to-stream-during-your-time-home/#3336a7ee5882>. accessed 18 March 2020.
- [4] Plangger, K., Watson, R.T., 2015. Balancing customer privacy, secrets, and surveillance: insights and management. *Bus. Horiz.* 58 (6), 625–633.
- [5] Prentice, C., Chen, J., Stantic, B., 2020. Timed intervention in COVID-19 and panic buying. *J. Retailing Consum. Serv.* 57, 102203 <https://doi.org/10.1016/j.jretconser.2020.102203>.

SECURITY ANALYSIS AND PRESERVING BLOCK-LEVEL DATA PREVENTION IN CLOUD STORAGE SERVICES

*Ms.R.Sandhiya¹,Ms.P.Prabavathy²,
Asst Professor,PG Dept of Computer Applications,
MKJC,Tamil Nadu, India.*

ABSTRACT:

Cloud compute setting in which probabilistic querying of outsourced data is a examine provider. donor or anyone else. Outsourcing offers the data owner scalability and a low initial venture. The could do with for seclusion may be due to the data being sensitive or otherwise confidential. Dynamic audit service is constructed base on the techniques, splinterconfiguration, random sample, and index-hash table, supporting attestableupdate to outsourced data and timely anomaly detection. To propose a method based on probabilistic query and sporadic verification for improving the performance of audit services. Final Results not only corroborate the helpfulness of our approaches, but also show our audit system verifies the truth with lower working out overhead and requiring less extra storage for audit metadata.

Index Terms—Security, IHT ,Cloud storage, Certificate Authority, Audit service

1. INTRODUCTION

In present days, the up-and-coming cloud-computing model is rapidly in advance force as an unconventional to traditional information technology. Cloud computing make on hand a scalability environment for emergent quantity of data and processes that work on a variety of services and function by means of on-demand self-services. One original aspect of this model shifting is that data are being centralized and outsourced into clouds. This category of outsourced storage legroom services in exhaust have turn out to be The cloud storage service (CSS) mitigates the load of continuation and storage management. However, if such a significant service is fragile to attacks or failures, it would take permanent losses to user since their data or account are stored into an unsure cargo space pool outside the enterprises. Cloud infrastructures are much more convincing and reliable than personal computing devices. If they are still susceptible to security bullying both from inside and external the cloud for the benefits of their control, there exist various motivations for cloud observe providers (CSP) to behave falsely toward the cloud users in addition, the difference of opinion once in a blue moon suffers from the lack of trust on cloud service provider. As a cause, their behaviors may not be known by the cloud users. then, it is compulsory for cloud tune providers to offer a scalable audit service to check the integrity and availability of the stored data. While Cloud total makes these advantages more appealing than ever, it also brings new challenging security threats towards users' outsourced data. Since cloud service provider is separate clerical units, data outsourcing is actually resign user's control over the destiny of their data. The exactness of the data in the cloud is being put at risk due to the subsequent reasons. First of all, even if the infrastructures beneath the cloud are much more dominant and reliable than private compute devices, they are silent facing the broad range of both internal and external threats for data integrity.

Traditional cryptographic technologies for data integrity and accessibility, based hash functions and on autographscheme cannot work on the outsourced data lacking a local facsimile of data. In accumulation, it is not a realistic solution for data validation by downloading them due to the limited transaction, especially for large-size files. Moreover, the solutions to audit the correctness of the data in a cloud environment can be formidable and expensive for the darken users. Therefore, it is critical to recognize public audit facility for Cloud Storage Service, so that data owners may remedy to a third party auditor (TPA), who has aptitude and capabilities that a common client does

not have, for from time to time auditing the outsourced data. This examination service is extensively insignificant for digital forensics and data assurance in clouds.

2. AUDIT SYSTEM ARCHITECTURE

The audit system architecture for outsourced data in clouds in which can work in an assessment service outsourcing approach. In this design, we reflect happening a data cargo space service containing four entities:

- A. **Data owner (DO):** who has facts files to be stored in the darken and relies on the cloud for data maintenance, can be an individual customer or an organization.
- B. **Cloud Storage Service Provider (CSP):** who provides data storage service and has enough luggage booth space to maintain client's data.
- C. **Third Party Auditor (TPA):** a trusted person who supervise or examine outsourced data under request of the numbers owner.
- D. **Authorized Application (AA):** who have the right to entrance and use stored data.

The data which the data owner wants to store in cloud first reaches the allowed application which will create digital cross and sends the data to the cloud cargo space. If the user needs to verify data means the proof request should be send to third party auditor (TPA), the TPA will retrieve the digital signature from the database and will send the substantiation request to the management server. The managing server in turn will generate the digital cross for the data stored in the cloud and it will send only that digital signature instead of the whole data to the TPA. The TPA will decrypt the digital signature and compares the message digest for verifying correctness of data.

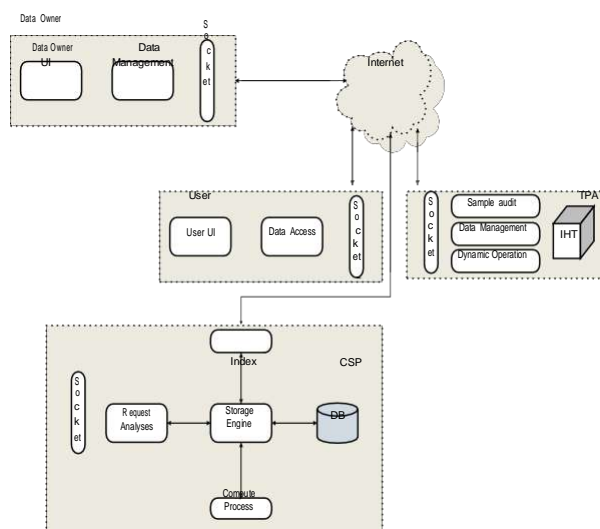


Fig.: System Architecture

This design is known as the audit service outsourcing due to data veracity authentication can be implemented by Third Party Auditor without help of data owner. Architecture contains the data owner and granted regulars need to enthusiastically interact with cloud service provider to access or update their data for various purpose purposes. However, we neither assume that cloud service bringer is trust to guarantee the security of stored data, or suppose that the data owner has the capability to collect the verifications of cloud service provider's fault after errors occur. Hence, third party auditor, as a trust third party (TTP), is used to ensure the storage safety measures of their outsourced data. We assume the third party surveyor is reliable and self-determining, and thus has no encouragement to join in concert with either the cloud service provider or the clients during the auditing development:

- a) TPA must be able to make customary check on the integrity along with ease of use of these delegated data at appropriate intervals;
- b) TPA must be able to take the evidence for the dispute about the contradiction of numbers in terms of authentic records for all data operations.

To facilitate privacy-preserving public audit for cloud data storage beneath the design, the protocol design ought to get subsequent defense and concert guarantee:

- A. Audit-without-downloading: to allow TPA (or other trade with the help of TPA) to authenticate the correctness of cloud data on stipulate without improving a copy of intact data or bring in other on-line burden to the cloud users;
- B. Verification-correctness: to formulate sure there exists no unethical CSP with the aim of can pass the inventory from TPA without indeed storing users' data intact;
- C. Privacy-preserving: to make sure that there exist no way for TPA to derive users' data from the in sequence collected for the period of the auditing process;
- D. High-performance: to allow TPA to perform auditing with least overheads in cargo space, communication and computation, and to maintain arithmetic audit sampling and optimized assessment schedule with a long adequate period of time.

The above process involves some procedures: TagGen, KeyGen, Update, Insert, Delete, algorithms as well as an interactive proof set of rules of retrievability. In regulate to recover security and show, we make use of following techniques to construct related algorithms and protocol.

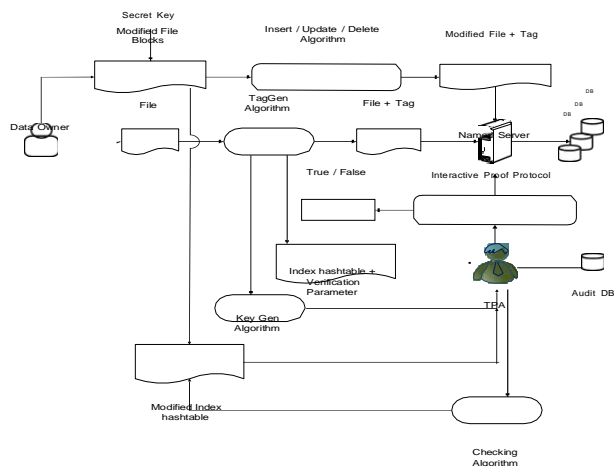


Fig.:DataFlowDiagram3.

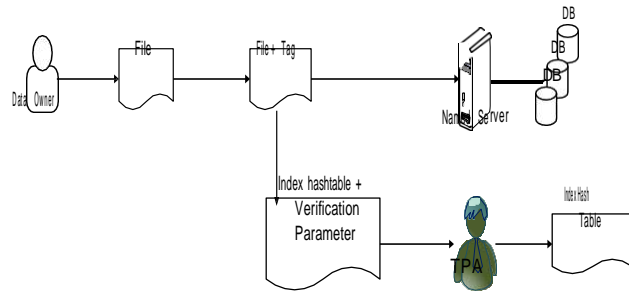
Methodology&Implementation

In this slice, we propose a cryptographic interactive audit idea to support our audit system in clouds. This scheme is construct on the standard model of interactive substantiation system, which can ensure the discretion of secret data (zero-knowledge property) and un decidability of worthless tags (soundness property).

To authorize the effectiveness of our approach, we have implement a prototype of an audit system based onour proposed solution. This classification have been developed in an investigational cloud computing system atmosphere, which is constructed within the support of the IaaS to provide powerful virtualization, scattered storage, and automated management

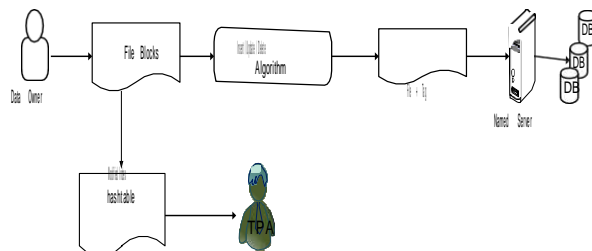
Index-Hash Table creation :

To support dynamic data operations, we begin auncomplicated IHT to record the changes of file blocks, as wellas engender the hash value of each block in the verificationprocess. The formation of our IHT is similar to that of fileslab allocation table in file systems. Generally, the IHT consists ofserial amount, block amount, version number,and random digit.



Vibrant Data Operations Module:

Dynamic data operations are available only to DOs or AAs, who cling to the secret key sk . Here, all operations are based on data blocks. In addition, to execute audit services, applications need to update the IHTs. It is compulsory for TPA and CSP to test the validity of modernized data.



5. CONCLUSION

Cloud compute releases the world of computing to a wider range of uses and increases the ease of usage by giving access through any kind of internet connection. Though with these increased ease of usage also come drawbacks. Privacy security is a key issue for cloud storage and is to be considered very imperative. To ensure that the risks of privacy have been mitigated a variety of techniques that may be used in order to achieve privacy. This paper has addressed some privacy approaches for overcome the issues in privacy on un trusted data stores in cloud compute. Categories the methodologies in the literature as encryption based method, access have power over based mechanisms, query integrity/ keyword hunt schemes, and audit ability schemes. The work is giving an efficient privacy-preserving storage compare to other works. Even though near are many approach in the prose for mitigating the concerns in privacy, no approach is fully urbane to give a privacy-preserving storagethat overcomes all the other isolation concerns. Thus to dealwith

the concerns of privacy, we need to develop privacy-preserving framework that overcomes the uncertainties in privacy security and egg on users to adopt cloud storage services more assertively.

6. REFERENCES

- 1) Yan Zhu, Hongxin Hue, Gail-Joon Ahn, Stephen S. Yau, "Efficient audit service outsourcing for data reliability in exhaust", Elsevier Journal Of Systems and Software, vol.85, pp.1083-1095, 2017.
 - 2) Bhagyaraj Gowrigolla, Sathyalakshmi Sivaji, M. Roberts Masillamani "Design and Auditing of Cloud Computing Security", IEEE intercontinental Conference on Information and Automation for Sustainability, 2016.
 - 3) Cong Wang furthermore Kui Ren, Wenjing Lou, Jin Li "Toward Publicly Auditable Secure Cloud Data Storage Services", IEEE International Journal On Networks, 2015.
 - 4) Q. Wang, C. Wang, K. Ren, W. Lou and J. Li, "Enabling public auditability and data dynamics for storage safety measures in cloud computing", IEEE Transactions on Parallel with Distributed Systems, vol. 22, no.5, pp.847-459, 2013.
 - 5) G. Ateniese, R. Burns, R. Curtola, J. Herring, L. Kissner and D. Song, "attestable figures possession at untrusted supplies", in Proc. Of CCS'07, pp.598-609.
 - 6) Daniel J. Abadi "Data supervision in the Ling Li, Lin Xu, Jing Li and Changchun Zhang "Study on the Thirdparty Audit in shade StorageService" IEEE International Conference on Cloud & Service compute, 2014.
- [8] Giuseppe Ateniese, Randal Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary Peterson "Provable factstenure at untrusted stores", in the ACM, 2007
- [9] B. Priyadarshini, P. Parvathi "Data Integrity in dim Storage", IEEE intercontinental consultation On advance In trade, knowledge And managing, 2012

A REVIEW ON IMPROVING SMART ATM SECURITY LEVELS

*Ms.P. Indumathy¹, Ms.V.Kaviya², Ms.V.Nisha³
Dept of Computer Applications,
MKJC, Tamil Nadu, India*

ABSTRACT

In the recent years ,Banking sectors and some of its related technologies had a high reach in customer satisfaction both economically and financially. Also in this spectacular modern world, everyone are closely dependent on computer technologies for gathering numerous records of data which they seek for. Automated Teller Machine which is abbreviated as ATM plays a major role in present days that was introduced to make people comfortable and easier to use money transactions. On June 27, 1967 the first ATM was installed in Enfield town in London by Barclays Bank. Basically Teller, Automated teller, Automated Transaction Money, Any Time Money (In India). In order to protect and safeguard money and details , Many ATM's uses Biometric technology like facial recognition, fingerprint and RFID (Radio frequency Identification) technology to prevent the data's that are not to be collected by the hackers. Even though some issues and troubles are happening in such ATM's. This paper is about how ATM are getting hacked and to get some tips to improve smart ATM security levels.

Keywords: Barclays bank, PIN, Facial recognition, fingerprint, RFID technology

INTRODUCTION

Automated teller machine is commonly known as ATM which is a technological computer that helps the bank account holders to maintain their money transactions. ATM 's are used to check out account balances, withdrawals, money depositing and transferring money from one bank account to another bank account. ATM was first invented by the three prominent inventors namely JOHN SHEPHERD-BARRON, DO DUC CUONG, DONALD WETZEL. At early 1960's, there was an American named Luther George Simjian invented the Bank graph which is a machine that only allows the bank account holders to deposit money and check it out. In 1967, a British inventor named John Shepherd-Barron was honoured for his invention of first ATM in London. The first

ATM was equipped in US(New York) in the year of 1969 with the slogan,"On September 2,Our Banks will open at 9Am and never close again".A year later in 1970,concept of Personal Identification Number(PIN) was generated by James good fellow for the verification of identity of account holders.ATM machines rapidly developed all over the world and installed nearly 1,00,000 in 1984.Then it reached more than three million in 2018 all over the worldwide.

METHODOLOGY

The ATM machine is architected with cash dispenser, keypad, display screen, screen button, speaker, ATM card reader, slot depositor etc., Inspite of all this it has a specialised Biometric scan technology such as facial recognition,fingerprint and RFID technology to protect all informations from scams and hackings.High resolution cameras are developed on facial recognitions where face cut of an individual,distance between mouth or nose can be analysed and captured.When the facial recognition is successfully completed,RFID automatically activated.Then the RFID reads the correct string and access the ATM or otherwise buzzer will on if the string is incorrect. But comparing to facial recognition and other security technologies for ATM fingerprint biometric technology is more better than other technologies.There are many types of ATM available in India namely Onsite ATM,Offsite ATM,White Label ATM,Yellow Label ATM,Brown Label ATM,Green Label ATM,Pink Label ATM,Orange Label ATM.Apart from all the activities like money transactions,viewing account balance ATM's are offering more facilities for the sake of customer needs.

Procedural Use of ATM:

The procedural steps to use an ATM properly are described below ;

- [1] Go to an ATM then insert your ATM card in which the bank logo should be inserted through the side and chip and take it out.
- [2] Next the bank account holder is requested to select language of his/her needs.
- [3] Use the Keypad to type your secret four digit PIN number.Make sure that the typed PIN number is not incorrect or else ATM card will be blocked.
- [4] Select your required transaction type like balance enquiry, withdraw money, deposit money, bill pay and others whatever you required.
- [5] Then you have to select the account type like current or savings or any other account.
- [6] Next enter the amount which you need to withdraw and press ok/enter.Always make sure that there is a available balance while withdrawing money.

[7] Then collect your money and receipt of the withdrawal as a proof where the receipt also contains the balance amount.

[8] Atlast stop the procedure by pressing cross/cancel button.

Risks and Issues in ATM:

There are many complexities in this existing system of an ATM which transacts via PIN and ATM-card based system. Now a days ATM scams are becoming high by using various technologies. Basically the criminals hack the PIN number with the people who are not aware of ATM pin code or pass code. In spite of having Biometric systems and all other crucial safety measures ATM's are being hacked by some culprits and frauds. Here are some issues to be focused where the thieves hack the ATM;

Card Shimming

It is a scam that focuses credit or debit cards developed with the EMV chip technology. It is about placing a fake card reader on point of systems and ATM's which resembles like original card.

It is a way of stealing personal information from ATM, credit card or debit card by fixing machines with some recording devices. Then the frauds can copy the data and pin numbers and access the money.

Card Stealing

Card stealing is about thieves will install a data-stealing software in card readers where all the card informations are stored in that and they steal the money or whatever the information they need.

Surfing Shoulder

It is the careless mistake of an individual while entering PIN number in ATM where they might not have noticed that the third person is looking over his/her shoulder to note the PIN number being entered. They just notice the PIN number and later they use it and drag money from their account. This is one of the major stealing method in ATM scams.

Guessing PIN number

This is actually a guessing process where the thief just guess the Four-Digit PIN number and try it by using various patterns. If the PIN number matches he/she will easily steal the money in very short period of time.

These mentioned scam techniques are to be noticed by all of us to protect our informations.

Improving Security Levels In ATM:

1. An ATM should be always monitored by bank staffs ,securities and also by CCTV so that we can easily find out if any scams had happened.
2. Jitter technology should be implemented so that fake card can be identified when it is inserted.
3. An ATM should have well versed safety lock and secured techniques so that unknown person can not access and hack the datas.
4. At the night time,be aware of the surroundings whether any doubtful person is near the ATM.
5. Networks must be secured, so that thieves can't hack the entire ATM network even is one ATM is hacked.
6. Keep your PIN number secret and do not share with anyone at anycause.
7. Make sure that while going onto an ATM nobody is following or spying you.
8. If you have any doubts regarding money transactions or any other issues don't ask help with the strangers.

CONCLUSION

Now a days banking sectors are becoming essential part of human life.Because everyone used to deposit or withdraw money for their regular needs.As soon as increase in ATM technology,frauds and thieves also increased in enormous in number.ATM's are the most comfortable and cost-effective method provided to the customers .The Frauds of the ATM not only make the financial loss to the Bank but also reduce the confidence level of the customers who use and believe ATM.Therefore it is the responsible of the banks to prevent and secure the ATM from frauds and thieves.Biometric is the most safest way to prevent ATM.The conclusion of this paper tells about what is ATM,how it is designed,what are the risks and issues in ATM,how to safeguard our ATM while doing transactions.

REFERENCES

1. Shaikha Al-Thani, "The Security of the ATM Machines in Relation to Students".
2. Deepak G. Deshekar, "Smart ATM Security Using Mobile Messaging".
3. Pavan S. Rane, Prashant P. Sawat, Sourabh B. Shinde, Prof. Nitin A. Dawande, "ATM Security".
4. Kavitha Hooda, "ATM Security".
5. Frimpong Twum, Kofi Nti, Michael Asante, "Improving Security Levels In Automatic Teller Machines (ATM) Using Multifactor Authentication".
6. [Http://www.ijaerd.com/papers/finished_papers/ATM_SECURITY-JAERDV05I0640547.pdf](http://www.ijaerd.com/papers/finished_papers/ATM_SECURITY-JAERDV05I0640547.pdf)
7. [Http://www.irjet.net/archives/v3/i7/IRJET-V3I765.pdf](http://www.irjet.net/archives/v3/i7/IRJET-V3I765.pdf)

AN OVERVIEW OF BLOCKCHAIN TECHNOLOGY

*Ms.R.Keerthana¹, Ms. D.Narmadha², Ms.S.Soundharya³
Dept of Computer Applications,
MKJC,Tamil Nadu, India*

ABSTRACT:

Blockchain has received much attention recently in technological industry. It is proven to be immutable. A blockchain technology allows transactions in a decentralized manner. In this paper, we provide the history, architecture, categories and industrial use of blockchain technology. Furthermore, the benefits and drawbacks of blockchain technology are briefly listed. Although the blockchain technology has great potential for the construction of the future interest system.

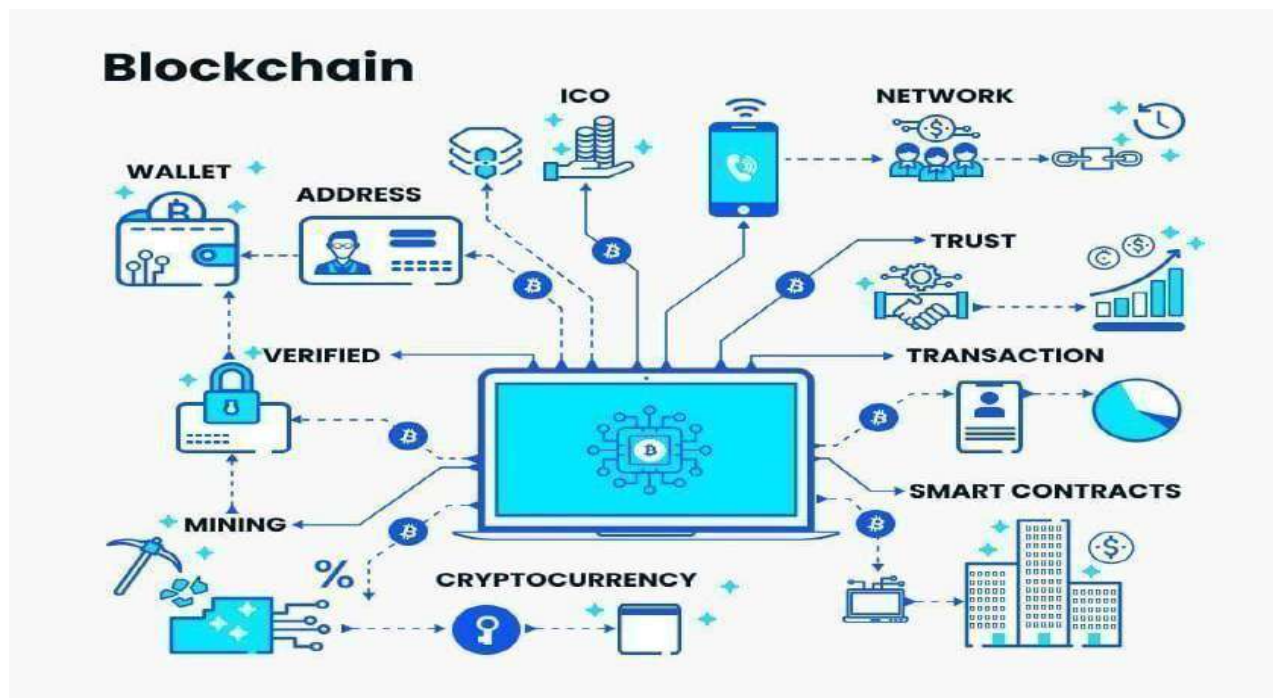
KEYWORDS: Cryptocurrencies, Blockchain, Database, Bitcoin, Network.

INTRODUCTION:

WHAT IS BLOCKCHAIN?:

In recent years, The blockchain technology plays a major role on technological development. Basically, a blockchain is cognate with cryptocurrencies like Bitcoin, Ethereum, Ripple. It is a particular type of database of record in which the transactions are maintained by a network of computers throughout the globe. A Blockchain stores different types of data in the blocks and then that are chained together. The term 'Blockchain' save record and keeping the data in back of the bitcoin network. This system of recording information makes impossible to change or hack the system. It is fundamentally a digital ledger of transactions (DLT) that is duplicated across the entire network of computer system on the block chain. Though a person perform a transactions, the transactions connect to the network then the computer algorithm regulate the transactions. If the

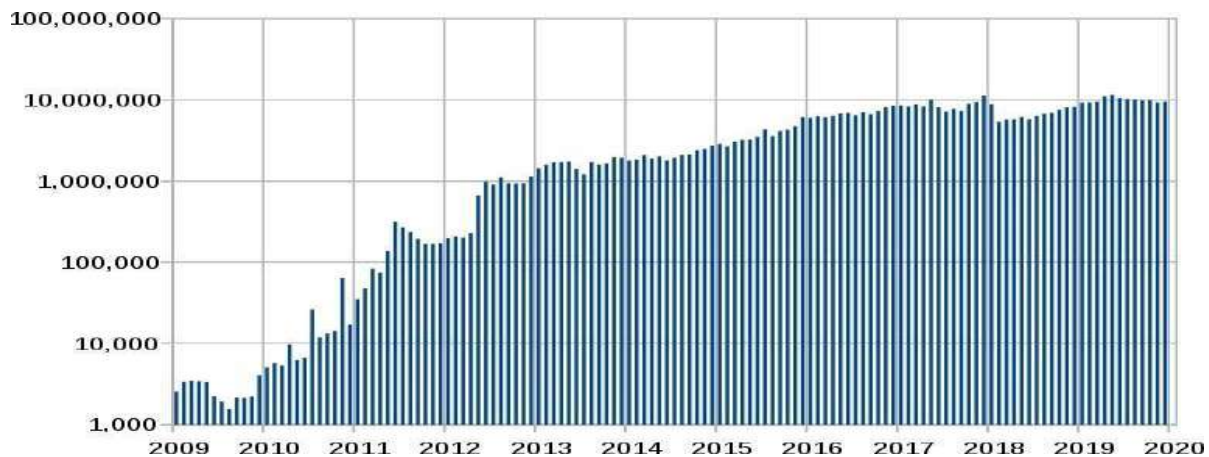
transations is established,the earlier transations and current transations are linked together and forma chain.This process is known as Block chain.



HISTORY OF BLOCKCHAIN:

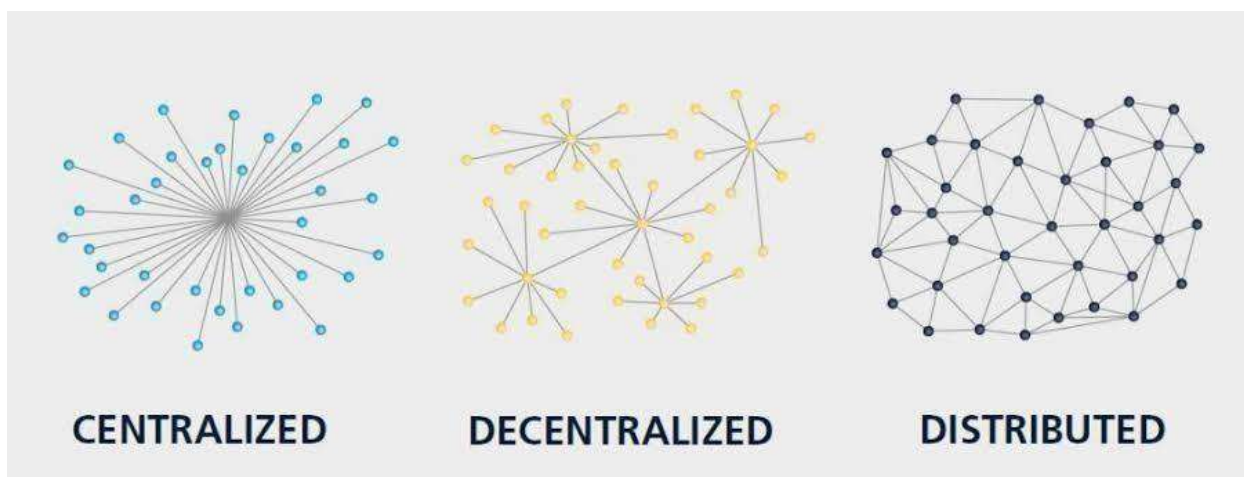
The Blockchain has the capability to play a major role on record keeping database systems throughout the world. This amazing technology was introduced to the world 10 years ago. This was introduced by unidentified user behind the online cash currency bitcoin, with the alias of Satoshi Nakamoto in 2008. The abstract of this paper was on the direct payment from one to the another source. This paper represented an electronic payment system based on cryptography.

Satoshi's paper stated that this digital currency cannot be duplicated, and one can spend it only one time. There are more different digital currencies like Litecoin, Dash, Dogecoin, etc., though Bitcoin takes the major and most popular cryptocurrency compared to others. Next to Bitcoin, Ethereum was launched which enables the block chain to work with loans and contracts. Ethereum was introduced in 2013 by programmer Vitalik Buterin. It is decentralized, open source blockchain with smart contract functionality. Ether is the endemic cryptocurrency of the platform.



ARCHITECTURE OF BLOCKCHAIN:

In general, Blockchain technology has the basic characteristics of decentralization, accountability and security. This blockchain technique can provide best operational processes and save costs efficiently. The demand and usage of applications built on blockchain architecture will only emerge. As a term, blockchain is a chain of blocks which contain specific information, but in a secure and genuine manner it is grouped together in a network. Especially, blockchain is a combination of computers linked to each other instead of a central server, so that the whole server is decentralized. The technique of blockchain allows digital data to be distributed, rather than copied. This distributed ledger contributes transparency, trust, and data security.



Blockchain architecture is being used very broadly in the financial industry. This architecture can serve the purposes for organizations and enterprises like,

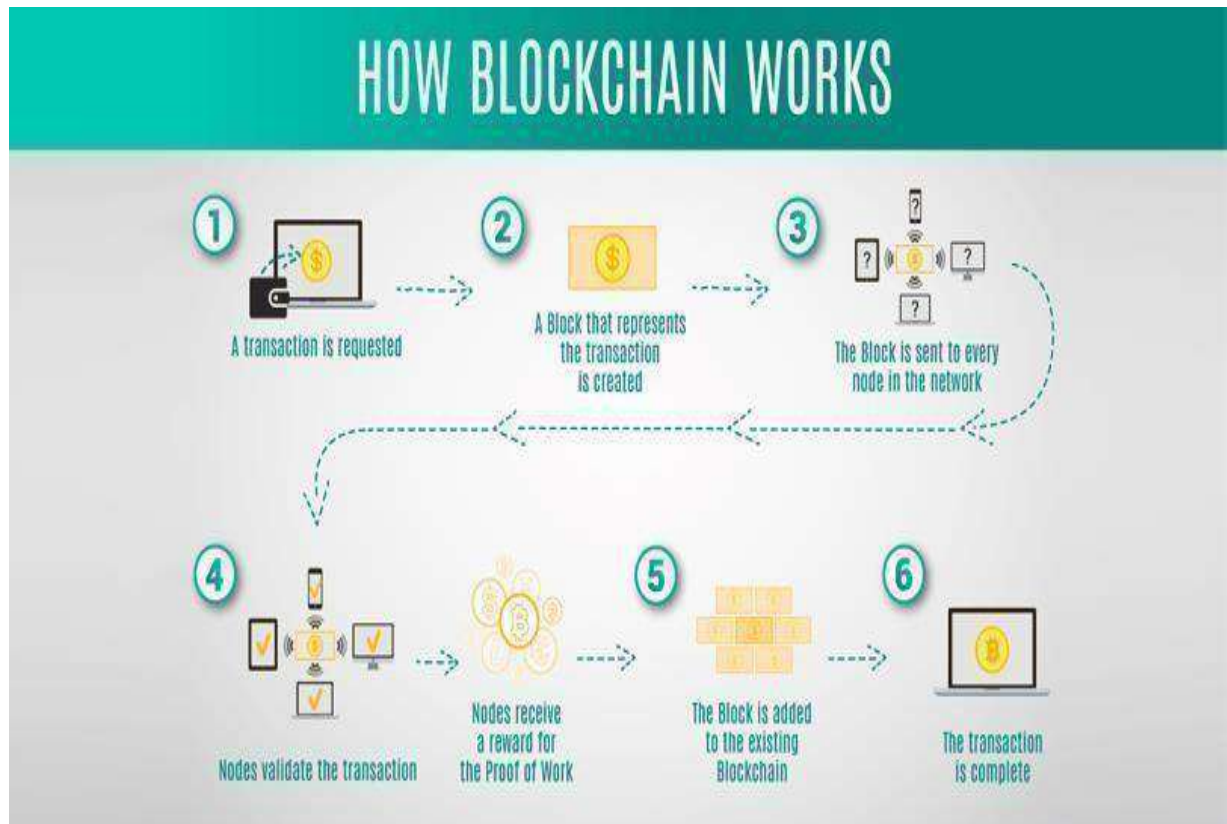
Cost reduction in banks and governmental institutions.

The history data of transactions can be viewed at any moment of time.

It guarantees high data security and validity.

HOW BLOCKCHAIN WORKS:

To explain how it works, the following diagram clearly gives details about the blockchain process by simple six steps,



TYPES OF BLOCKCHAIN:

Blockchain has derived greatly in the last few years and based on its different attributes, they can be divided into multiple types.

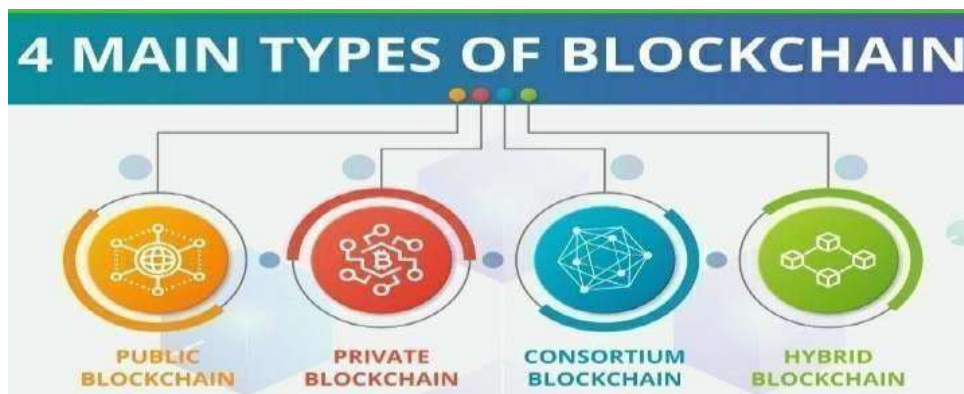
There are four main types of blockchain technology

Public Blockchain

Private Blockchain

Consortium Blockchain

Hybrid Blockchain



PUBLIC BLOCKCHAIN:

Public blockchains are initiated to the public and any separate can involve in the decision-making process by becoming a node, but users may or may not be benefited for their involvement in the decision-making process. No one in the network has ownership of the ledgers and are publicly open to anyone participated in the network. The users in the blockchain use a shared consensus mechanism to reach on a decision and maintain a copy of the ledger on their local nodes. The most basic use of blockchain is for mining and exchanging cryptocurrencies. Thus, the most familiar public blockchains are Bitcoin and Litecoin blockchains. Public blockchains are mostly secure if the users strictly follow security rules and methods. However, it is only risk when the participants don't follow the security protocols deeply.

Example: Bitcoin, Ethereum, Litecoin.

PRIVATE BLOCKCHAIN:

These types of blockchains are not initiated to the public and are initiated to only a group of people or organizations and the ledger is shared to its participated members only. Thus, private blockchains are comparable in use a public blockchain but have a small and restrictive network. Private blockchain networks are deployed for voting, supply chain management, digital identity, asset ownership, etc.

Examples : multichain and hyperledger projects (Fabric, Sawtooth), corda, etc.

CONSORTIUM BLOCKCHAIN:

A consortium blockchain is a semi-decentralized type where more than one organization conduct a blockchain network. This is hostile to what we saw in a private blockchain, which is managed by only a single organization. More than one organization can act as a node in this type of blockchain and exchange information or do mining. Consortium blockchains are typically used by banks, government organizations, etc.

Example:Energy Web Foundation,R3,etc.

HYBRID BLOCKCHAIN

A hybrid blockchain is a combination of the private and public blockchain.It uses the features of both types of blockchains that is one can have a private permission –based, system as well as a public permission-less system.With such a hybrid network,users can control who gets access to which data stored in the blockchain . only a selected section of data or records from the blockchain can be allowed to go public keeping the rest as confidential in the private network.the hybrid system of blockchain is flexible so that users can easily join a private blockchain with multiple public blockchains.A transaction in a private network of a hybrid blockchain is usually verified within that network.But users can also release it in the public blockchain to get verified.The public blockchains increase the hashing and involve more nodes for verification.This enhances the security and transparency of blockchain network.

Example: Hybrid blockchain is Dragonchain.

BENEFITS OF BLOCKCHAIN :

The blockchain is nothing short of a game-changing technology for anyone who chooses to use and master it, let's discuss the benefits of blockchain.

1. TRUST:

Enable trust between participants who don't know each other.

To create trust in the blockchain, one needs to verify relations between a few objects like Unique tokens, each containing a history of each previous owner.

2. DECENTRALIZED STRUCTURE :

Enable real-time data sharing among businesses like suppliers and distributors while reducing points of weakness.

3. IMPROVED SECURITY AND PRIVACY:

Creates an permanent records of transactions with End-to-End encryption, which reduces fraud and unauthorized activity.

4. REDUCED COSTS :

Creates performance by reducing manual tasks such as aggregating data and by basing reporting and auditing.

5. SPEED:

Ignore intermediaries so transactions are handled faster than conventional methods.

6. **VISIBILITY AND TRACEABILITY:**

Tracks the origins of a variety of items, such as medicines, to confirm they're legitimate instead of counter and organic items to confirm they're indeed organic.

7. **IMMUTABILITY :**

Ensures transactions can't be changed or deleted.

8. **INDIVIDUAL CONTROL OF DATA :**

Gives entities the ability to decide what digital data they wants to share, with whom and for how long, with limits enforced by smart contracts.

9. **TOKENIZATION :**

Converts value of an asset into a digital token recorded and shared via blockchain, Non-fungible tokens are used to sell digital art.

10. **INNOVATION:**

Leaders across multiple industries are exploring and implementing blockchain-based systems to solve intractable problems and improve long-standing cumbersome practices, such as verifying the info on a job resume.

DRAWBACKS OF BLOCKCHAIN :

Each coin has a flip side. Blockchain is a notch above its infancy today, and there are a number of drawbacks with the technology that need to be handled before it can be widely used for everyday transactions.

1. **SCALABILITY:**

Blockchain's application Bitcoin is massively popular. However, it can only handle 7 transactions per second. Where hyperledger can handle 10,000 and visa 56,000. The practical use of blockchain gets a bit hard to imagine with the issue of scalability in view.

2. **STORAGE :**

Since blockchain databases are stored indefinitely on all network nodes. The issues of storage surfaces. With the increasing number of transactions, the

size of the database will only expand, and there is no way personal computers can store unlimited data which only gets appended. To put this in perspective, the Ethereum blockchain is expanding at the speed of 55 GB/year and as on 2017. It is at 180 GB.

3. **PRIVACY** :

Data on a public blockchain is encrypted and anonymous, but lies in the hands of all nodes in the network. So, everyone in the network has a rightful access to this data there is a possibility someone could track down the identity of a person in the network through transactional data.

4. **REGULATIONS** :

Regulatory regimes in the financial arena are a challenge for blockchain's implementation. Blockchain applications will have to lay down the process of pinpointing the culprit in case a fraud takes place. Which is a bit of a challenge. Other regulatory aspects of blockchain technology will need to be laid down first in order to facilitate its wide adoption.

5. **SECURITY** :

Satoshi Nakamoto highlighted the '51% attack' when he launched Bitcoin. The attack can be simply put like this-if 51% of the nodes in a network lie. The lie will have to be accepted as truth. Therefore, everyone in the network will have to constantly have a watch on it to perceive any unwanted influence.

BLOCKCHAIN FOR EVERY INDUSTRIAL USE :

1. **LEGAL** :

"Smart Contracts" gathered on the blockchain track parties arrangement. Terms, transfer of ownership, and delivery of goods and services without the need for valid interruption.

2. **TRAVEL AND HOSPITALITY** :

In this passengers store their evidence "single travel ID" on the blockchain for use in lieu of travel record, identification cards, loyalty program IDs, and payment data.

3. **HEALTHCARE**:

Electronic medical records stored in a blockchain, accessed and updated via biometrics. Allow for the democratization of quiet data and lighten the difficulty of transferring records among workers.

4. FOOD :

Using this blockchain to store food supply chain data offers enlarge traceability of product origin, batching, processing, expiration, storage temperatures, and shipping.

5. RETAIL :

Protected peer-to-peer(P2P) marketplaces can track P2P retail transactions, with product information, shipment. And bills of lading input on the blockchain, and settlements made via Bitcoin.

6. SUPPLY CHAIN :

By promoting a distributed ledger, companies within a supply chain gain transparency into shipment tracking, deliveries, and progress among other suppliers where no inherent hope exist.

7. EDUCATION :

Educational institutions could utilize the blockchain to store graduating data around judgements, degrees, and transcripts, as well as verification of knowledge transfer between parties.

8. GOVERNMENT :

In Government the blockchain offers assurance as a technology to store personal identity information, criminal backgrounds, and “e-citizenship”, licensed by biometrics.

9. INSURANCE :

When autonomous vehicles and other smart devices communicate status updates with insurance providers via the blockchain, premium costs decrease as the need for analyzing and authenticating data exit.

10. ENERGY :

Decentralized energy transfer and distribution are possible via micro-transactions of data sent to blockchain, validated, and re-dispersed to the grid while securing payment to the submitter.

CONCLUSION:

In this paper, we present the overview on blockchain. We first give an overview of blockchain technology including architecture and types. Furthermore we listed some benefits and drawbacks of blockchain technology. Finally industrial uses of blockchain which shows the possible ways to

use in many industrial application. Also we plan to investigations on blockchain technology in future.

REFERENCES:

- [1] Pilkington,"Blockchain Technology:principles and applications." Research Handbook On Digital Transfer-ations(2016):225.
- [2]Crosby.,M.,Pattanayak,P.,Verma,S.,&Kalyanaraman,V.(2016).Blockchain Technology: Beyond Bitoin. Applied Innoa-Tion,26-10.
- [3] Atzori,Marcella. "Blockchain Tehnology And eentralied Goernane:Is The State Neessary?"(2015).
- [4] Malinoa,Katya, And Andreas Park."arket Design with Blockchain Tehnology".(2017).
- [5]Bashir,Imran.Mastering Blockchain. Packt Publishing Ltd,2017.

REVIEW ON ARTICLE ABOUT CLOUD COMPUTING

G.Sasirekha,¹S. Deepthi², K. Sathiya priya³

*Dept of Computer Science
MKJC, Tamil Nadu, India.*

ABSTRACT

This study helps organization and individuals to understand how cloud computing can provide them with reliable, customised and cost-effective services in a wide variety of applications. Cloud computing is a way of computing, where most of our data is stored in the cloud eg like internet. In this paper we have tried to explore various cloud computing Services, Application, Characteristics, Top trends and Security issues; We give various example for cloud services delivered by the most common Cloud Services Provider (CSP) such as Google, Microsoft, and Amazon. We have also discussed cloud computing services models and their benefits.

INTRODUCTION

Cloud computing is internet-based computing, whereby shared resources, software and data are provided to computers and different devices on demand, like the electricity grid. Cloud computing could be a paradigm shift following the shift from mainframe to client-server within the early Eighties. Details are abstracted from the users, who now haven't want for experience in or management over the technology infrastructure "in the cloud" that support them.

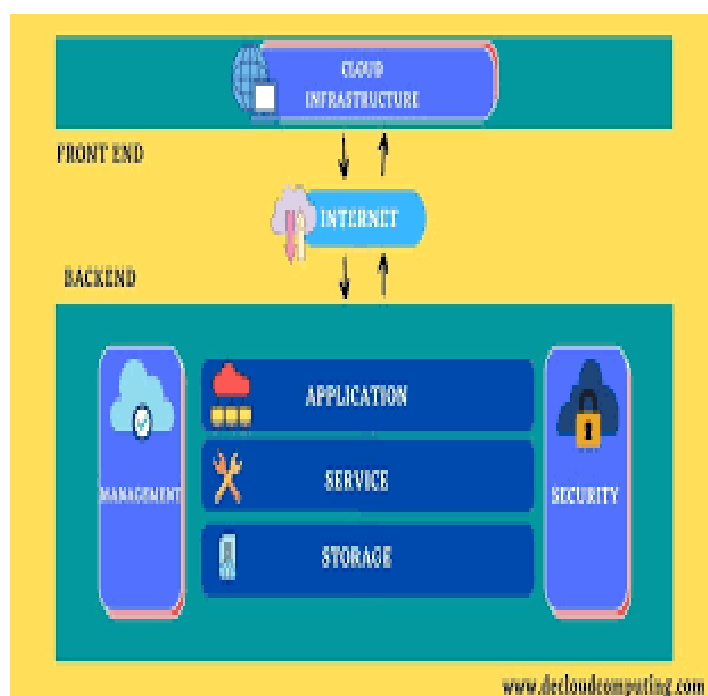
WHAT IS CLOUD COMPUTING?

Cloud computing is that the on-demand delivery of IT resources over the web with pay-as-go evaluation. instead of shopping for, owning and maintaining physical knowledge centers and servers, you can access technology services, such as computing power, storage, and info, on an as-needed basis from a cloud supplier like Amazon internet Services(AIS).

ARCHITECTURE OF CLOUD COMPUTING:

CloudConnect enables you to transfer valid data from the field level to the cloud. Combined with other information and its evaluation, CloudConnect opens up completely new possibilities:

- **Better product quality thanks to a big-data analysis of all relevant parameters**
- **Automatic process optimization thanks to the evaluation of movement data and the turnover rate of mobile assets**
- **Preventive maintenance – the evaluation of key machine and robot metrics to increase availability**



Front of the cloud computing design refers to the shopper facet if cloud facet of cloud computer shopper to access the cloud computing services/resource. For example use of an internet browser to access the cloud type.

CLIENT INFRASTRUCTURE:

Shopper infrastructure refers to the leaf finish element. It contains the appliance and user interfaces that area unit needed to access the cloud platform.

BACK END:

Face refers to the cloud itself that is employed by the service supplier. It contains the resources further as manages the resources and suppliers security mechanisms. Along with this it includes immense storage, virtual application, virtual machines, traffic controls mechanisms, deployments models etc.

APPLICATION:

Application in backend refers to a package or platform to that shopper accesses. Means it Provides the services in backend as per the shopper demand.

SERVICES:

Services in backend refers to the main 3 varieties of cloud based mostly services like Saas, Paas and Iaas. Also manages, which type of services the user accesses.

CLOUD RUNTIME:

Runtime cloud in face refers to produce of execution and runtime platform/environment to the virtual machine.

STORAGE:

Storage in face refers to produce versatile and ascendible storage service and management of keep knowledge.

INFRASTRUCTURE:

Cloud infrastructure in backend refers to hardware and package parts of cloud love it includes servers, storage, network devices, virtualization package etc.

MANAGEMENT:

Management in backend refers to management of face parts like applications, service, runtime cloud, storage, infrastructure, and alternative security mechanism etc.

SECURITY:

Security in face refers to implementation of various security mechanism in the backend for secure cloud resources, system, files and infrastructure to end-users.

INTERNET:

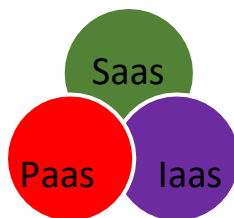
Net affiliation acts because the medium or bridge between frontend and backend and establishes the interaction and communication between frontend and backend.

CLOUD SERVICE MODELS:

INFRASTRUCTURE AS A SERVICE (IAAS).

PLATFORM AS A SERVICE (PAAS).

PACKAGE AS A SERVICE (SAAS).

**INFRASTRUCTURE as a SERVICE (IaaS):**

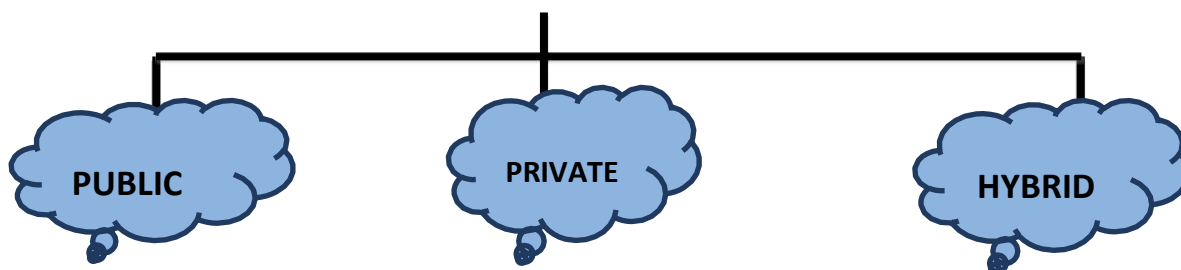
A trafficker provides consumer pay-as-you-go access to storage, networking, server and alternative computing resources within the cloud.

PLATFORM as a SERVICE (PaaS):

A service supplier offers access to a cloud-based surroundings during which users will build and deliver applications. The supplier provides underlying infrastructure.

SOFTWARE as a SERVICE (SaaS):

A service supplier delivers package and applications through the web. Users buy the package and access it via the net or trafficker APLs.

TYPES OF CLOUD COMPUTING:**TYPES OF CLOUD COMPUTING**

SECURITY:

Use of latest technologies and security technique to shield your knowledge ,applications and infrastructure escort cloud computing.New threats square measure arising everyday.

WHY CLOUD SECURITY:

LinkedIn- In 2012, 6.5 Million usernames and passwords square measure exposed by hackers.Sony- money knowledge , coming moving picture project and music were printed publically by hackers.I cloud- personalpictures of users were exposed to public by hackers.

IS CLOUD SECURITY VERY A PRIORITY:

- + Mix of art and science.
- + Art-Authentication ought to be outlined with user expertise in mind.
- + Science- you have got to come back up with new ways in which of securing your application.

VULNERABILITY,THREAT AND RISK:

- + Vulnerability refers to the weakness of associate degree quality that may be exploited by one or a lot of wrongdoer.
- + A threat is any event that has the potential to bring damage to a corporation or individual.
- + Risk refers to the potential for loss or injury once a threat exploits a vulnerability.

-Risk=Threat x Vulnerability

-Risk Management is a key to cloud security

MONITORING DATA:

- + Anomalies behavior are tracked.
- + Amazon Web service- Cloud Watch .

Set Alarms
View Graphs and statistics
Monitor and react to resources

GAINING VISIBILITY:

- + Need to know who is accessing through whose access point.
- + Cloud Trail is a logging service which can be used to log the history of API calls.
- + Used to identify which user/hacker from AWS Management console requested the particular service.

MANAGING ACCESS:

- + List of user who are holding authorization will be generated
- + Wipe the culprit or hacker out of the system.
- + AWS IAM-Identity and access management .
- + Erase and session and revoke the role.
- + Granular permission.
- + Secure access to application running on the environment.
- + Free-to-use.

ADVANTAGES OF CLOUD COMPUTING:

BACK-UP AND RESTORE INFORMATION:

Once the info is hold on in the cloud, it's easier to urge back-up and restore that information mistreatment the cloud.

IMPROVED COLLABORATION:

Cloud application improves collaboration by permitting teams of individuals to quickly and simply share data within the cloud via shared storage.

EXCELLENT ACCESSIBILITY:

Cloud permits North American country to quickly and access store data anyplace, anytime within the whole world, using an online association .An internet cloud infrastructure increase organization productivity and potency by making certain that our information is usually accessible.

LOW MAINTENANCE VALUE:

Cloud computing reduces each hardware and computer code maintenance prices for organizations.

MOBILITY:

Cloud computing permits North American country to simply access all cloud information via mobile.

SERVICES WITHIN THE PAY-PER-USE MODEL:

Cloud computing offers Application Programming Interfaces (APIs) to the users for access services on the cloud and pays the costs as per the usage of service.

UNLIMITED STORAGE CAPABILITY:

Cloud offers North American country a large quantity of storing capability for storing our necessary information like DOCUMENTS, IMAGES, AUDIO, VIDEO, ETC. IN ONE PLACE.

INFORMATION SECURITY:

Data security is one in every of the most important blessings of cloud computing. Cloud offers several advanced options associated with security and ensures that information is firmly hold on and handled.

DISADVANTAGES OF CLOUD COMPUTING:

A list of the disadvantage of cloud computing is given below

WEB PROPERTY:

As you recognize, in cloud computing, each information (image, audio, video, etc.) is hold on on the cloud, and that we access these information through the cloud by mistreatment the web association. If you are doing not have sensible web property, you can not access these information. However, we've got no the other thanks to access information from the cloud.

VENDOR LOCK-IN:

Vendor lock-in is that the biggest disadvantage of cloud computing. Organizations might face issues once transferring their services from one seller to a different. As totally {different|completely different} vendors offer different platforms, that may cause problem moving from one cloud to a different.

RESTRICTED MANAGEMENT:

As we know, cloud infrastructure is totally owned , managed, and monitored by the service supplier, that the cloud users have less management over the perform and execution of services at intervals a cloud infrastructure. Although cloud service suppliers implement the most effective security standards to store necessary data. But,

SECURITY:

Before adopting cloud technology, you ought to remember that you just are going to be causation all of your organization's sensitive data to a 3rd party, i.e., a cloud computing service supplier. whereas causation the info on the cloud, there could also be an opportunity that your organization's data is hacked by Hackers.

CONCLUSION:

Cloud Computing is outpacing the IT business. Real business price may be realised by customers of all sizes. Cloud solutions square measure straightforward to accumulate, don't need future contracts and square measure easier to rescale and down as required. Proper coming up with and migration services square measure required to confirm a thriving implementation. Public and personal Clouds may be deployed along to leverage the most effective of each. Third party watching services guarantee client have gotten the foremost out of their cloud surroundings. Security Compliance and watching is doable with careful coming up with and analysis.

REFERENCE:

https://www.powershow.com./view0/68f816-N2JjO/Cloud_Computing_powerpoint_ppt_presentation#.
<https://www.researchgate.net/publication/352477780>.
<https://www.zdnet.com>.

An Algorithm to implement Higher Order Mining with Logistic Regression and VARMA Method For Forecasting the yield of sugarcane

Mrs.M.Deepanayaki, PhD Research scholar¹,

Department of Computer Science, Periyar University, Salem, Tamilnadu, India

Dr. Vidyaathulasiraman, AP/HOD²,

Department of Computer Science, Government Arts & Science College (W),

Bargur, Tamil nadu, India

ABSTRACT

This paper proposes an algorithm to implement the higher order mining [1] with a classification technique and Time series trend analysis VARMA- Vector Auto Regression Moving Average technique. Datasets of Area, Production and yield is taken to implement the concept is taken from the ministry of Agriculture and Farmer welfare. After the Preprocessing and transformation of data, classification technique with highest accuracy is taken to fit (Since the data has more independent variable and label is with binary, Logistic regression is preferred). After the classification process, with the predicted information, the data is applied to the VARMA (Vector Auto Regression Moving-Average model) to forecast the production of sugarcane in the coming years. VARMA model is chosen since the dataset contains dataset contains more than one time-dependent variable.

KEYWORDS: Higher order mining, Time series trend Analysis, VARMA, etc.,

I. INTRODUCTION

Agriculture has become increasingly high-tech over the years. Now a days agriculture professionals

DM	Association mining	Clustering	Classification	Trend analysis	Rule structure
HOM					
Association mining					
Clustering					
Classification					

digging into the data and use it to make highly informed decisions. The advancements of data scientists are making this reality possible, in the present and also for the future. Data mining [2] methods used for data analysis, created with the aim to find out specific dependence, relations and rules related to data and making them out in the new higher level quality information. Data mining algorithms and techniques are thus used for transforming data into business information and thereby improving the decision making

process. Higher Order Mining (HOM) is a form of data mining that is applied over Non-primary, Derived data or Patterns [2]. HOM paradigm reveals further potential for knowledge discovery, including the delivery of rules and patterns with semantics that are closer to human intuition and are thus more appropriate for human inspection.

I. ALGORITHMIC CATEGORIZATION OF HOM APPROACHES

Higher order mining is the process of combination of various data mining techniques such as association rule, classification techniques, clustering etc.,

The above table is used to know the possibilities of performing higher order mining.

The colored portion reveals that it is possible to do

Association mining (1st order) and association mining (2nd Order)

Association mining (1st order) and clustering (2nd order) and so on.,

The uncolored portion reveals that it is not possible to perform classification and Association mining, classification and Clustering, Classification and Rule structure.

Hence the proposed algorithm used the combination of a Classification technique (Logistic regression) and a Time series Trend analysis technique (VARMA)

II. DATA COLLECTION

Researches on Agricultural field become an inevitable because of its need to have a boom for the enrichment of healthy human life. Today Agricultural organizations are in the position to work with very huge amounts of farming and its related data. Dataset Taken From Directorate of Economics and Statistics, Ministry of Agriculture and Farmers Welfare.

Attributes are

District	object
Year	object
Season	object
Area (Hectare)	int64
Production (Tonnes)	float64
Yield (Tonnes/Hectare)	float64
class	object

As sugarcane is cultivated by planting in either January-February, July-August or October-November, with maturity duration of 12-18 months, it is challenging task for predicting crop yield, as time series problem across various states of India.

III. TECHNIQUES & TOOLS USED

- Data Preprocessing : Standard scaler, Imputer
- Data transformation: Principle Component Analysis
- Classification technique : LOGISTIC REGRESSION
- Forecasting technique : VARMA

Data Preprocessing:

Preprocessing and retrieval of significant data from this abundance of agricultural information is very much necessary, which enable the reduction of manual tasks and easier data extraction directly from electronic sources, transfer to secure electronic system of documentation which will enable production, cost reduction, higher yield and higher market price. It is necessary to collect and store raw data of agricultural enterprises which is very ample and diverse, in an organized form, and their integration enables the creation of agricultural information system. Data processing and therefore the varied methodologies related to it will scale back the complexness of the agricultural information to make proper choices and decisions.

The steps of Data Preprocessing usually fall into two categories:

1. Selecting data and attributes for the analysis.
2. Creating/changing the attributes.
3. Machine Learning (Data mining) is 80% of preprocessing and 20% of model making
4. While implementing Machine learning models data preprocessing is the common factors of success of a model, i.e., if there is correct data preprocessing and feature engineering, that model is more likely to produce noticeably better results as compared to a model for which data is not well preprocessed. There are 4 main important steps for the preprocessing of data.

- Splitting of the data set in Training and Validation sets
- Taking care of Missing values
- Taking care of Categorical Features
- Normalization of data set

The main idea behind the train test split is to convert original data set into 2 parts

- train
- test

where train consists of training data and training labels and test consists of testing data and testing labels.

Standard scaler, Imputer

StandardScaler **removes the mean value of the given data and scales each feature/variable to unit variance**. This operation is performed feature-wise in an independent way. StandardScaler can be influenced by errors (if they exist in the dataset) since it involves the estimation of the empirical mean and standard deviation of each feature.

For various reasons, many real world datasets contain missing values, often encoded as blanks, NaNs or other placeholders. A better strategy is to impute the missing values, i.e., to infer them from the known part of the data.

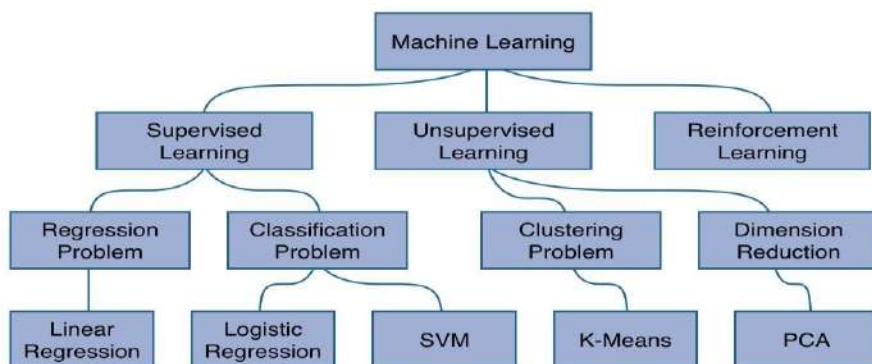
Data transformation:

Data transformation is **the process of converting data from one format or structure into another format or structure**. Data transformation is critical to activities such as data integration and data management. Perform data mapping to define how individual fields are mapped, modified, joined, filtered, and aggregated.

PCA :

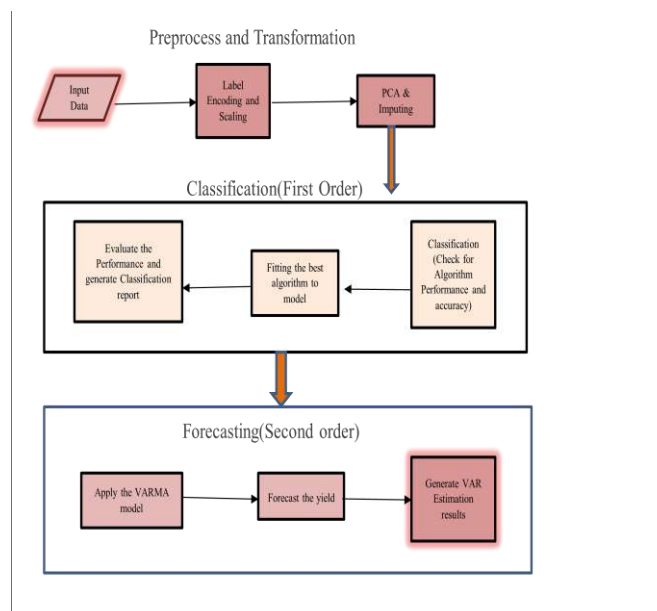
Principal component analysis (PCA) is **a technique for reducing the dimensionality of such datasets**, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize variance.

Machine Learning(Data Mining) technique:



- Machine Learning is broadly classified into three categories depending on the nature of the learning ‘signal’ or ‘feedback’ available to a learning system.
- Supervised learning:** Computer is presented with inputs and their desired outputs. The goal is to learn a general rule to map inputs to the output.
- Unsupervised learning:** Computer is presented with inputs without desired outputs, the goal is to find structure in inputs.
- Reinforcement learning:** Computer program interacts with a dynamic environment, and it must perform a certain goal without guide or teacher.

IV. ARCHITECTURE



LOGISTIC REGRESSION

Logistic regression is one of the most Machine Learning algorithms, comes under the Supervised Learning technique. It is **used for predicting the categorical dependent variable using a given set of independent variables**. Logistic regression predicts the output of a categorical dependent variable. It predicts the probability of occurrence of an event by fitting data to a logistic_function. Hence, it is also known as **logistic regression**. Since, it predicts the probability, the output values lies between 0 and 1 (as expected).

FORECASTING TECHNIQUES -VARMA

Time series forecasting is needed to make scientific predictions based on historical time stamped data. It involves building models through historical analysis and using them to make observations and drive future strategic decision-making. VARMA (Vector autoregressive moving average) is used in the proposed algorithm as it is the combination of Auto regression and Moving average method which is an effective method to forecast the dataset that contains more than one time-dependent variable. The vector autoregressive moving average (VARMA) model is one of the statistical analyses frequently used in several studies of multivariate time series data in economy, finance, and business. It is used in numerous studies because of its simplicity. Furthermore, it can be used to predict and forecast time series data.

V. ALGORITHM

Algorithm: Higher order mining algorithm

Input: Dataset of sugarcane yield of India from the year 1997 to 2018

Attributes: District, Year, Season, Area, Production and yield and label (Class: High and Low)

Output: Forecast the yield

Step 1: Importing the necessary libraries

Step 2: Read the dataset(apv_India.csv) and distributing the dataset into X and y components for data analysis.

Step 3:Pre-processing part on training and testing set such as fitting the Standard scale.

i) Convert categorical attributes (District, Year, Season, class) to numerical attributes

Step 4: Applying the PCA function into training and testing set for analysis.

Step 5: Check for model algorithms performance and select the best to fit

Step 6 : Fitting the best model (Logistic Regression) To the test set & training set

Step 7:Visualizing the results

Step 8: Forecast the future production of sugarcane based on predictions made by the model to the Time series data mining algorithm (VARMA) – Vector Auto Regression Moving-Average (VARMA)

Step 9: Evaluating the forecast model

I. CONCLUSION

Agricultural machine learning is not a secretive trick or magic, but a set of well-defined models that collect specific data and apply specific algorithms or models to achieve expected results. The proposed work determines the influence of Location, season and Area of cultivation that influences the yield of crop. Higher order mining requires a change in perspective for knowledge discovery, from the analysis of data to the analysis of patterns (Trained model).

REFERENCES

- [1] Changsheng Zhu, Christian Uwa Idemudia, Wenfang Feng “Improved logistic regression model for diabetes prediction by integrating PCA and K-means techniques” 2352-9148/ © 2019 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).
- [2] Shenzhi Li, Tianhao Wu, William M. Pottenger, Lehigh University, “Mining Higher-Order Association Rules from Distributed Named Entity Databases”
- [3] Jyotshna Solanki, Prof. (Dr.) Yusuf Mulge, “Different Techniques Used in Data Mining in Agriculture”, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 5, May 2015
- [4] B.Milovic, V.Radojevic, “Application of Data mining in Agriculture”, Bulgarian Journal of Agricultural Science, 21(no 1)2015, 26-34
- [5] **Ramesh A. Medar , Vijay S. Rajpurohit and Anand M. Ambekar ,** “ Sugarcane Crop Yield Forecasting Model Using Supervised Machine Learning” *I.J. Intelligent Systems and Applications*, **2019, 8, 11-20** Published Online August 2019 in MECS (<http://www.mecs-press.org/>) DOI: 10.5815/ijisa.2019.08.02
- [6] Arulbalaji, P. and Gurugnanam, B. “Evaluating The Normalized Difference Vegetation Index Using Landsat Data By Envi In Salem District, Tamilnadu, India” ISSN: 2230-9926 , *International Journal of Development Research*, Vol. 4, Issue, 9, pp. 1844-1846, September, 2014
- [7] Phusanisa Charoen-Ung, Pradit Mittrapiyanuruk, “Sugarcane Yield Grade Prediction using Random Forest and Gradient Boosting Tree Techniques” Conference Paper · July 2018 DOI: 10.1109/JCSSE.2018.8457391
- [8] Monitoring and mapping of seasonal vegetation trend in Tamil Nadu using NDVI and NDWI imagery Journal of Applied and Natural Science 11(1): 54 - 61 (2019) ISSN : 0974-9411 (Print), 2231-5209 (Online) journals.ansfoundation.org

A REVIEW PAPER ON CLOUD COMPUTING IN INDUSTRIES

Ms. Kaviyanjali¹

*Dept of Computer Application,
MKJC, Tamil Nadu, India*

Ms.P. Arishitha²

*Dept of Computer Application,
MKJC, Tamil Nadu, India*

ABSTRACT

Attributable to the instant growth of computing resources over the past years, has let to more flexibility. Cloud computing has been flourished has a prominent and this technology appeared as a new solution in IT field. The future of all levels of the manufacturing industry will incorporate this technology to stay more securely connected with consumers and the supply chain. It appears as a unique and attractive computing pattern. It has various avails atop grid computing and other computing. Around the world, 90% of companies are using some type of cloud technology. In this paper, we have given a brief evolution of cloud computing by reviewing more than 10 articles on cloud computing. The outcome of the review explores and conclude the face of the IT industries before and after computing and also express the influence, pros and cons of cloud computing.

KEY WORDS:

Cloud Computing, SAAS, IAAS, PAAS, Private cloud, public cloud, hybrid cloud.

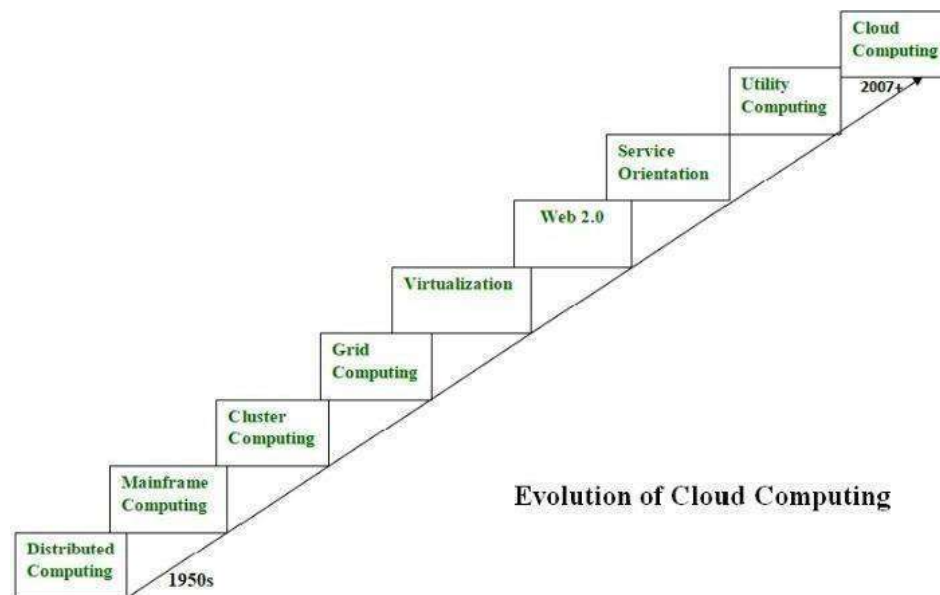
1 .INTRODUCTION

Cloud computing is the delivery of different services through the internet, including data storage, servers, databases, networking, and software. Cloud-based storage makes it possible to save files to a remote database and retrieve them on demand. Cloud computing provides an alternative to the on-premises datacenter. With an on-premises data center, we have to manage everything, such as purchasing and installing hardware, virtualization, installing the operating system, and any other required applications, setting up the network, configuring the firewall, and setting up storage for data. After doing all the setup, we become responsible for maintaining it through its entire lifecycle. But if we choose cloud computing, a cloud vendor is responsible for the hardware purchase and maintenance. They also provide a wide

variety of software and platforms as a service. We can take any required services on rent. The cloud computing services will be charged based on usage.

2 .EVOLUTION OF CLOUD COMPUTING

Cloud computing is all about renting computing services. This idea first came in the 1950s. In making cloud computing what it is today, five technologies played a vital role. These are distributed systems and its peripherals, virtualization, web 2.0, service orientation, and utility computing.



.DISTRUBUTED SYSTEM

It is composed of multiple independent systems but all are depicted as a single entity to the users. The aspiration of distributed systems is to share resources and also use them completely and efficiently. Distributed systems possess components such as scalability, concurrency, continuous availability, heterogeneity, and independence in failures. But the main problem with this system was that all the systems were required to be present at the same geographical location. Hence to solve this problem, distributed computing led to three more types of computing and are-mainframe.

||.MAINFRAME COMPUTING

Mainframes which first came into existence in 1951 are highly capable and respectable computing measures. These are answerable for handling large data such as massive input-output operations. Even today these are used for quantity processing tasks such as online transactions etc. These systems have almost no downtime with high fault tolerance. After distributing computing, these increased the processing efficiency of the system. But these were very lavish. To reduce this cost, cluster computing came as an alternative to mainframe technology.

III. CLUSTER COMPUTING

In 1980s, cluster computing came as an alternative to mainframe computing. Each machine in the cluster was connected to each other by a network with high frequency. These were way cheaper than those mainframe systems. These were equitably capable of high computations. Also new nodes could easily be added to the cluster if it was required. Thus, the problem of the cost was solved to some expansion but the problem related to geographical constraints still pertained. To solve this, the concept of grid computing was introduced.

V GRID COMPUTING:

In 1990s, the concept of grid computing was introduced. It means that different systems were placed at entirely different geographical locations and these all were connected via the internet. These systems belonged to different organizations and thus the grid consisted of different nodes. Despite it solved some problems but new problems emerged as the distance between the nodes expanded.

3 .MAIN TYPES OF CLOUD COMPUTING

There are four main types of cloud computing: private clouds, public clouds, hybrid clouds, and multi clouds.

PRIVATE CLOUDS-private clouds are defined as cloud environments solely dedicated to a single end user or group, where the environment usually runs behind that user or groups firewall. All clouds become private clouds when the underlying IT infrastructure is dedicated to a single customer with completely isolated access.

PUBLIC CLOUDS- public clouds are cloud environments typically created from IT infrastructure not owned by the end user. Some of the largest public cloud providers include Amazon web Services (AWS), Google cloud, IBM cloud, and Microsoft azure.

HYBRID CLOUDS-A hybrid cloud is a seemingly single IT environment created from multiple environments connected through local area networks(LAN), wide area networks(WAN), virtualprivate networks(VPN),and API.

MULTI CLOUDS-Multi clouds are cloud approach made up of more than one cloud service, from more than one cloud vendor-public or private. All hybrid clouds are multi clouds, but not all multi clouds are hybrid clouds. Multi clouds become hybrid clouds when multiple clouds are connected by some form of integration or orchestration.

There are also three main service models for cloud computing:1 .Infrastructure as a service(IAAS),2 .platform as a service(PAAS),and 3 .software as a service(SAAS). In the IAAS model, cloud service provider(CSP) only manage the hardware and virtualization layers.

(IAAS)-IAAS means a cloud service provider manages the infrastructure for you-the actual servers, network, virtualization, and data storage-through an internet connection.

(PAAS)-PAAS means the hardware and an application-software platform or provided and managed by an outside cloud service provider, but the user handles the apps running on top of the platform and the data the app relies on.

(SAAS)-SAAS is a service that delivers a software application-which the cloud service provider manages-to its users. SAAS apps are web applications or mobile apps that users can access via web browser.

HOW CLOUD COMPUTING INFLUENCING THE IT INDUSTRY

The traditional way of building an IT environment is to buy servers, hardware, licenses and to install the software. This is a long and costly process, involving many infrastructure demands and long deployment cycles. This fully IT internal model may be common place, but IT as we know it today is being replaced by newer technologies. Lately, cloud computing is causing a major shift in the IT industry. New technologies have been developed, and now there are various ways to virtualize IT systems and to access the needed application on the internet, through web based applications. This means no IT costs for hardware or servers.

The IT infrastructure will be crucially changed, as more applications are being moved to private or public clouds. Software developers will have to adjust the ways they create and deliver applications. The effort to maintain the data is also diminished. Although clients may not be comfortable with this fact, they should understand the data in the cloud can be safer than being in-house. Enterprise cloud providers that offer a managed cloud solution have security experts on

REVIEW ON ARTICLE ABOUT ARTIFICIAL INTELLIGENCE

Ms.S. Chandra Praba¹, Ms.P. Yogeshwari², Ms.R.S Batma Jayani³

*Dept of Computer Science
MKJC, Tamil Nadu, India*

ABSTRACT:

Artificial Intelligence(A.I) multi disciplinary field whose automateactivities that presently require human intelligence.Recent successes in A.I include computerize medical diagnosticians and systems that automatically customize hardware to particular user requirements. The major problem areas addressed in A.I can be summarized as Perception, Manipulation, Reasoning, Communicationand Learning. In simple Artificial Intelligence is an approach to make acomputer, a robot, or a productto think how smart human think. Themain aim of A.I is to improve computer functions which are related to human knowledge.

Keywords: Artificial Intelligence ,Machine Learning ,Deep Learning, Neural Network, Natural Language Processing, Computer Vision, Cognitive Computing ,Self-Driving Cars , Robotics

The role of Artificial Intelligence in future technology² The birth and evolution of AI

The start of AI is believed to be made by Alan Turing with his question "CAN MACHINE THINK?" [1]. The Turing test, developed by Turing in 1950, is a test of a machine's ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human. The test set some requirements to build a truly intelligent machine that requires knowledge representation, natural language, machine learning, automated reasoning, vision, and robotics for the full test. Since then, the term AI was first introduced by John McCarthy and it was closely associated with the field of "symbolic AI", which was popular until the end of the 1980s. In the 1990s, the new concept of "intelligent agent" [13] emerged. An agent is a system that perceives its environment and undertakes actions that maximize its chances of being successful. To overcome some of the limitations of symbolic AI, sub-symbolic methodologies such as neural networks, fuzzy systems, evolutionary computation, and other computational models started gaining popularity, leading to the term "computational intelligence" emerging as a subfield of AI. Different approaches and methods are being used in AI. Two major methodologies or beliefs are the top-down and bottom-up methods. The top-down theorists believe in mimicking the human brain's behavior with computer programs, whereas the bottom-up theorists believe that the best way to achieve AI is by building electronic replicas similar to the human brain's complex network of neurons. Recently, the term AI encompasses the whole conceptualization of a machine that is intelligent in terms of both operational and social consequences.³ AI applications and future technology AI is ubiquitous and is not only limited to computer science but has evolved to include other areas like health [14–17], security [18–21], education [22, 23], music [24–28], art [29, 30], and business [31, 32] application. Many AI applications are

deeply embedded in the infrastructure of every industry. AI is expected, in a few years, to touch nearly all the industries [33] and there are plenty of ways AI is and can transform certain industries. AI is currently being utilized for a wide range of activities including medical diagnosis, electronic trading platforms, robot control, and remote sensing. It has been used to develop and advance numerous fields and industries, including finance, healthcare, education, transportation, and robotics. AI researchers have created many tools to solve the most difficult problems in computer science and other fields. The current AI performance ranges between sub-human, optimal, and super-human performance. A wide range of tasks can be solved by AI applications including facial recognition, speech recognition [34], object recognition [35], image classification [36] and surpassing human-level intelligence in The Game of Go [11], Chess [37], Dota 2 [38], and StarCraft II [39–43]. Another focus of AI technologies lies between the areas of healthcare and privacy with the advancement of federated learning [44] and privacy-perceiving machine learning [45]. AI used as clinical decision support systems for medical diagnosis, computer-aided interpretation of medical images, and companion robots. It can even produce music usable in a medical setting by computer-generated music for stress and pain relief. Moreover, initiatives like Google Magenta [46], conducted by the Google Brain team, want to seek out if AI can be capable of making compelling art and music. One more active field of AI research is the use of AI to create other AI. This includes Google's AutoML project to evolve a new neural network topologies with new architectures and topologies exceeding the performance of all previously published ImageNet [47] performance. This is also extended to the current research of Generative Adversarial Networks [48] and the work done by the research team from the visual computing group of the Technical University of Munich and Stanford University developed Face2Face [49], a program which animates the face of a target person, transposing the facial expressions of an exterior source. Since then, other methods have been demonstrated based on deep neural networks, from which the name "Deep Fake" [50] was taken. Recently, new research directions, focus and initiatives have arisen including research in quantum machine learning [51], hierarchical reinforcement learning [52], Bayesian deep learning [53], affective computing & Human-Centered AI [54, 55], neuroscience [56, 57], self-driving cars [58], and conversational agents [59, 60].

4 Opportunities, Limitations and Ethics of AI

Given the exponential rise of interest in AI, major studies have started on the impact of AI on society, not only in technological but also in legal, and ethical areas. This also includes the spec-

ulation that autonomous super AI may at some point supersede the cognitive capabilities of humans. This future scenario is called the "AI SINGULARITY" [61, 62], defined as the ability of machines to build better machines by themselves. Current AI researchers are more focused on developing systems that are excellent at tasks in a narrow range of applications. This focus is at odds with the idea of the pursuit of artificial general intelligence (AGI) [63, 64] that could mimic all different cognitive abilities related to human intelligence such as self-awareness and emotional knowledge. Current AI development and the status of our hegemony as the most intelligent species on earth, further societal concerns are raised. However, AI technologies still limited to very specific applications. One limitation of AI is the lack of "commonsense" the power to judge information beyond its acquired knowledge. AI is also limited in terms of emotional intelligence. AI can only detect basic human emotional states such as anger, joy, sadness, stress, pain, fear, and neutrality. Emotional intelligence is one of the next frontiers of higher levels of personalization. The computer science principles driving AI forward, are rapidly advancing and it is important to assess its impact, not only from a technological standpoint but also from a social, ethical and legal perspective.

INTRODUCTION:

Artificial Intelligence is that it depends on who you ask. A layman with a fleeting understanding of technology would link it to robots. They'd say Artificial Intelligence is a terminator like-figure that can act and think on its own. If you ask about artificial intelligence to an AI researcher, would say that it's a set of algorithms that can produce results without having to be explicitly instructed to do so. And they would all be right. So to summaries, Artificial Intelligence meaning is: An intelligent entity created by humans, Capable of performing tasks intelligently without being explicitly instructed, Capable of thinking and acting rationally and humanely.

LITERATURE REVIEW:

The term Artificial Intelligence was reviewed by various technologists and search encyclopedias. According to them Artificial Intelligence (A.I) is the science and engineering of making intelligent machines (original definition by John McCarthy who coined the term 'Artificial Intelligence').

ence' in 1955), The science of making machines do things that would require intelligence if done by men (definition offered by A.I pioneer Marvin Minsky in 1968), The science of making machines smart (Demis Hassabis, CEO and founder of Deep Mind, now part of Google), is an intelligent machine (Google's Avinash Kaushik), Is the next, logical step in computing : a program that can figure out things for itself. It's a program that can reprogram itself (Jim Sterne author of Artificial Intelligence for Marketing), anything that makes machines act more intelligently (IBM's definition) a constellation of technologies that extend human capabilities by sensing, comprehending, acting and learning –allowing people to do much more" (Accenture's definition)

DESIGN AND IMPLEMENTATION:

Building an AI system is a careful process of reverse-engineering human traits and capabilities in a machine, and using its computational prowess to surpass what we are capable of. To understand How Artificial Intelligence actually works, one needs to deep dive into the various subdomains of Artificial Intelligence and understand how those domains could be applied into the various fields of the industry.

Machine Learning: ML teaches a machine how to make inference and decisions based on past experience. It identifies patterns analyses past data to infer the meaning of these data points to reach a possible conclusion without having to involve human experience. The automation to reach conclusions by evaluating data, saves a human time for businesses and helps them make a better decision.

Deep Learning: Deep Learning is an ML technique. It teaches a machine to process inputs through layers in order to classify, infer and predict the outcome.

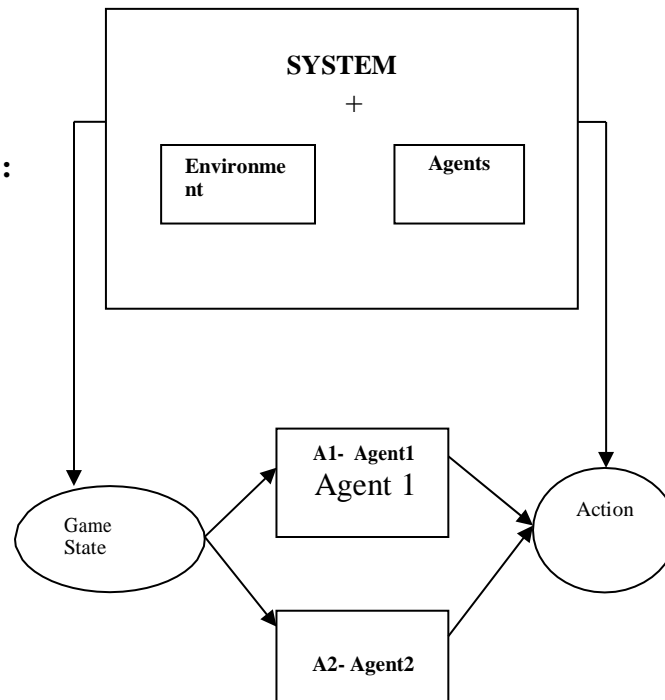
Neural Networks: Neural Networks work on the similar principles as of Human Neural cells. They are a series of algorithms that captures the relationship between various underlying variables and processes the data as a human brain does.

Natural Language Processing: NLP is a science of reading, understanding , interpreting a language by a machine. Once a machine understand what the user to communicated, it response accordingly.

Computer Vision: Computer vision algorithms tries to understand an image by breaking down an image and studying different parts of the objects. This helps the machine classify and learn from a set of images, to make a better output decision based on previous observations.

Cognitive Computing: Cognitive computing algorithms try to mimic a human brain by analyzing speech/images/objects in a manner that a human does and tries to give the desired output.

ARCHITECTURE :



METHODOLOGY:

In recent years, there have been advances in artificial intelligence (AI) for a broad range of machine learning techniques, such as deep learning, reinforcement learning, and transfer learning. These are considered efficient and advanced tools to address non-trivial technical challenges in both science and engineering. AI-based solutions have been adopted to address challenges in many applications, for example smart energy systems, intelligent manufacturing, intelligent transportation, healthcare, and public safety, among others. In parallel, the Internet of Things (IoT) has experienced rapid development in the past decade. It deploys a massive number of smart terminals, including sensors, actuators, and

establishes ubiquitous connections between smart devices and objects, with which one can perceive the world more clearly and control various systems more accurately. IoT also interconnects various separated intelligent systems into a vast smart world.

At present, there is an increasing demand for the convergence of AI and IoT to tackle programming issues in both scientific and engineering applications, such as high-performance data processing and analysis for intelligent decision-making of large-scale complex systems, and lightweight machine learning-based solutions in IoT-driven applications. Novel methods, models, algorithms, and tools are considered worth further research to improve AI and IoT driven solutions in terms of efficiency, scalability, security, and resilience, which significantly benefits the scientific and engineering programming community.

This Special Issue will focus on state-of-the-art research efforts and new perspectives in scientific and engineering computing driven by the convergence of AI and IoT, with particular focus on scientific innovations in methodologies, approaches, and algorithms to enable the efficient, sustainable, and secure operation of complex systems, such as smart grids, or transportation systems. This Special Issue can serve as a forum for researchers and professionals to explore and develop knowledge and insights into AI and IoT driven applications in various domains. We welcome both original research and review articles.

Potential topics include but are not limited to the following:

Scientific programming theories, methods and tools for the convergence of AI and IoT, Security, scalability, reliability, and portability in machine learning and IoT driven applications, AI and IoT driven framework for ubiquitous sensing and data analytics, AI and IoT driven techniques in cloud-edge computing environments, Architecture, operation paradigms, and resource management for the convergence of AI and IoT systems, Cost-benefit analysis of AI and IoT enabled systems against legacy systems, Test beds, implementation, and practices of AI and IoT driven solutions in different application domains, e.g., smart grid, advanced manufacturing, intelligent transportation, healthcare, smart cities, and public safety

FUTURE ENANCEMENT:

As humans, we have always been fascinated by technological changes and fiction, right now, we are living amidst the greatest advancements in our history. Artificial Intelligence has emerged to be then extbigthingin the field of technology. Organizations across the world are coming up with break through innovations in artificial intelligence and machine learning. Artificial intelligence is not only impacting the future of every industry and every human being but has also acted as the main driver of emerging technologies like big data, robotics and IoT. Considering its growth rate, it will continue to act as a technological innovator future. Hence, there are immense opportunities for trained and certified professionals to enter are warding career. As the setechnologies continue to grow, they will have more and more impact on the social setting and quality of life.

CONCLUSION:

To sum up, Artificial Intelligence is widespread in the lives of humans. The break of the 21st century led to the development of artificial intelligence that has further expanded the use of technology in numerous area. All in all, AI has numerous applications that make it beneficial for humans in day-to-day lives. While many scholars have warned mankind against AI, the artificial intelligence future is highly important for the development of technology.

References:

- [1] A. M. Turing. "Computing machinery and intelligence". In: Parsing the Turing Test. Springer, 2009, pp. 23–65.
- [2] J. McCarthy. "Artificial intelligence, logic and formalizing common sense". In: Philosophical logic and artificial intelligence. Springer, 1989, pp. 161–190.
- [3] J. McCarthy. Artificial intelligence: a paper symposium: Professor Sir James Lighthill, FRS. Artificial Intelligence: A General Survey. In: Science Research Council, 1973. 1974.
- [4] S. Russel et al. Artificial Intelligence: A Modern Approach. 3rd ed. Prentice Hall, 2010.
- [5] D. Chen et al. Autonomous Driving using Safe Reinforcement Learning by Incorporating a Regret-based Human Lane-Changing Decision Model. 2019.

- [6] A. Ferdowsi et al. "Robust Deep Reinforcement Learning for Security and Safety in Autonomous Vehicle Systems". In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (2018). DOI:
- [7] P. Palanisamy. Multi-Agent Connected Autonomous Driving using Deep Reinforcement Learning. 2019.
- [8] S. Wang et al. "Deep Reinforcement Learning for Autonomous Driving". In: a preprint arXiv:1811.11329 (2018).

BIO MEDICAL INFORMATICS*Ms.M.Haripriya¹, Ms.B.Swathi², Ms.B.Sneha³**Dept Of Computer Application,MKJC.***ABSTRACT:**

Biomedical informatics lacks a clear and theoretically grounded definition. Many proposed definitions focus on data, information, and knowledge, but do not provide an adequate definition of these terms. Leveraging insights from the philosophy of information, we define informatics as the science of information, where information is data plus meaning. Biomedical informatics is the science of information as applied to or studied in the context of biomedicine. Defining the object of study of informatics as data plus meaning clearly distinguishes the field from related fields, such as computer science, statistics and biomedicine, which have different objects of study

KEYWORDS:

Biomedical informatics, scientific discipline, data, information, knowledge, definition, philosophy of information

INTRODUCTION:

Biomedical informatics has been an “emerging field” for decades. Concern about medical information and the desire to computerize health care are hardly new. Though originally focused on traditional paper-based medical records and their management rather than electronic medical records, the American Health Information Management Association (AHIMA) was founded in 1928 as the American Association of Medical Record Librarians.

Definition can help the field address practical issues, such as:

*Educational program design: provide a clear vision of our field to students, guide curriculum development and evaluation within training programs

*Administrative decisions: make a clear and consistent case for resources to administrators, to guide informatics units (academic and service-oriented) with respect to hiring faculty or staff, relationship to other organizational units and performance metrics

*Communication: including internal communication among informaticians and external communication with those outside of our field; a definition can help match current and potential collaborators, guide informatics societies such as the American and International Medical Informatics Associations (AMIA and IMIA, respectively), and help funding agencies and members of the general public understand our role and contributions

*Research agenda: provide a basis for identifying fundamental research questions, and to distinguish basic research in informatics from applied work

Definition of informatics:

Informatics study information (data + meaning, in contrast to focusing exclusively on data), with usage and effects. The practitioners must understand the context or domain, in addition to abstract properties of information and their representation.

DISCUSSIONS:

*Implication 1: Defining informatics as the study of data + meaning clearly distinguishes informatics from important related fields

Defining the central object of study of informatics as data + meaning allows us to distinguish informatics as a science from computer science, mathematics, statistics, the biomedical sciences and other related fields. It also clarifies the role of each of these fields in informatics.

Computer science is primarily the study of computation. Computer scientists seek to provide solutions to general problems by classifying computational problems in terms of formal abstract properties and deriving effective, efficient algorithms for solving them. The meaning of the data being manipulated by an algorithm is not important. An algorithm to find the shortest path connecting two nodes in a network depends only on the length of the edges, not whether the edges and nodes represent a geographical map, computer network, or social network. On the other hand, computer science plays an important role in informatics. There can be no information without data, and computers are the best medium we have for reliably storing, transmitting, and manipulating data. Thus, some informatics develop methods that allow computers to process data as if the computer understands the meaning; and to produce tools that allow human beings to make more sense of data displayed by the computer, thereby turning it into information.

*Implication 2: Computation is an important tool for informatics, but is not the primary object of study and is neither a necessary nor sufficient condition for informatics

In our definition, information, not computation, is the primary object of study of informatics. Many activities in informatics have nothing to do with computation (computers). Within health care, time-based, source-based, and problem-oriented medical records are all important informatics products that predate computers. Thus a central concern in informatics is: what information is needed and how it is best represented to support a specific set of human

activities. Information architecture and book shows are all important informatics tools that do not depend on computers. Computation is increasingly important as the amount of available information increases exponentially.=

*Implication 3: The emphasis on meaning allows us to see why some informatics problems are easier than others

This definition allows us to understand why some informatics problems are easier than others. Consider the banking system. Clearly it is quite complex and involves a great deal of data and meaning. In contrast to biomedicine, we hear no arguments regarding the suitability of computers to track accounts. We argue that in the case of banking, there is a very narrow semantic gap.

Biomedical informatics advancing the national health agenda: the AMIA 2015 Year in review in clinical and consumer informatics. There are significant advances in establishing policies for EHR feature implementation, it increased interoperability for these to gain traction. Decision support systems improve practice behaviours, evidence of their impact on clinical outcomes is still lacking. Progress in natural language processing (NLP) suggests that we are approaching but not achieved truly interactive NLP systems. Prediction models are becoming more robust by the lack of interoperable clinical data records. Consumers can and will use mobile applications for improved engagement.

Biomedical Informatics on the Cloud

A Treasure Hunt for Advancing Cardiovascular Medicine

In the digital age of cardiovascular medicine, the rate of biomedical discovery can be greatly accelerated by the guidance and resources required to unearth potential collections of knowledge. A unified computational platform leverages metadata to not only provide direction but also empower researchers to mine a wealth of biomedical information. As an internet-based computing solution, cloud computing gives shared resources on demand for data storage, processing, and dissemination is reliable, cost-effective, and customizable to suit individual user's needs. These services provide users with a simple way to access databases, servers, storage and applications.

Biomedical and health informatics approaches remain essential for addressing the COVID-19 pandemic

Colleagues provide an overview of the rationale, design, infrastructure, and deployment of the National COVID Cohort Collaborative (NC3) an open science community focused on analyzing individual-level data from many centers, was developed by the Clinical

and Translational Award Program, the National Center for Translational Sciences, and the scientific community to enable rapid collaboration among clinicians, researchers, and data scientists to identify treatments and specialized care and subsequently reduce the immediate and long-term consequences of COVID-19.

CONCLUSION:

Biomedical informatics is the application of the science of information as data plus meaning to problems of biomedical interest. Since its beginnings, biomedical informatics innovations have been developed to support the needs of various stakeholders including biologists, clinicians/clinical researchers, epidemiologists, and health services researchers. The inclusion of biomedical informatics in the translational medicine team may thus help enable a trans-disciplinary paradigm shift towards the development of the next generation of groundbreaking therapies and interventions.

REFERENCES:

- 1.Greenes RA, Shortliffe EH : Commentary : Informatics in biomedicine and health care. Acad Med. 2009, 84: 818-820. 10.1097/ACM.0b013e3181a81f94.
- 2.Shortliffe EH, CiminoJJ : Biomedical informatics: computer applications in health care and biomedicine. 2006, New York, NY: Springer, 3.
3. AHIMA facts. 2007.[cited 2007 December 17]; Available from: <http://www.ahima.org/about/about.asp>.

DATA SCIENCE AND BIG DATA

R.Harshavardhini¹, A.Lokeshwari², S.Sowmiya³

Dept of computer Science, MKJC,TN,India.

Abstract-

Data science encompasses a set of principles, problem definitions, algorithms, and processes for extracting no obvious and useful patterns from large data sets. Many of the elements of data science have been developed in related fields such as machine learning and data mining. Data science has attracted a lot of attention, promising to turn vast amounts of data into useful predictions and insights. In this article, we ask why scientists should care about data science. To answer, we discuss data science from three perspectives: statistical, computational, and human. Although each of the three is a critical component of data science, we argue that the effective combination of all three components is the essence of what data science is about.

Keywords-Data science, data science components, data science process, data science jobs roles, tools of data science, challenges of data science technology, applications of data science, prerequisite of data science, big data, types of big data, characteristics of big data, Advantages of big data.

INTRODUCTION

Data Science is the area of study which involves extracting insights from vast amounts of data by the use of various scientific methods, algorithms, and processes. It helps you to discover hidden patterns from the raw data. The term Data Science has emerged because of the evolution of mathematical statistics, data analysis, and big data. Data Science is an interdisciplinary field that allows you to extract knowledge from structured or unstructured data. Data science enables you to translate a business problem into a research project and then translate it back into a practical solution.

In short, we can say that data science is all about:

- ✓ Asking the correct questions and analyzing the raw data.
- ✓ Modelling the data using various complex and efficient algorithms.
- ✓ Visualizing the data to get a better perspective.

Understanding the data to make better decisions and finding the final result.

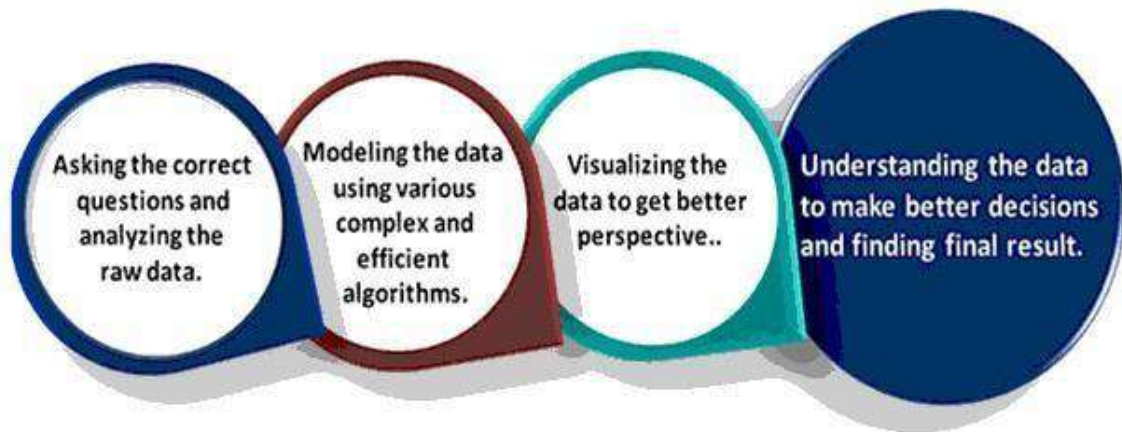


Fig 1.1 Data science

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

Types Of Big Data:

- Structured ·
- Unstructured
- Semi-structured

II Why Data Science?

Here, are significant advantages of using Data Analytics Technology:

Data is the oil for today's world. With the right tools, technologies, algorithms, we can use data and convert it into a distinctive business advantage. Data Science can help you to detect fraud using advanced machine learning algorithms. It helps you to prevent any significant monetary losses. Allows to build intelligence ability in machines. You can perform sentiment analysis to gauge customer brand loyalty. It enables you to take better and faster decisions. Helps you to recommend the right product to the right customer to enhance your business,

Following are some main reasons for using data science technology:

With the help of data science technology, we can convert the massive amount of raw and unstructured data into meaningful insights.

Data science technology is opting by various companies, whether it is a big brand or a start-up. Google, Amazon, Netflix, etc, which handle the huge amount of data, are using data science algorithms for better customer experience.

Data science is working for automating transportation such as creating a self-driving car, which is the future of transportation. Data science can help in different predictions such as various survey, elections, flight ticket confirmation, etc.

III Data Science Components:

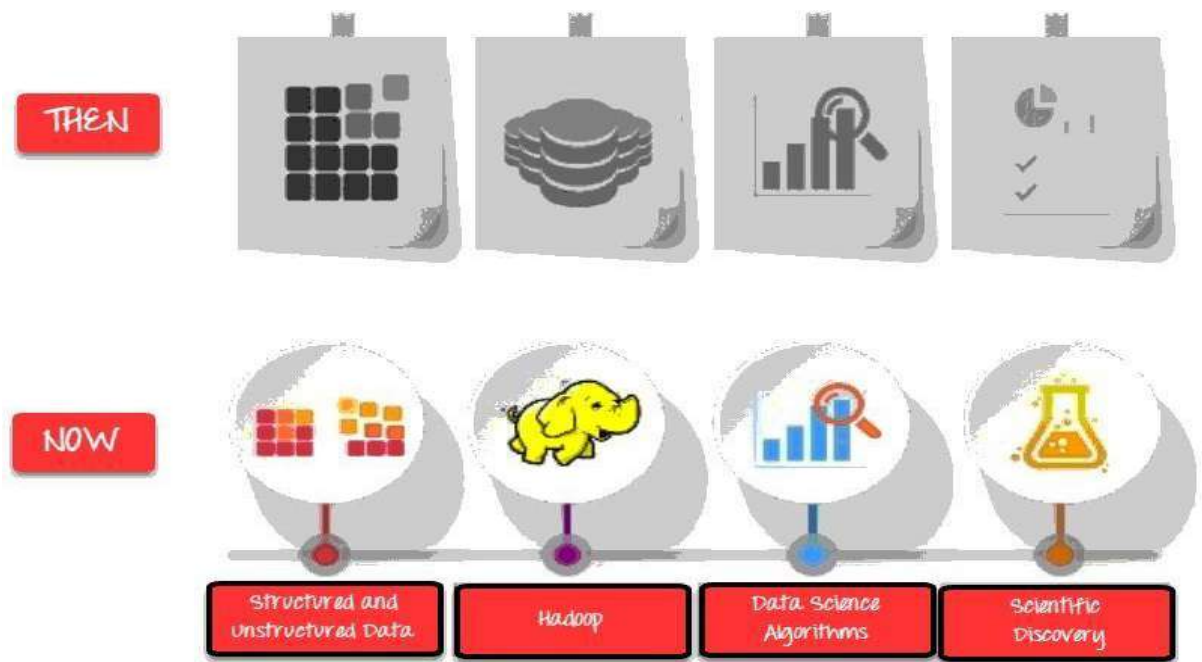


Fig 1.2 Evolution of Data science

Statistics:

Statistics is the most critical unit of Data Science basics. It is the method or science of collecting and analyzing numerical data in large quantities to get useful insights.

Visualization:

Visualization technique helps you to access huge amounts of data in easy to understand and digestible visuals.

Machine Learning:

Machine Learning explores the building and study of algorithms which learn to make predictions about unforeseen/future data.

Deep Learning:

Deep Learning method is new machine learning research where the algorithm selects the analysis model to follow.

Domain Expertise:

In data science, domain expertise binds data science together. Domain expertise means specialized knowledge or skills of a particular area. In data science, there are various areas for which we need domain experts.

Advanced computing:

Heavy lifting of data science is advanced computing. Advanced computing involves designing, writing, debugging, and maintaining the source code of computer programs.

Data engineering:

Data engineering is a part of data science, which involves acquiring, storing, retrieving, and transforming the data. Data engineering also includes metadata (data about data) to the data.

Mathematics:

Mathematics is the critical part of data science. Mathematics involves the study of quantity, structure, space, and changes. For a data scientist, knowledge of good mathematics is essential.

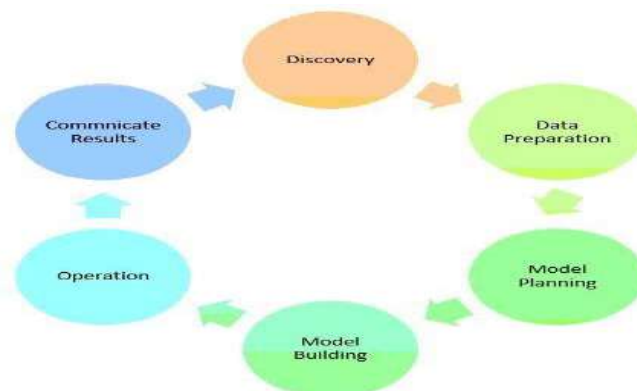
IV Data Science Process:

Fig 1.3

1. Discovery:

Discovery step involves acquiring data from all the identified internal & external sources which helps you to answer the business question.

The data can be:

- ✓ Logs from web servers
- ✓ Data gathered from social media
- ✓ Census datasets

- ✓ Data streamed from online sources using APIs

2. Preparation:

Data can have lots of inconsistencies like missing value, blank columns, incorrect data format which needs to be cleaned. You need to process, explore, and condition data before modelling. The cleaner your data, the better are your predictions.

3. Model Planning:

In this stage, you need to determine the method and technique to draw the relation between input variables. Planning for a model is performed by using different statistical formulas and visualization tools. SQL analysis services, R, and SAS/access are some of the tools used for this purpose.

4. Model Building:

In this step, the actual model building process starts. Here, Data scientist distributes datasets for training and testing. Techniques like association, classification, and clustering are applied to the training data set. The model once prepared is tested against the "testing" dataset.

5. Operationalize:

In this stage, you deliver the final baseline model with reports, code, and technical documents. Model is deployed into a real-time production environment after thorough testing.

6. Communicate Results:

In this stage, the key findings are communicated to all stakeholders. This helps you to decide if the results of the project are a success or a failure based on the inputs from the model.

V Data Science Jobs Roles:

- ✓ Most prominent Data Scientist job titles are:
- ✓ Data Scientist
- ✓ Data Engineer
- ✓ Data Analyst
- ✓ Statistician
- ✓ Data Architect
- ✓ Data Admin
- ✓ Business Analyst
- ✓ Data/Analytics Manager

Data Scientist:**Role:**

A Data Scientist is a professional who manages enormous amounts of data to come up with compelling business visions by using various tools, techniques, methodologies, algorithms, etc.

Languages:

R, SAS, Python, SQL, Hive, Matlab, Pig, Spark

Data Engineer:**Role:**

The role of data engineer is of working with large amounts of data. He develops, constructs, tests, and maintains architectures like large scale processing system and databases.

Languages:

SQL, Hive, R, SAS, Matlab, Python, Java, Ruby, C + +, and Perl

Data Analyst:**Role:**

A data analyst is responsible for mining vast amounts of data. He or she will look for relationships, patterns, trends in data. Later he or she will deliver compelling reporting and visualization for analyzing the data to take the most viable business decisions.

Languages:

R, Python, HTML, JS, C, C+ + , SQL

Statistician:**Role:**

The statistician collects, analyses, understand qualitative and quantitative data by using statistical theories and methods.

Languages:

SQL, R, Matlab, Tableau, Python, Perl, Spark, and Hive

Data Administrator:**Role:**

Data admin should ensure that the database is accessible to all relevant users. He also makes sure that it is performing correctly and is being kept safe from hacking.

Languages:

Ruby on Rails, SQL, Java, C#, and Python

Business Analyst:**Role:**

This professional need to improves business processes. He/she as an intermediary between the business executive team and IT department.

Languages:

SQL, Tableau, Power BI and, Python

VI What is Big Data?

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently. Big data is also a data but with huge size.

Types of Big Data:

- Structured
- Unstructured
- Semi-structured

Structured:

Any data that can be stored, accessed and processed in the form of fixed format is termed as a 'structured' data. Over the period of time, talent in computer science has achieved greater success in developing techniques for working with such kind of data (where the format is well known in advance) and also deriving value out of it. However, nowadays, we are foreseeing issues when a size of such data grows to a huge extent, typical sizes are being in the rage of multiple zetta bytes.

Unstructured:

Any data with unknown form or the structure is classified as unstructured data. In addition to the size being huge, un-structured data poses multiple challenges in terms of its processing for deriving value out of it. A typical example of unstructured data is a heterogeneous data source containing a combination of simple text files, images, videos etc. Now day organizations have wealth of data available with them but unfortunately, they don't know how to derive value out of it since this data is in its raw form or unstructured format.

Semi-structured:

Semi-structured data can contain both the forms of data. We can see semi-structured data as a structured in form but it is actually not defined with e.g. a table definition in relational DBMS. Example of semi-structured data is a data represented in an XML file.

VII Characteristics of Big Data:

Big data can be described by the following characteristics:

- ✓ Volume
- ✓ Variety
- ✓ Velocity
- ✓ Variability

(i) Volume – The name Big Data itself is related to a size which is enormous. Size of data plays a very crucial role in determining value out of data. Also, whether a particular data can actually be considered as a Big Data or not, is dependent upon the volume of data. Hence, ‘Volume’ is one characteristic which needs to be considered while dealing with Big Data solutions.

Variety – The next aspect of Big Data is its variety.

Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications. Nowadays, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. are also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analyzing data.

(iii) Velocity – The term ‘velocity’ refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data.

Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks, and social media sites, sensors, Mobile devices, etc. The flow of data is massive and continuous.

(iv) Variability – This refers to the inconsistency which can be shown by the data at times, thus hampering the process of being able to handle and manage the data effectively.

VIII Advantages of Big Data Processing:

Businesses can utilize outside intelligence while taking decisions, Access to social data from search engines and sites like facebook, twitter are enabling organizations to fine tune their business strategies.

Improved customer service

Traditional customer feedback systems are getting replaced by new systems designed with Big Data technologies. In these new systems, Big Data and natural language processing technologies are being used to read and evaluate consumer responses.

Early identification of risk to the

Early identification of risk to the product/services, if any

Better operational efficiency

IX Scope of future use:

Companies' Inability to handle data:

Data is being regularly collected by businesses and companies for transactions and through website interactions. Many companies face a common challenge – to analyze and categorize the data that is collected and stored. A data scientist becomes the savior in a situation of mayhem like this. Companies can progress a lot with proper and efficient handling of data, which results in productivity.

Revised Data Privacy Regulations:

Countries of the European Union witnessed the passing of the General Data Protection Regulation (GDPR) in May 2018. A similar regulation for data protection will be passed by California in 2020. This will create co-dependency between companies and data scientists for the need of storing data adequately and responsibly. In today's times, people are generally more cautious and alert about sharing data to businesses and giving up a certain amount of control to them, as there is rising awareness about data breaches and their malefic consequences. Companies can no longer afford to be careless and irresponsible about their data. The GDPR will ensure some amount of data privacy in the coming future.

Data Science is constantly evolving:

Career areas that do not carry any growth potential in them run the risk of stagnating. This indicates that the respective fields need to constantly evolve and undergo a change for opportunities to arise and flourish in the industry. Data science is a broad career path that is undergoing developments and thus promises abundant opportunities in the future. Data science job roles are likely to get more specific, which in turn will lead to specializations in the field. People inclined towards this stream can exploit their opportunities and pursue what suits them best through these specifications and specializations.

Programming & Frameworks Training**An astonishing incline in data growth**

Data is generated by everyone on a daily basis with and without our notice. The interaction we have with data daily will only keep increasing as time passes. In addition, the amount of data existing in the world will increase at lightning speed. As data production will be on the rise, the demand for data scientists will be crucial to help enterprises use and manage it well.

Virtual Reality will be friendlier:

In today's world, we can witness and are in fact witnessing how Artificial Intelligence is spreading across the globe and companies' reliance on it. Big data prospects with its current innovations will flourish more with advanced concepts like Deep Learning and neural networking. Currently, machine learning is being introduced and implemented in almost every application. Virtual Reality (VR) and Augmented Reality (AR) are undergoing monumental modifications too. In addition, human and machine interaction, as well as dependency, is likely to improve and increase drastically.

Block chain updating with Data science:

The main popular technology dealing with crypto currencies like Bit coin is referred to as Blockchain. Data security will live true to its function in this aspect as the detailed transactions will be secured and made note of. If big data flourishes, then Iot will witness growth too and gain popularity. Edge computing will be responsible for dealing with data issues and address them.

X The Future of Big Data in India:

The Indian industrial ecosystem is changing. Big data is opening unprecedented opportunities that were unimaginable even a few years ago. The demand for data analytics in India is on the rise. It has led to an increase in the demand for data scientists in the country. Candidates opting for a career in data science should know about big data technologies and tools like Hadoop, Hive, Spark Streaming, and others. Several industries in India, like e-commerce, manufacturing, and retail, have taken up big data to ensure customer satisfaction and business growth. The future of big data in India is bright.

XI Flexible career options:

When it comes to job positions and roles, Big Data is one of the most versatile career options. As Analytics is a crucial tool used in many different fields, you get a host of job titles to choose from including:

- ✓ Big Data Engineer
- ✓ Big Data Analyst
- ✓ Big Data Analytics Architect
- ✓ Big Data Solution Architect
- ✓ Analytics Associate
- ✓ Metrics and Analytics Specialist
- ✓ Big Data Analytics Business Consultant
- ✓ Business Intelligence and Analytics Consultant

XII Conclusion:

Big Data is a game-changer. Many organizations are using more analytics to drive strategic actions and offer a better customer experience. A slight change in the efficiency or smallest savings can lead to a huge profit, which is why most organizations are moving towards big data.

Data science education is well into its formative stages of development; it is evolving into a self-supporting discipline and producing professionals with distinct and complementary skills relative to professionals in the computer, information, and statistical sciences.

References:

1. Website link: DataScienceCentral.com
2. Website link: WhatsTheBigData.com
3. Website link: InsideBIGDATA.com
4. Website link: StarbridgePartners.com/Data-Science-Report
5. A.Chebotko, A. Kashlev, and S. Lu,"A Big Data Modeling Methology for Apache Cassandra," proc. Of IEEE Int. Congress on Big Data,2015,pp.238-p.245.

COMPUTER GRAPHICS, VISION, ANIMATION AND GAME SCIENCE

Ms.Nabeela Nousheen¹ , Ms.P.Vaideeswari² , Ms.A.Preethika³

Dept of Computer Science ,MKJC,Tamil Nadu, India.

ABSTRACT:

Computer graphics deals with generating images with the aid of computers. Today, computer graphics is a core technology in digital photography, film, video games, cell phone and computer displays, and many specialized applications. Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do. Computer animation is the process used for digitally generating animated images. The more general term computer-generated imagery encompasses both static scenes and dynamic images, while computer animation only refers to moving images.

Computer game science is designed around a set of core courses that introduce the fundamentals of computer science (programming, data structures, graphics and artificial intelligence), math (statistics, linear algebra and logic), and games (games and society, game design, game engines and multiplayer games).

KEYWORDS:

Graphics ,Vision, Animation and Game Science.

I. INTRODUCTION

Today, computer graphics is a core technology in digital photography, film, video games, cell phone and computer displays, and many specialized applications. Computer graphics deals with the generating images with the aid of computers. It is a vast and recently developed area of Computer Science. Computer vision tasks include methods for acquiring, processing, analyzing and understanding digital images and extraction of high-dimensional data from the real world in order to produce numerical or symbolic information. Computer Animation is a form of pictorial presentation which refers to simulated motion pictures showing movement of drawn objects. Game design is art of applying design and aesthetics to create a game for entertainment or for educational, exercise or experimental purposes.

II. COMPUTER GRAPHICS

Computer Graphics is the creation of pictures with the help of a computer. The end product of the computer graphics is a picture it may be a business graph, drawing and engineering. In Computer Graphics, two or three-dimensional pictures can be created that are used for research.

Definition:

Computer Graphics is a sub-field of computer science which studies methods for digitally synthesizing and manipulating visual content. Although the term often refers to the study of three-dimensional graphics, it also encompasses two-dimensional graphics and image processing.

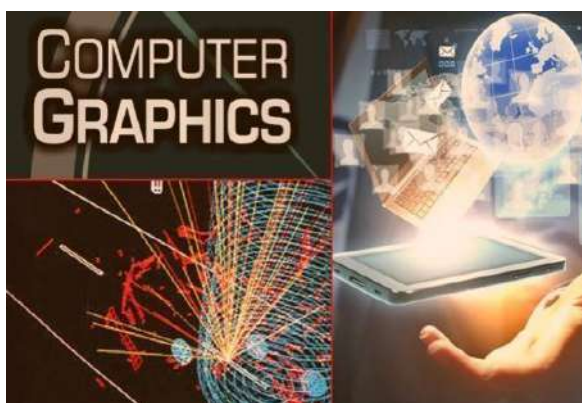


Fig:1.1

Importance:

Interactive Computer Graphics allows the physician to interpret this large volume of data in a new and useful ways. Computer Graphics has also expanded the boundaries of art and entertainment. Movies such as, Jurassic Park make an extensive use of computer graphics to create images that test the bounds of imagination.

Uses:

Computer Graphics is used in the areas like;

- ✓ Cartography.
- ✓ Visualization of measurement data (2D and 3D).
- ✓ Visualization of Computer simulations.
- ✓ Medical diagnostics.
- ✓ Drafting and Computer design.
- ✓ Preparation of publications.
- ✓ Special effects in movies.

Computer Games.

Computer Graphics used in Education such as, it is receiving much attention in the development of interactive educational software, multimedia systems and many other applications. If it is skilfully and relevantly used, it can be an important component of computer-assisted instruction, which is an Educational application area with tremendous potential.

Types of Graphics:

Types of Graphics	Composition	File Formats
Raster Graphics	Pixels	.psd, .bmp, .jpg, .gif, .png, .tiff
Vector Graphics	Points on a Cartesian plane	.ai, .svg, .eps, .pdf

How do computer graphics work?

A Graphics card works along the same principles. The CPU, working in conjunction with software applications, sends information about the image to the graphics card. The graphics card decides how to use the pixels on the screen to create the image. It then sends that information to the monitor through a cable.

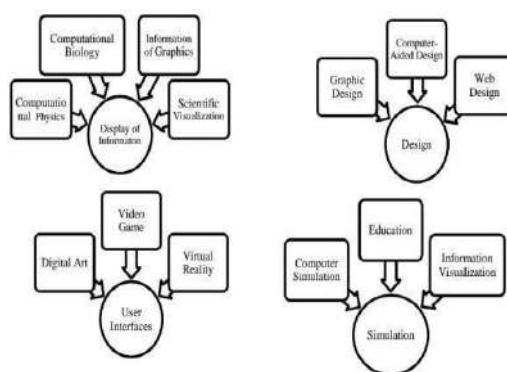


Fig:1.2

Basic elements of computer graphics:

Line, Shape, Textures, Color, Value, Space.

4 Elements:

Modeling- representation choices, geometric processing.

Rendering-geometric transformation, visibility, simulation of light.

Interaction-input/output devices, tools.

Animation-lifelike characters, natural phenomena, their interactions, surrounding environments.

Advantages of Graphics:

7 Ways of Graphic Design benefits your business

Employee pride and productivity.

Saves time.

Saves money.

High-quality visuals increase viewer interactions.

Graphic Communication: it's more than a trend.

Attract and keep customers interested.

Strengthens your brand.

III. COMPUTER VISION

Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs and take actions or make recommendations based on that information.

How does Computer Vision work?

Computer vision needs lots of data. It runs analyses of data over and over until it discerns distinctions and ultimately recognize images. Two essential technologies are used to accomplish this: a type of machine learning called deep learning and a convolutional neural network (CNN).



Fig:1.3

Computer Vision applications:

There is a lot of research being done in the computer vision field, but it's not just research. Real-world applications demonstrate how important computer vision is to endeavors in business, entertainment, transportation, healthcare and everyday life. A key driver for the growth of these applications is the flood of visual information flowing from smart phones, security systems, traffic cameras and other visually instrumented devices.

Google Translate lets users point a smart phone camera at a sign in another language and almost immediately obtain a translation of the sign in their preferred language.

The development of self-driving vehicles relies on computer vision to make sense of the visual input from a car's cameras and other sensors. It's essential to identify other cars, traffic signs, lane markers, pedestrians, bicycles and all of the other visual information encountered on the road.

IBM is applying computer vision technology with partners like Verizon to bring intelligent AI to the edge, and to help automotive manufacturers identify quality defects before a vehicle leaves the factory.



Fig:1.4

System Methods:

The organization of a computer vision system is highly application-dependent. Some systems are stand-alone applications that solve a specific measurement or detection problem, while others constitute a sub-system of a larger design which, for example, also contains sub-systems for control of mechanical actuators, planning, information databases, man-machine interfaces, etc. There are, however, typical functions that are found in many computer vision systems.

Uses:

Enables computers and systems to derive meaningful information from digital images, videos and other visual inputs

Take actions or make recommendations based on that information.

IV.ANIMATION

For the past two decades, the most prominent feature of the technology-based learning environment has become animation.

The most famous animation characters in the history of animation

- ✓ Thaumatrope
- ✓ Phenakistoscope

- ✓ Zoetrope
- ✓ Praxinoscope
- ✓ Kinstoscope (an example of computer animation which is produced from motion capture techniques)

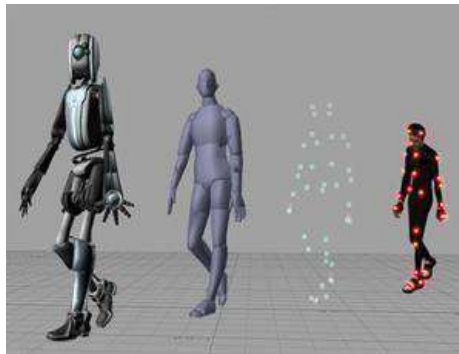


Fig:1.5

When it comes to new forms of animation, firstly let us define traditional animation - a system of animating in which the illusion of movement is presented by photographing a sequence of individual drawings on consecutive frames of film. On the other hand, computer animation is a form of pictorial presentation which refers to simulated motion pictures showing movement of drawn objects.



Fig: 1.6

9 Types of Animation Styles

Traditional / 2D Animation. 2D animation probably doesn't need much of an explanation.

3D Animation. 3D animation was quite revolutionary when it was introduced.

- ✓ Stop Motion Animation.

- ✓ Rotoscope Animation.
- ✓ Motion Capture.
- ✓ Typography Animation.
- ✓ Mechanical Animation.
- ✓ Claymation.

French cartoonist and animator Émile Cohl is often referred to as "the father of the animated cartoon."

Pixelation is a form of stop motion that uses real people and real environments to create unreal videos. It uses the stop motion method of taking a still photo, moving things around, and then taking another photo, but the subject matter is usually real people instead of puppets.

A Good Career Choice :

A career in animation is one of the most lucrative and most-sought-after courses these days. With attractive salaries and the personal freedom it offers, a career in animation could be the right choice for you. Both movies, video games, and other forms of media use computer animation. With movies like Ice Age, Kung Fu Panda, Incredibles 2, and others being extremely popular among audiences, computer animation for is something that is not going away any time soon.

Animations for the web are in extremely high demand, and as a result, there's no shortage of work when it comes to possessing a skill for the craft.

V .GAME SCIENCE

The Master of Science (MS) in Game Science and Design is a program that seeks to give students a comprehensive understanding of how successful game products are created in a player-centric environment. ... The MS in Game Science and Design is a STEM-designated degree

Concept of science used in Gaming:

The premise we entertain is that several classes of video games can be viewed as a type of cultural tool that is capable of supporting three key elements of scientific literacy: content knowledge, process skills, and understanding the nature of science

The B.S. in Computer Game Science is no longer accepting students. Effective Fall 2021, the B.S. in Computer Game Science has been renamed to the B.S. in Game Design and Interactive Media.

The Computer Game Science major gives students a strong foundation in introductory information and computer science, an extensive education in technologies and design practices

associated with computer games, and an opportunity to focus in two areas of particular interest to the student. Students who complete the major will be able to create interactive and human-centered game designs; implement games using skills in modeling, graphics, software engineering, hardware architectures, human interfaces, and aesthetics; and evaluate games and game technology for their use in education, art, and social change.



Fig:1.7

The video game industry is an ever-evolving area, and those interested in the field may ask how computer science impact games do. In today's highly technological world, game developers are constantly looking for new fresh ideas to broaden the range of consumers seeking to purchase and play an array of video games. As computer science has progressed, the video game industry has implemented various computing technology concepts. Computer science has improved graphics, enhanced multiplayer games, enabled cloud-based and on- demand gaming, and availability of virtual and augmented reality.

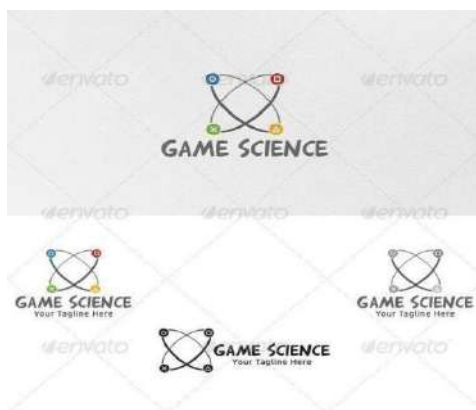


Fig: 1.8

Gaming science: the “Gamification” of scientific thinking.

Game science was founded by Lou Zocchi.

Game science published the board game The Battle of Britain (1968), the war games Mig Killers (1977), and Strike Team Alpha (1978), and the role-playing games Star Patrol (1977; originally called Space Patrol), Superhero: 2044 (1977), the second edition of Empire of the Petal Throne (1984), and TWERPS (1987).

Game science also produces dice, including several types of nonstandard dice

Game science is an American game company that produces role-playing games and game supplements.

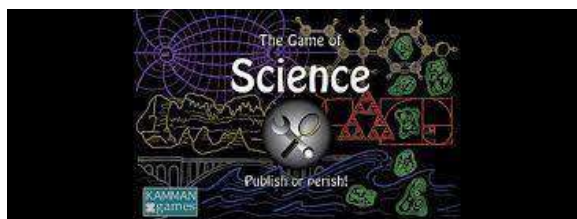


Fig: 1.9

SCOPE:

The scope of graphics and design is very wide.

First it is about making cool designs that will suite people taste, and i want you to know that you can't satisfy every one.

It involve learning How to use graphics programmes to produce nice designs.

And know that not all designs have to be complex you can make a simple design that will worth a thousand Dollar and also make a very complex design and the client might not even like it.

Graphics and design works with passion "if you don't have passion for it you can't gain Pleasure from it"

The present state of computer vision technology

Computer vision technology of today is powered by deep learning algorithms that use a special kind of neural networks, called convolutional neural network (CNN), to make sense of images. These neural networks are trained using thousands of sample images which helps the algorithm understand and break down everything that's contained in an image. These neural networks scan images pixel by pixel, to identify patterns and "memorize" them. It also memorizes the ideal output that it should provide for each input image (in case of supervised learning) or classifies components of images by scanning characteristics such as contours and colors. This memory is then used by the systems as the reference while scanning more images.

And with every iteration, the AI system becomes better at providing the right output. Following are a few areas in which computer vision technology is being used or tested. Job prospects in Animation are increasing at a rapid pace as the entertainment and gaming industry is in its development stages. The film and advertising industry needs the help of animators, and hence, new jobs are being created



Fig: 1.10

The gaming industry has an array of career choices to offer. It makes up for a large segment on the employment front too. According to statistics shared by the American Gaming Association, game industry jobs provide employment to as many as 1.7 million individuals, with the employment rate growing by 62,000 jobs (on average) every year.

FUTURE PLANS:

Computer graphics research and hardware has matured as a field to the point that high-quality computer graphics is becoming ubiquitous. Computer graphics shortly will be where word processing is today: everyone uses it, but there are very few people doing basic research in word processing. All of the challenges lie in the applications and use of this technology to enable advances in many fields. This panel will combine experts in computer graphics and associated technology with experts from a few applications areas to discuss the possibilities and future ways that computer graphics can advance discovery in many fields. With further research on and refinement of the technology, the future of computer vision will see it perform a broader range of functions. Not only will computer vision technologies be easier to train but also be able to discern more from images than they do now. This can also be used in conjunction with other technologies or other subsets of AI to build more potent applications. For instance, image captioning applications can be combined with natural language generation (NLG) to interpret the objects in the surroundings for visually challenged people. Computer vision will also play a vital role in the development of artificial general intelligence (AGI) and artificial superintelligence (ASI) by giving them the ability to process information as well as or

even better than the human visual system. Computer animation has been considered for many years as



Fig: 1.11

a new media for advertisement and special effects in films. More recently, the fast development of powerful super workstations has led to new areas like multimedia, interactive games and Virtual Reality. What's the future of game development? It constantly changes. New game development trends emerge every year. And technology advances, bringing opportunities for game development studios to innovate. There's so much for game development studios to consider if they want to produce their next hit game. That's why we recently surveyed over 500 game development professionals to get their take on the biggest challenges, trends, and opportunities facing the industry today.

V.CONCLUSION

Computer graphics will continue to get more sophisticated. Their 3-D photorealistic capabilities and ability to predict changes over time have revolutionized product development and marketing, as well as scientific research and education. They are responsible for superior special effects in movies and on television. Many newspapers and magazines use only computer-generated graphics. They add an aesthetic and emotional dimension to text. Computer graphics affect everyone's life in almost every aspect every day. Computer vision typically requires a combination of low level image processing to enhance the image quality(eg: remove noise ,increase contrast) And higher level pattern recognition and image understanding to recognize features present in the image. Computer animation has made a impact on entertainment ,movie industries and kids. Computer animation made entertainment more interesting and enjoyable. Huge companies such as Walt Disney and Pixar, were greatly influenced by the invention of computer animation. Also the effects of children from the technology can be both bad and good. Additionally computer animation is not a replacement for traditional animation. just another tool. Finally, although the technology today is advanced,

there's a greater future in store for computer animation. Thinking of the game as a part of a bigger educational process is really in the core mind-set that this project wants to promote. This project aimed as much at using alternative and innovative methods to teach through coding digital games and playing games as part of learning, as at developing the skills of teachers in extending academic goals to understand, support and include the whole child: not only their academic subject skills but also social, emotional and behavioural skills.

VI. REFERENCES

- [1] Abbas, P. (2012). The significant role of multimedia in motivating EFL learners' interest in English language learning, I.J. Modern Education and Computer Science, 4, 57-66.
- [2] Ainsworth, S. (2008). How do animations influence learning? In. D. Robinson & G. Schraw (eds.), Current Perspectives on Cognition, Learning and Instruction: Recent Innovations in Educational Technology that Facilitate Student Learning, pp 37-67, Information Age Publishing, UK.
- [3] Baddeley, A. (1998). Human Memory, Allyn and Bacon, Boston, USA.
- EDDY, A. 2003. On with the show... again.
- [4] <http://www.gamespy.com/bizbuzz/january03/bizbuzz49/> (January)
- [5] FOLEY, J., VAN DAM, A., FEINER, S., AND HUGHES, J. 1990. Computer Graphics: Principles and Practice, 2nd Edition. Addison Wesley.
- [6] FUCHS, H., KEDEM, Z., AND NAYLOR, B. 1980. On visible surface generation by a priori tree structures. In Proceedings of SIGGRAPH, ACM SIGGRAPH / Addison Wesley, Computer Graphics Proceedings, Annual Conference Series, 124–13 arXiv:2109.07855 (cross-list from cs.CV) [pdf, other]
- [7] Evaluating Continual Learning Algorithms by Generating 3D Virtual Environments
Enrico Meloni, Alessandro Betti, Lapo Faggi, Simone Marullo, Matteo Tiezzi, Stefano Melacci

DIGITAL TECHNOLOGY ON PSYCHOLOGY TREATMENTS*T.Shaktippriya¹, K.Sneha², R.Lekshana Priya³**Dept of computer science MKJC,TN, India.***ABSTRACT:**

The psychological treatment of mental health problems is beginning to undergo a sea-change driven by the widespread availability of digital technology. In this paper we provide an overview of the developments to date and those in the pipeline. We describe the various uses of digital interventions and consider their likely impact on clinical practice, clinical services and the global dissemination of psychological treatments. Technological advances can also lead to improved treatments, using virtual reality, pharmacological treatments and psychotherapy. The authors illustrate how technology can further professional development, whether by improving the quality of one's work or transforming research and clinical insights into innovative business practices or product development. We note the importance of online clinics, blended treatment, digital assessment and digital training.

KEY WORDS: Digital technology, Digital health, Psychological treatment, Blended treatment, Training, Dissemination.

I.INTRODUCTION:

The psychological treatment of mental health problems is beginning to undergo a fundamental change. This change is being driven by the widespread availability of “digital technology” by which we mean computers, the internet, mobile devices such as smart phones, and mobile software applications (apps). In this paper we describe the various uses of digital interventions and consider their likely impact on clinical practice, Clinical services and the global dissemination of psychological treatment Technology has thoroughly changed the field of psychology. It has introduced new ways to provide treatment. It has made learning easier. It has affected dissemination. of knowledge and how research is conducted .It has even created new issues that need further exploration. Let’s examine how technology has contributed to the study of psychology.

*Fig:1.1*

II. DIGITAL TECHNOLOGY:

Digital technologies are electronic tools, systems, devices and resources that generate, store or process data. Well known examples include social media, online games, multimedia and mobile phones. Digital learning is any type of learning that uses technology. It can happen across all curriculum learning areas.

*Fig:1.2*

Simple Examples of Digital Technology

Video technologies for working from home.

GPS, What3Words and Location.

5G and Virtual Reality.

Smart homes.

Cloud and SaaS.

Artificial Intelligence (AI)

Deep fake and AI.

Bitcoin and blockchain.

III .DIGITAL HEALTH:

Digital health technologies use computing platforms, connectivity, software, and sensors for health care and related uses. These technologies span a wide range of uses, from applications in general wellness to applications as a medical device. They include technologies

Marudhar Kesari Jain College For women,vaniyambadi

intended for use as a medical product, in a medical product, as companion diagnostics, or as an adjunct to other medical products (devices, drugs, and biologics). They may also be used to develop or study medical products. The broad scope of digital health includes categories such as mobile health, health information technology (IT), wearable devices, telehealth and telemedicine, and personalized medicine.

BENEFITS OF DIGITAL HEALTH:

Digital health has the potential to prevent disease and lower healthcare costs, while helping patients monitor and manage chronic conditions. It can also tailor medicine for individual patients. The result is increased efficiency and improved medical outcomes.

1. Prevention before treatment:

Digital health technologies help patients self-manage their health conditions through regular monitoring and tracking of symptoms. More importantly, it is a tool for the early detection of significant changes to disease progression in a patient, before lung health has been irreversibly compromised. Digital health platforms are therefore of immense value for both respiratory disease patients and those classified as 'at risk'.

2. Re-modelling the patient-doctor relationship:

Digital health systems engage patients with their health care provider, making them co-designers of their care and treatment plans. The quick, direct and shared access to the current health status of the patient increases the sense of partnership, trust, and transparency between patient and doctor. Important factors affecting health such as time of day, environmental stimulants, the use of medications and adherence to medications can all be logged in real-time and can be used to present a clear profile of the sensitivities of the patient's condition.

3. Expanding the reach of health-care professionals:

Digital health innovations aim to reduce health-care professionals' administrative burden and other repetitive aspects of their jobs. This expands their time for actual patient-contact and monitoring. This is particularly important for patients or clinics located in rural areas or for home-care/outpatients for travel is difficult or not recommended. With clinical-grade technologies in their pocket, patients are equipped to provide their health information to their physician at any time.

4. Leveling the platform:

The access to affordable medical technologies in clinics reduces the financial burdens associated with disease management for both the clinic and patients. Many digital health platforms also a gateway to online communities where patients can find encouragement and

engagements with others experiencing similar health issues. Psychological treatment is sometimes called ‘psychotherapy’ or ‘talking therapy’.

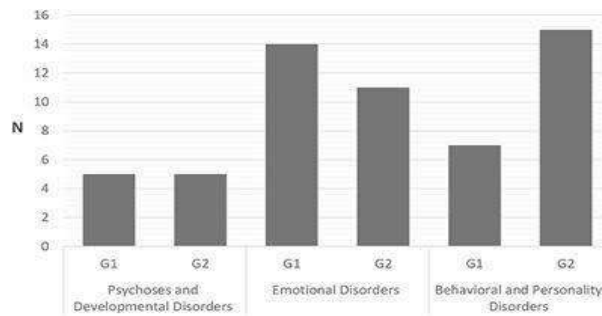
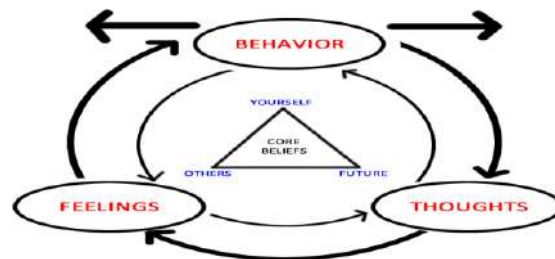


Fig:1.3

It involves talking about your thoughts with a professional to:

Better understand your own thinking and behaviour understand and resolve your problems recognise symptoms of mental illness in yourself reduce your symptoms change your behaviour improve your quality of life. Evidence shows that psychological treatments work well for emotional, mental and behaviour issues. Psychological treatments are useful for people of all ages, including children. They can help people from different cultural, social and language backgrounds. You can have psychological treatment in an individual session, as part of a group, or online.



don't know where to begin, you're not alone. The world of psychotherapy can seem vast and confusing to those new to it. A wide range of therapeutic approaches exists, yet no single type is right for everyone.

COGNITIVE BEHAVIORAL THERAPY (CBT):

Cognitive Behavioral Therapy is a short-term, goal-oriented therapy that focuses on the link between our thoughts (cognition) and our actions (behaviour). It is intended to help people change thought patterns that cause unhealthy, unproductive behaviour.

Psychotherapists often use this approach with people suffering from anxiety, depression, stress, or phobias. It's also popular among people seeking to overcome harmful habits and addictions, such as smoking, overeating, or gambling.

PSYCHODYNAMIC THERAPY:

Psychodynamic therapy emphasizes how certain life events and relationships, both past and present, affect your current feelings, relationships, and choices. Its goal is to help you acknowledge and understand negative feelings and repressed emotions so you can resolve internal psychological conflicts, and improve life experiences, self-esteem, and relationships. This approach is a popular treatment for people who are depressed.

A psychodynamic therapist will encourage you to speak openly about a range of issues to help you to uncover different memories, experiences, or dreams that helped shape your life. In particular, you will explore the reasons why you have taken certain adverse decisions or actions in the past to help you avoid making similar unfavourable choices in the future. You can also use this new understanding of yourself to resolve current problematic situations and enhance relationships.

Psychodynamic therapy may sometimes be an effective short-term therapy, but it often takes a year or longer to obtain enduring benefits.

DIALECTICAL BEHAVIOR THERAPY (DBT):

Dialectical behavior therapy (DBT) treatment is a form of cognitive behavior therapy. Its primary aim is to give people the skills to regulate their emotions, handle stress in a healthy manner, and improve relationships, and live mindfully. Originally developed to treat people with borderline personality disorder, DBT is now used to treat a variety of mental conditions and is believed to be especially helpful for people with seemingly uncontrollable, intense negative emotions or those who may incline toward self-harm.

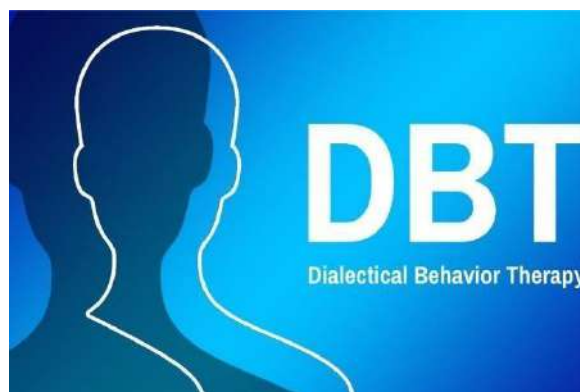


Fig: 1.5

DBT differs from CBT in that it teaches you that your experiences are real and shows how to accept yourself, even with your unique challenges and life experiences. The treatment usually consists of both one-on-one sessions with a psychotherapist and therapist-led group sessions where the participants develop and practice skills and behaviors needed for a more

manageable daily life. In both situations, DBT patients learn how to label emotions, handle angry feelings and navigate conflict without giving into impulsive tendencies, and develop awareness of their feelings during the present moment. Like CBT, DBT also has a homework component. DBT is ideally done both in groups and in individually sessions simultaneously. People generally find this mixture quite helpful.

HUMANISTIC/EXPERIENTIAL THERAPY:

Unlike behavioral-based therapies, humanistic/experiential therapy focuses on a person's individual nature rather than the collection of behaviors that make up a specific psychological category. The therapy's holistic approach emphasizes the whole person, especially their positive behaviours and their ability to grow, heal and find self- actualization through self-exploration. People with depression, anxiety and panic disorders, and low self-esteem often seek this approach.

Humanistic therapy consists of two popular techniques: Gestalt therapy and client-centered therapy. Gestalt therapy helps people to centre on "here and now" feelings and experiences rather than their perception of the root causes of those feelings. Your therapist will help you explore feelings and experiences through creative and experiential techniques, such as guided re-enactments, role-playing, exaggerated movement, and other exercises. The goal is to arouse emotions in different situations, allowing the person being treated to become aware of and understand those emotions as they happen.

Client-or person-centered therapy centers on the idea that people are capable of deciding for themselves the psychological areas they want to explore and know best how to go about it. Known as a "non-directive" form of therapy, the therapist does not guide the client toward any particular direction or outcome but creates a supportive environment for clients as they investigate their identity, feelings, experiences or emotions. You can expect your therapist to listen to your point of view with empathy, warmth, respect, and non-judgment, and to encourage your growth and self-realization.

Sustainable and Responsive Health care :

With an improvement in life expectancy, the number of people living with chronic diseases has been increasing steadily. This has led to a significant rise in the costs of healthcare services. These changes in the healthcare sector have contributed to a growing shift towards digital health approaches for receiving treatments. Groundbreaking technologies are opening a new era of smart care and wearables. We covered Smart Contact Lenses on this article here as a

way patients are taking control of their own health and empowering themselves to a better living.

Digital platforms can help patients receive quicker access to healthcare services thereby improving the quality of treatment they receive as well as the prognosis of their health issues.

These platforms have also helped to reduce the burden on the healthcare professionals as well as facilities like clinics and hospitals by pioneering the concept of self-care by patients themselves.

IV. DIGITAL BLENDED TREATMENT:

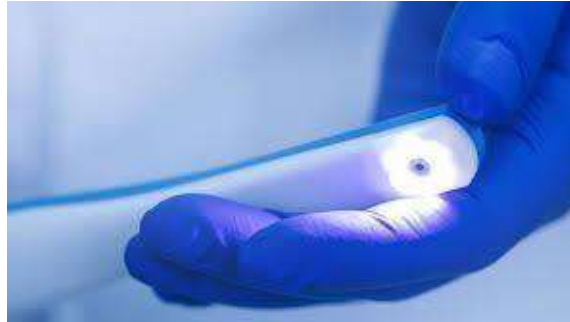
The concept of “blended treatment” is a new one. Generally, it refers to face-to-face treatments which include a digital intervention or component (Wentzel, van der Vaart, Bohlmeijer, & van Gemert-Pijnen, 2016) although the clinician involvement need not be literally face-to-face; for example, it could be via telephone or vide conferencing. Blended treatment is gaining in popularity, a particularly early adopter being the Netherlands (Ruwaard & Kok, 2015)

Blended treatment can take a variety of forms. It can be programme-led as in what may be termed “supervised digital treatment”. This involves a digital programme delivering the intervention (as in autonomous digital treatment and supported digital treatment) while a clinician closely supervises its use and makes additional contributions as needed. These may include communicating regularly with the user (via whatever medium seems suitable), providing advice, controlling the pace at which the intervention is delivered, and deciding if and when further treatment modules should be made available. This role requires that the clinician be intimately familiar with the digital intervention and the treatment upon which it is based. Preliminary evidence suggests that supervised digital treatment may dramatically reduce the amount of clinician time needed without sacrificing effectiveness (Stott et al., 2013; Wild et al., 2016).

More commonly, blended treatment resembles conventional clinician-led treatment. In this case the digital intervention is used to make the treatment more efficient by taking over one or more tasks ordinarily undertaken by the clinician (for example, providing education) or it is used to make the treatment more effective as in the case of virtual reality-based exposure (Freeman et al., 2016, Valmaggia et al., 2016).

The research on digital treatment:

Digital treatment has been the focus of an impressive amount of research given that this is a young field. Many of the website-based interventions have been tested in randomised controlled trials (RCTs) and their findings have been the focus of numerous systematic reviews (Andersson et al., 2014, Richards and Richardson, 2012). Surprisingly, app-based interventions have barely been studied (Donker, et al., 2013b). Here is a brief summary of the main findings.



1. Direct-to-user digital treatments are popular and can access underserved groups. A leading example is Mood GYM, a free online intervention for depression that has been available since 2001 (Christensen, Griffiths, & Korten, 2002) and has been used by over three-quarters of a million people. An important shortcoming of direct-to-user interventions is that completion rates are low if there is no accompanying support. Certain forms of psychopathology may prove to be more amenable to direct-to-user treatment than others. The eating disorders bulimia nervosa and binge eating disorder might be particularly suitable as binge eating is a repeated highly aversive experience which responds well to self-help interventions (Wilson & Zandberg, 2012) yet many sufferers do not seek treatment because of the associated shame and secrecy (Hart, Granillo, Jorm, & Paxton, 2011).

2. Online clinics can obtain clinically relevant change on a large scale. A good example is Mind Spot, an Australian online clinic. In its first year of operation it provided supported digital treatment to over 2000 adults with depression or certain anxiety disorders (Titov et al., 2015). Over 70% completed the treatment with limited external input and their intent-to-treat outcome was similar to that of equivalent face-to-face treatments.

3. Supported digital interventions are more effective than unsupported ones. This is a consistent finding although, depending on the context and the specific intervention, the difference is not necessarily great (Baumeister, Reichler, Munzinger, & Lin, 2014). It is generally thought that the explanation lies in better treatment adherence in the presence of support (Mohr, Cuijpers, & Lehman, 2011).

4. When accompanied by support, digital

Not surprisingly, many important questions have yet to be answered. Here are some examples. First, as there have been few head-to-head comparisons of different digital interventions for the same mental health problem, it is not clear which ones are the most effective ones nor is their relative cost-effectiveness known (Donker et al., 2015). These are significant omissions since this information is needed by policy makers.

Second, it is not known whether the functionality of a digital intervention has a bearing on its effectiveness. This too is an important omission as the answer has major implications for the design of future treatments. It is possible that the age of the user may be relevant with younger users complying better with more interactive treatments. The nature of the psychopathology being addressed may also need to be taken into account when designing interventions; for example, users with depression may struggle to complete interventions which require sustained concentration. In addition, there is a need for research on how these interventions work; who is accessing them; who benefits most; and whether the changes last. Also, more needs to be known about any negative effects of digital treatment (Rozental et al., 2014).

V. TRAINING:

Clinical Psychologist Role and Training:

By Owen Kelly, PhD Updated on April 10, 2020. A clinical psychologist is a mental health professional with highly specialized training in the diagnosis and psychological treatment of mental, behavioral and emotional illnesses, including obsessive-compulsive disorder (OCD).

The Clinical Psychologist's Role:

Clinical psychologists do not prescribe medications to treat mental illness. Rather, they use psychological techniques, such as cognitive-behavioral therapy (CBT) and psychoanalytic therapy.

In addition to the delivery of psychotherapy, psychologists may perform psychological testing or research, or they may teach.

VI DISSEMINATION:

The contribution of digital technology to the global dissemination of psychological treatments: The challenges to disseminating empirically supported psychological treatments are global; to our knowledge, their effective coverage does not exceed 50% in any country. Even this figure is likely to be a significant over-estimate. The coverage in low and middle income countries is

much lower. A recent estimate from India and China, two relatively well resourced middle income countries, revealed a treatment gap exceeding 90% for common mental disorders and alcohol use disorders (Patel et al., 2016), the two mental health conditions for which psychological treatments are recommended as first line interventions by the mhGAP program (WHO, 2008).

There are a number of other challenges which will need to be addressed in the process of dissemination. Foremost amongst these are language (the vast majority of digital innovations are in English); the restricted coverage of internet-enabled devices; limited internet literacy in vast sections of the global population; and the lack of clearly defined regulatory procedures to ensure privacy and confidentiality of digital health data. Notwithstanding these limitations and challenges, we are bullish about the prospect of digital technology being transformative in improving the global availability of psychological treatments. Our enthusiasm is influenced by several factors: the ingenuity with which digital technology is being applied in diverse ways for diverse goals; the demonstrated successes of digital technology in a variety of other health care domains; the rapid growth in internet coverage and internet literacy, in particular among young people who are potentially the most important group for targeting psychological treatments for common mental health problems; and the ever-increasing speed of data access and the reducing cost of internet-enabled devices. .

While the “digital divide” undoubtedly remains a problem, particularly in low-resource settings, the divide is closing and there is no reason to think that this will not continue. Digital interventions that can be used without support are of particular importance as they have enormous potential to improve access, and additionally they have the value of being inherently empowering. They need to be optimised and “task sharing” needs to be expanded to embrace digital self-help. National and international organisations concerned with mental health need to endorse and support digital technologies as they are likely to be transformative. Above all, the international psychological treatment community must strive to engage digital entrepreneurs and innovators, particularly those who are championing initiatives in global health, to partner with them to exploit the many opportunities for using digital technology to transform mental health care worldwide.

SCOPE:

Advances in biotechnology and information technology are poised to transform well-being research. This article reviews the technologies that we predict will have the right most impact on both measurement and intervention in the field of positive psychology over the next decade.

These technologies include: psychopharmacology, non-invasive brain stimulation, virtual reality environments, and big-data methods for large-scale multivariate analysis. Some particularly relevant potential costs and benefits to individual and collective well-being are considered for each technology as well as ethical considerations. As these technologies may substantially enhance the capacity of psychologists to intervene on and measure well-being, now is the time to discuss the potential promise and pitfalls of these technologies.

Future Use:

The use of electronic and communication technologies as a therapeutic aid to healthcare practices is commonly referred to as telemedicine or eHealth. The use of such technologies as a supplement to mainstream therapies for mental disorders is an emerging mental health treatment field which, it is argued, could improve the accessibility, effectiveness and affordability of mental health care. Mental health technologies used by professionals as an adjunct to mainstream clinical practices include email, SMS, virtual reality, computer programs, blogs, social networks, the telephone, video conferencing, computer games, instant messaging and podcasts.

There is uncertainty around the ethical and legal implications of digital technologies in the mental health context, including the use of artificial intelligence, machine learning, deep learning, and other forms of automation. Ethical and legal issues tend to not be explicitly addressed in empirical studies on algorithmic and data-driven technologies in mental health initiatives. Concerns have been raised about the near-complete lack of involvement of mental health service users, the scant consideration of algorithmic accountability, and the potential for overmedicalization and techno-solutionism

CONCLUSION :

To conclude now that I have finished the module and have completed the tasks I feel that I have learnt an incredible amount of new skills and have deepened my knowledge of digital technology. After reflecting on how much my skills have improved I feel that I am now able to use a lot more sites and apps that I would not have been able to use before I started the Digital Technologies module. Technology is extremely important in education as said in the Curriculum for Excellence “The technologies provide frequent opportunities for active learning in creative and work-related contexts.” (Scottish Government, 2004) I feel that this shows how technology can positively impact the classroom as it can make turn simple lessons, into interactive ones which will be enjoyable be to all pupils. As a future educator I now have a

large amount of technology skills and knowledge that I will be able to teach children and I will be able to use it to enrich the pupil's learning.

REFERENCE:

- 1.<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5214969/>
- 2.https://www.researchgate.net/publication/312575121_The_impact_of_digital_technology_on_psychological_treatments_and_their_dissemination
- 3.https://link.springer.com/chapter/10.1007/978-3-540-77006-0_3
- 4.https://www.researchgate.net/publication/249363078_Economic_paper_blending_optimization_model_with_competing_materials
5. <https://www.mdpi.com/1660-4601/17/1/153>
- 6.<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5214969/>
- 7.<https://link.springer.com/article/10.1007/s41811-020-00088-1>
- 8.https://www.googleadservices.com/pagead/aclk?sa=L&ai=DChcSEwIU2uPN3I3zAhWGMyoKHZglAyUYABAAGgJ0bQ&ae=2&ohost=www.google.com&cid=CAASEuRo6q6WOPpkccIIItRZ_Lh_cA&sig=AOD64_0oq8Z7_FpTu95to0KJ9qzP7VOyw&q&nis=1&adurl&ved=2ahUKEwjH4dvN3I3zAhXqyTgGHYP0Cr4Q0Qx6BAgGEAE
- 9.<https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190218058.001.0001/oxfordhb-9780190218058-e-13>
- 10.<https://www.online-psychology-degrees.org/study/technology-changing-psychology/>

ARTIFICIAL INTELLIGENCE TRENDS IN DEEP LEARNING ANALYTICS

*Mrs. S. Sowmiyasree¹, Mrs.V.Deepa²
Assistant Professor ,Dept of Computer Science
Vivekanandha College for Women.Tamil Nadu, India*

Abstract.

Deep Brain Chain is a decentralised, low-cost and privacy-protecting AI computing platform that uses block chain technology. It is basically a decentralised neural network According to the developers, the project aims to modify from the present NEP-5 chain to a native substrate token with decentralised governance and deliver a cloud computing in AI environment DeepBrain Chain is additionally a secure data trading platform which maximises the worth of knowledge while ensuring data privacy by separating data ownership from data usage.

Keywords :- Block chaining, Deep chain, Crypto currency, Ethereum, Bitcoin

1.INTRODUCTION.

Deep Brain Chain (DBC) is a blockchain-based computing platform for AI that was designed to lower user costs for processing power. This platform operates as a decentralized neural network and it is a Artificial Intelligence Computing Platform driven by blockchain. This DBC project is for global AI computing resource sharing and resource scheduling because many small businesses do not money to buy expensive GPU servers, but many companies offered a large number of GPU servers which are idle. Scheduling global resources and increasing the utilization efficiency of resources are of positive significance regarding the AI business computing costs reduction. Its vision aims at providing a decentralized AI Computing platform, which is low cost, private, flexible, and safe, It serves the interests of several parties.

Firstly, the Miner's main income is rewarded with token from mining.

Secondly, AI companies just pay small amounts to run.

Thirdly, the Chain uses the smart contract to separate the data provider and data trainer physically. Thus, it protects the data of the provider. The informat upon need to utilize the blockchain innovation, at that point, it is hard to transform it A blockchain is a product like a convention for the email framework. Be that as it may, this is additionally called a meta innovation since it can't be actualized without a web association as it influences different advancements. An information base application work in innovation and it works with an

assortment of parts associated with the Internet,[1] now and then in bitcoin, blockchain, and etherium blockchain. Numerous individuals are distrustful and now and again allude to it as other virtual monetary forms or electronic licenses, nonetheless, they all allude to the conveyed booklet This Blockchain isn't an e-cash, yet the innovation behind the blockchain e-money is (810001). Additionally, e-money is an electronic permit. Monitor who holds these passes. When all is said in done, without blockchain bitcoin are not utilized however without bitcoin, we are not used to this innovation.

2. LITERATURE REVIEW.

This research consists of scholarly research papers from reputed journals, conferences and books which are consisting of 40 well known resources top world universities journals by IEEE, ACM, Elsevier and Science Direct. The purpose of this literature review is to spot the prevailing tools, approaches, methodologies and truly big data and blocks chains which are consistently working in industry. In literature review we observed its latest technological trends and researchers and industrialists who are simultaneously working on it and improving this technology based on the existing shortcomings. Blockchain is perhaps the most advanced money globe-trotter organization development and exchange supporting through Bitcoin, etherium as of now, if the money is transported off an individual, the significant bank will charge a trade cost of 0.2 percent from the sender or recipient's record. To overcome the issue, we separate and improve the Blockchain pack containing diverse information times for its benefit, the money exchange, the trading of assets for execution.

3. BLOCKCHAIN TECHNOLOGIES.

Blockchain Pathways:

A chain set Transfer quantifies The first to make burrow for himself Starting. The pack thusly dispatched is being checked by countless center points ie. Web million [2]

this information for some, PCs is offered after its. By them, Lonely Information is being constrained by billions of PCs on the Sovereign PC. Anyone can get to thus not having a spot It is decentralized and it is managed by everyone and can be murdered.

It is encoded (mixed) and set aside. The blockchain stays unaltered

The checked pack is a group (59%), so the data in it will be modified when anyone in the North ponders the chain saving. Therefore, it isn't simply novel chronicle of pollution yet furthermore a record with an alternate tail Forgery of any chronicle or making it pointless. is worth an

enormous number of rupees .There is a whole chain lying or is to make events and groupings. So this industry makes or changes to destroy or it can't be deformed.

Blockchain Techniques

This Blockchain advancement is an astounding force to be reckoned with so any person who needs data can similarly followed.

- Widespread ·
- Unchangeable ·
- Exposed

All of the three properties are the fundamental segments

As of now, our aggregate [2] will move and reliant on bank rules and rules made in a united laborer PC through a client PC. For this whole trade, they may any issue glanced in the bank, for this time bank measure for all trades not managing that time.

Considering google's concentrated laborer PC, the customers improving client PCs worked under these trades. This is a concentrated model. Since this blockchain development not under a lone specialist PC to get to instead of countless PCs worked this cooperation at a time. This infers that an individual can eagerly exchange information with their friend through the Blockchain, unlike the association's frail association like WhatsApp, Facebook, and Twitter ETC..,

Transparency.

For instance: If u are sending one An individual To B individual send through BTC ie. (Bitcoin) No one comprehends to encode and send like this code

Ex:1 MF1bhsFLKBzzz9vpC23u6

The information just encoded anyway they got individual gets the information in sensible idea.

Immutability.

Any change of this web will absolutely change the encryption so it will not be modified [3][4] or dispensed with by some other individual. These three are that have been seen so far above. The real segments unequivocally guarantee these ideas. For example, the people who stay in school quarters are inclined to record their money moves reliably. Exactly when someone forms the name An instead of B so we instructed as wrongly made that will be recognized at this point in this new Blockchain development prevalently include through all

works under the implementations technique of hash encode coding methodology of the environment. Right when we have a couple of corrections made, the SMS sent off all the customer.

Important Activities of Blockchain:

Blockchain advancement diverse trade-related [5] Information about the diary and spreading

- Each trade make one hash work
- One hash is described as numbers, strings and words consolidate.
- Each trade has delegated demands and that each solicitation will monitor everything
- Each hash implied the particular trade, yet it works the previous trade
- Every trade in little corrections moreover rolls out an improvement new hash code.

Structure of the Blockchain technology.

In Blockchain, the square is the variety of real trades. In a blockchain structure, any center point can start a trade and broadcast to all centers present in the association. Association centers affirm the trade using the old trades, when the trade is endorsed following stage is added to the current blockchain. The quantity of trades that happened in regards to that time span is accumulated as an element of the square and subsequently, set aside into the blockchain block. In Bitcoin [5][6]. A square may contain more than 500 trades in general, the ordinary size of a square is around 1 MB. It may grow up to 8 MB or sometimes higher (as of March 2018). Greater squares can help in planning incalculable trades in one go. The detail of the blockchain structure is depicted in Fig. 3.3 Each piece of the Figure depicted underneath.

Two sections: Block Header and List of Transactions Block header includes Metadata about a square.

- Previous square hash in each square gets from the past block. Blockchain structures use past squares' hash to make the hash of the new square it makes the blockchain deliberately planned. Mining bits of knowledge used to assemble the square: The instrument [6] ought to be adequately be wildered, to make the blockchain fixed bitcoin Mining:

Bit coin:

Bitcoin made progressed trades possible without a "trusted in the representative." The development allowed this to happen at scale, all around the planet, with cryptography doing

what establishments like business banks, financial regulators, and public banks used to do: affirm the legitimacy[7][8] of trades and secure the trustworthiness of the crucial asset. Bitcoin has also been used as an investment of amount in the various places although several procedure oriented regulatory agencies have issued investor alerts about bitcoin. Bitcoin was only a little better than holding conventional cash, but that the slight difference made it a better asset to hold the money. It was announced in September 2020, that they will start to accept tax payments in bitcoin from February 2021 at foreign countries like Switzerland and etc...

Bitcoin is a decentralized, freely available report. There is no accepted untouchable controlling the record. Anyone with bitcoin can participate in the association, send and get [9][10] bitcoin, and even hold a copy of this record if they need to. In that sense, the record is "trustless" and direct.

Ethereum

Ethereum is a decentralized, open-source blockchain including quick arrangement convenience. Ether is the neighborhood cryptographic cash off the stage. It is the second-greatest cryptographic cash by market capitalization, after Bitcoin. Ethereum is the most adequately used blockchain.

Cryptocurrency

Computerized cash, computerized cash, or crypto is a high-level asset expected to work as a method of exchange wherein particular coin ownership records are taken care of in a record existing in a kind of automated database using strong cryptography to get trade records, to control the creation of coins.

4. BLOCKCHAIN VS DEEP BRAINTECHNOLOGY.

According to an end of year report by DeepBrain Chain, the project grew substantially in network users, GPU providers, and geographic reach in 2020. The machine learning and distributed cloud storage provider also noted developmental progress and other achievements. DeepBrain Chain uses blockchain to compensate GPU providers for supplying AI computing and storage nodes to its decentralized global network. It was the third project to conduct a public token sale on the Neo blockchain, doing so in Dec. 2017. The DBC token has multiple utilities, like purchasing computing power and rewarding network miners..DeepBrain Chain's computing network grew exponentially in 2020. In Jan. 2020, there have been but 100 GPUs, growing to 2,000 by December. The DeepBrain Chain network of GPU processors expanded

across Asia, North America, Europe, and Australia and witnessed a quite 90% utilization[19] rate in 2020. Lastly, DeepBrain Chain offered support for about 30 GPU cloud platforms that utilize the AI computing network. The platforms function portals for users to rent or lease GPU space and receive DBC tokens for settlement.

In recent years, Blockchain has been heralded as a major disrupter on track to rattle almost every industry from agriculture and manufacturing to FS and insurance. 2019 revealed the emerging shared belief that blockchain is real and that it can serve as a pragmatic solution to business problems. Even leaders, initially [19][20] wary of tech-based solutions, appeared to see potential in the technology. In 2020, the global blockchain survey improved much better in this technology.

5. EXISTING BLOCKCHAIN.

Interoperable blockchain can support data decency while better-getting patients' modernized characters. Smart arrangements can be made to fill in as the entryway to store standardized information. Immaturity of the Technology. Blockchain is a new technology, represents a complete shift to a decentralized network and might lead to organizational transformation, including changes in strategy, structure, process, and culture.

Blockchain enables [11] PCOR and Precision Medicine pieces of information

Cost issues. Blockchain Technology has initial costs to pay. The users have to pay for the transactions and computational power

Immaturity of the Technology.

Blockchain is a decentralized network represents a complete shift and lead to organizational transformation, including changes in strategy, structure, process the results and finding the culture.

Latency issues.

Time factor is most critical issues in implementations, since not appropriate for massive transactions, due to complexity of the problem and verification process of them.

Data malleability issues: Data malleability is a potential issue in the Blockchain implementation. The signatures do not provide guarantee of the ownership. An attacker can modify and rebroadcasts a transaction which can cause problems in transaction confirmation.

6. PROPOSED METHODOLOGY

Improved precision by dispensing with human consideration [12] in the affirmation Cost declines by shedding pariah affirmation. Decentralization makes it harder. Transactions are more secure, private, and capable transparent of advancement and gives a monetary other choice and way to deal with get singular information for occupants of countries with shaky or youthful government's data integrity and immutability: Participants can reduce fraud while strengthening regulatory compliance. Once a record has been stored in the ledger, it can only be deleted after a consensus. All transactions will be digitally time-stamped with a cryptographic hash code, a unique 64-digit alpha-numeric signature is recorded corresponding to every single transaction High availability and Accessibility: Due to decentralized networks, blockchain Technology data would be complete, timely and accurate Reliability: blockchain Technology it is not regulated by a single control center and here's no single point of failure

7. RESULTS AND DISCUSSION.

The critical perspective to grasp about the wire of public-key cryptography in advanced cash systems, for instance, Bitcoin infers that the mathematical limits that set up open key cryptography are by and large decided in one [13][14] course and the private key contains alphanumerical characters that give a customer access and control over their resources for their relating computerized cash address. The private key is used to sign trades that grant the customer to spend their resources.

8. FUTURE ENHANCEMENT.

Blockchain and bitcoin are genuinely hard for people who are not worked with development and programming headway. So something to accompany redesigns is to develop mechanical assemblies to simplify trades.. Taking care of data in the blockchain is exorbitant so make a [15][16] answer for store the data off the tie and send them blockchain techniques are involved some times. Make laws to accept blockchain development for the business is significant for the change of the monetary

9. REFERENCES.

- [1] Jong-Hyouk Lee, Sejong University, Republic of Korea, jonghyouk (at) sejong.ac.kr
- [2] Swan M. Blockchain: Blueprint for a New Economy. O'Reilly Media, Inc., 2015.
- [3] Zhao, J. L., Fan, S. & Yan, J. Overview of business innovations and research opportunities in blockchain and introduction to the special issue, Financial Innovation, 2016, 2:28, 2-7.
- [4] Iansiti, M. & Lakhani, K. R. The Truth About Blockchain, Harvard Business Review, 2017 January-February, 1-11
- [4] Swan M Blockchain: Blueprint for a New Economy[M]O'Reilly Media, Inc. (2015)
- [5] uan Wang, Development Status and Prospects of Blockchain Technology [J] Journal of Automation, 42 (4) (2016), pp. 481-494
- [6] Guo Y, Liang C ,Blockchain application and outlook in the banking industry[J]Financial Innovation, 2 (1) (2016), p. 24
- [7] Yermack D, Is Bitcoin a real currency An economic appraisal[M],Handbook of digital currency. (2015), pp. 31-43
- [8] Reid F, Harrigan M, An analysis of anonymity in the bitcoin system[M]
- [9] Grinberg RBitcoin: An innovative alternative digital currency[J],Hastings Sci. & Tech. LJ, 4 (159) (2012)
- [10] S.Nakamoto. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. [Online]. Available: <https://bitcoin.org/bitcoin.pdf>
- [11] H.jia.S.Ding.X. Xu. "The latest research progress on spectral clustering" DOI 10.1007/s00521-s013-1439-2, Springer-Verlag, 2013
- [12] J.A. Dev , Bitcoin mining aceleration and performance quantification, in: Proceedings of the 2014 IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE), IEEE, 2014, pp. 1–6 .
- [13] T. Aste, P. Tasca, T. Di Matteo, Blockchain technologies: the foreseeable impact on society and industry, Computer 50 (9) (2017) 18–28.
- [14] Z.Zheng, S. Xie, H. Dai, X. Chen, H. Wang, An overview of blockchain technology: architecture, consensus, and future trends, in: 2017 IEEE International Congress on Big Data (BigData Congress), 2017, p. 557564.
- [15] S.Nakamoto, Bitcoin: a peer-to-peer electronic cash system, 2008.Www.Bitcoin.Org.<https://bitcoin.org/bitcoin.pdf>[Online]. Available:
- [16] G. Wood, Ethereum: a Secure Decentralized Generalized Transaction Ledger YellowPaper, Ethereum Project. Yellow Pap., 2014, p. 132.

- [17] V. Buterin, A Next-Generation Smart Contract and Decentralized Application Platform, Ethereum, 2014
- [18] R.C. Merkle, "Protocols for public key cryptosystems," In Proc. 1980 Symposium on Security and Privacy, IEEE Computer Society, pages 122-133, April 1980.
- [19] Fahn S, Elton R. (1987) Unified Parkinson's Disease Rating Scale (UPDRS). In: Fahn S, Marsden CD, Calne D, Goldstein M. (eds) Recent Developments in Parkinson's Disease. Florham Park, NJ: Macmillan Health Care Information, pp. 153–163
- [20] DeLong MR. (1990) Primate models of movement disorders of basal ganglia origin. Trends in Neurosciences 13(7): 281–285

INTERNET OF THINGS (IOT) BASED NETWORK SECURITY, CYBER SECURITY & DATA SECURITY

*A.Indhumathi¹, R.Ayesha Riyaz², B.Pushpavalli³
Dept of Computer Application ,
Marudhar Kesari Jain College for Women ,TN, India.*

ABSTRACT

Along with the growing threat of cyberattacks, cybersecurity has become one of the most important areas of the Internet of Things (IoT). The purpose of IoT cybersecurity is to reduce cybersecurity danger for organizations and users through the protection of IoT assets and privacy. This paper presents a four-layer IoT cyber risk management framework. This paper also applies a linear programming method for the allotment of financial resources to multiple IoT cybersecurity projects. setup existing data security solutions to the Internet of Things (IoT) is not straightforward because of device heterogeneity, highly dynamic and possibly unsafe environments, and large scale.

Introduction:

The Internet of Things (IoT) has created a new model in which a network of machines and devices capable of communicating and collaborating with each other are driving new process variation in enterprises. Network security is a broad term that covers a introduction of technologies, devices and processes. In its simplest term, it is a set of rules and configurations designed to protect the integrity, confidentiality and accessibility of computer networks and data using both software and hardware technologies. Prevalent and ever-increasing cybersecurity attacks to IoT systems have caused people and organizations a wide range of issues in opinion, compliance, finance, and business operations. The Internet of Things (IoT) has created a new paradigm in which a network of machines and devices capable of communicating and collaborating with each other are driving new process innovations in enterprises. Pervasive and ever-increasing cybersecurity attacks to IoT systems have caused people and organizations a wide range of issues in reputation, compliance, finance, and business operations.

What is IOT SECURITY:

IoT security refers to the methods of safeguard used to secure internet-connected or network-based devices. The term IoT is especially broad, and with the technology continuing to evolve, the term has only become broader. From watches to thermostats to video game encourage

nearly every technological device has the ability to interact with the internet, or other devices, in some capacity.

Network security



Networks provide a huge moment for threat actors to remotely control others' IoT devices. Because networks involve both digital and physical components, on-layout IoT security should address both types of access points. Protecting an [HYPERLINK "https://searchsecurity.techtarget.com/tip/Use-an-IoT-security-architecture-to-protect-networks-end-to-end"](https://searchsecurity.techtarget.com/tip/Use-an-IoT-security-architecture-to-protect-networks-end-to-end) includes ensuring port security, disabling port forwarding and never opening ports when not needed; using antimalware, firewalls and intrusion detection systems/incursion prevention systems; blocking unauthorized IP (Internet Protocol) addresses; and ensuring systems are patched and up to date. Hardware, software and connectivity will all need to be secure for IoT objects to work effectively. Without security for IoT, any connected object, from refrigerators to manufacturing bots, can be hacked. Once hackers gain control, they can usurp the object's functionality and steal the user's digital data.

Cybersecurity in IoT Architecture:



1. Cybersecurity at the Perception Layer
2. Cybersecurity at the Network Layer
3. Cybersecurity at the Processing Layer
4. Cybersecurity at the Application Layer
5. Cybersecurity at the Service Management Layer

>>Cybersecurity at the Service Management Layer:

There is a paucity of studies on IoT cybersecurity risk management in savants and in the industry. Therefore, this literature review is not limited to cybersecurity risk management in the IoT. Previous studies are broadly categorized into approximate and quantitative approaches to cybersecurity risk management.

The CMMI Institute's Cybermaturity Platform shares a similar risk management approach with OCTAVE. However, one unique feature of the Cybermaturity Platform is a risk-based roadmap, which is a custom-made to order list of action items prioritized based on the risks most applicable to the organization.

>>IoT Cybersecurity Technology Developers:

Many IoT managers do not have satisfactory expertise to develop IoT cybersecurity solutions. Hence, IoT managers need to be aware of the trends of IoT cyber technology **developers'** communities. As the IoT market grows, so does the market for the IoT cybersecurity technology developers. Cybersecurity technology developers are provide IoT devices with the latest tools in order to secure the transfer of data, prevent hacking, and keep isolation standards. New technology developments such as 5G, serverless, and edge/fog computing IoT systems should be able to provide better protection against various security cyberattacks.

>> Cybersecurity Technologies:

The cybersecurity technologies at this layer refer to internal cyber technology assets. The IoT cybersecurity technology developers at the IoT cyber ecosystem are highly relevant and may help the organization find and authorize strong IoT cybersecurity automation. private cybersecurity technologies should support the organization's overall cybersecurity goals as well as the IoT.

>>cybersecurity goals:***IoT Cyber Performance Layer:**

Once a cyber solution is identified and a resource allocation selection is made at the IoT cyber assessment layer, the IoT cyber performance movement kick in. While this layer is important for the whole of risk management, the discussion will be brief, as the activities mostly follow the decisions made at the risk assessment layer. The three major activities at the IoT cyber performance layer are utilization, monitoring and control, and continuous improvement.

>>Implementation:

The implementation of the new IoT cyber framework includes IoT cyber technology development, testing, deployment, new policy development, training, and user acceptance. Organizations need to develop selection criteria to evaluate and choose among commercially available cybersecurity technologies analyze in the IoT cyber ecosystem layer. The implementation must take into account the ease, benefit, and usefulness of cyber monitoring and control systems.

While commercially available cybersecurity platforms facilitate the development of monitoring and control systems, these platforms can limit the developer in terms of the solution.

DATA SECURITY in IOT**Definition:**

Data security is the practice of protecting digital information from unauthorized access, corruption, or theft throughout its entire lifecycle. IoT systems are at high security risks for certain reasons. They do not have well defined perimeters, are highly dynamic, and continuously change because of mobility. In addition IoT systems are highly different with respect to communication medium and protocols, platforms, and devices. IoT systems may also include “objects” not designed to be connected to the Internet.

**Types of Data Security:**

Access Controls. This type of data security measures includes limiting both physical and digital access to critical systems and data. ...

Authentication. ...

Backups & Recovery. ...

Data Erasure. ...

Data Masking. ...

Data Resiliency. ...

Encryption. ...

Data Auditing

RESEARCH DIRECTIONS:

Developing complete security and privacy solutions for IoT requires revisiting almost all security techniques we may think of. Encryption protocols need to be engineered so to be efficient and scalable for deployment on large-scale IoT systems and devices with limited computational resources. Benchmarks are needed to perform detailed computation of such protocols.

Software running on the devices must also be secured. Major challenges here arise from the fact that many IoT devices are based on processors such the ARM processor, which have differences in the instruction sets with respect to other conventionally used processors. Such diversity has an conclusion for example on the techniques for protecting software from attacks, such as return-oriented programming attacks, as such techniques must be tailored to the specific information set of the platform of interest.

IoT security issues:

The more ways for devices to be able to connect to each other, the more ways threat actors can obstruct them. Protocols like HTTP (Hypertext Transfer Protocol) and API are just a few of the channels that IoT devices rely on that hackers can intercept. The IoT umbrella doesn't strictly include internet-based devices either. Appliances that use Bluetooth technology also count as IoT devices and, therefore, require IoT security. Oversights like this have contributed to the recent spike in IoT-related data breaches.

Below are a few of the IoT security challenges that continue to threaten the financial split of both individuals and organizations.

CONCLUSION:

The IoT has been a foundational component for smart cities, smart network, smart manufacturing, smart health, driverless cars, and idler, to name a few. As a growing number and variety of connected devices are introduced into the IoT networks, the potential security presentation grow exponentially. A lack of security in the IoT systems opens up opportunities for intruders and hackers to access critical infrastructure and sensitive data. The main emphase of this paper was to highlight major security issues of IoT particularly, focusing the security attacks and their countermeasures. Due to lack of security mechanism in IoT devices, many IoT devices become soft targets and even this is not in the victim's knowledge of being infected. In this paper, the security requirements are discussed such as confidentiality, integrity, and authentication, etc.

REFERENCE:

- 1.J. S. Kumar and D. R. Patel, "A survey on internet of things: Security and privacy issues," International Journal of Computer Applications, vol. 90, no. 11, 2014.
- 2.Hejazi, D.; Liu, S.; Farnoosh, A.; Ostadabbas, S.; Kar, S. Development of use-specific high-performance cyber-nanomaterial optical detectors by effective choice of machine learning algorithms. Mach. Learn. Sci. Technol. 2020, 1, 025007.

A STUDY ON DIGITAL MARKETING AND ITS IMPACTS ON CONSUMER BUYING BEHAVIOUR

*Ms.S. Ranjitha¹, Ms.K.Divya²
Assistant Professor, Dept of BBA
Marudhar Kesari Jain College for Women,Tamil Nadu, India.*

ABSTRACT OF THE STUDY

The world is shifting from analog to digital and marketing is no exception. As technology development is increasing, the use of digital marketing, social media marketing, search engine marketing is also increasing. Internet users are increasing rapidly and digital marketing has profited the most because it mainly depends on the internet. Consumer's buying behavior is changing and they are more inclined towards digital marketing rather than traditional marketing. The purpose of this review paper is to study the impact of digital marketing and how important it is for both consumers and marketers. This paper begins with an introduction of digital marketing and then it highlights the mediums of digital marketing, the difference between traditional and digital marketing, and the pros, cons, and importance of digital marketing in today's era.

Keywords—digital marketing, internet, online advertising, internet marketing

INTRODUCTION OF THE STUDY

Marketing refers to the steps that the company takes to promote the buying of any products or services. The company seeks customers or consumers for their products or services via the help of marketing. Digital Marketing refers to the marketing of any product or service in digital form. For example, marketing using smart phones, computers, laptops, tablets, or any other digital devices. Digital marketing is a form of direct marketing that links consumers with sellers electronically using interactive technologies like emails, websites, online forums and newsgroups, interactive television, mobile communications etcetera.

'Digital marketing' term was first coined in the 1990s. Digital marketing is also known as 'online marketing', 'internet marketing', or 'web marketing'. It is known as 'internet marketing' because with the rise of the internet there is also high growth of digital marketing. The major advantage of digital marketing is that marketers can sell their products or services 24

hours and 365 days, lower cost, efficiency gain, to motivate the customer for more purchase and improve customer services. It helps many-to-many communications because of its excessive degree of connectivity and is generally completed to sell services or products in a timely, relevant, non-public, and cost-powerful manner.

In 2005, there were around 1.1 billion internet users which consist of 16.6 percent of the population at that time. In 2020, the number of internet users is around 4.8 billion and the percentage population has increased to 62 percentages. And there is a direct connection between digital marketing and the internet. Countries like India and China have the highest number of internet users so they have a great opportunity.

VARIOUS CHANNELS OF DIGITAL MARKETING

Digital marketing consists of various channels which are medium used by the marketer to promote their products or services. As an advertiser one, the main aim is to select the channel which is best for communication and give maximum return on investment (ROI).

The list of important digital marketing channels is given below:

A. Social Media

In the current era, social media marketing is one of the most important media in digital marketing. It is the fastest-growing digital channel. Social media marketing is the process of gaining traffic or sites through social media sites. According to Neil Patel, "Social media marketing is the process of creating content that you have tailored to the context of each social media platform to drive user engagement and sharing". The number of internet users by population has increased from 16.6 to 62 percent in 15 years and social media marketing has benefited the most in that.

1) Facebook: It is the number one social media platform. A company can promote their product and services on Facebook.

2) LinkedIn: Professional write their profiles on LinkedIn and can share with others. The company also build their profile and LinkedIn connect these two dots companies and professionals.

3) Google+: It is Google's social network, user can easily connect based on their common interest and friendship.

4) Twitter: Its strategy is to increase brand awareness and sales, attract new followers, and lead and boost conversions.

5) Pinterest: It is a social media platform in which visual content is available and the user can share or store with others.

B. Email Marketing

When a message is sent through email about any product or service to any potential customer it is known as email marketing. It is a simple digital marketing channel to understand. Email marketing is used to sell a product using discounts and event ads, increase brand awareness, and direct people to their business websites

C. Affiliate Marketing

In affiliate marketing, the company rewards subsidiaries for every customer or visitor they bring to the company's website by their marketing efforts or strategy on behalf of the company. According to Pat Flynn's Smart Passive Income, "Affiliate marketing is the process of earning a commission by promoting other people's (or company's) products. You find a product you like, promote it to others, and earn a piece of the profit for each sale that you make". There are 4 different parties involved in affiliate marketing:

1) The Merchant: Sometimes it can be the seller, the brand, or retailer. This party produced a product to sell. It can be an individual or startup or big fortune company.

2) The Affiliate: This party is also known as a publisher. . It also can be an individual or startup or big fortune company. They take a commission from the merchant for every service or product they sell. The affiliate brings customers to the merchant.

3) The Customer: The customer or consumer is an important part of the whole system. They go to affiliates and affiliate redirect them to merchants by taking his commission. Without a customer, the affiliate cannot earn a commission.

4) The Network: Network works as an intermediate between affiliate and merchant. Affiliates require a network to promote products or services.

D. Search Engine Marketing

A search engine is a web based tool that helps the user to find the information they are looking for. Examples of a search engine are Google, Yahoo, Bing, Baidu, etc. Search engine marketing refers to any activity that increases a user's websites rank in any search engine. There is two types of Search engine marketing search engine optimization (SEO) and paid search.

According to Neil Patel, Search engine optimization is the art of ranking high on a search engine in the unpaid section. It is also known as organic marketing or organic listing. In general, the higher the rank of the webpage in the search engine more visitors will visit that webpage.

In paid search one need to pay to get a higher rank in search engine. In paid search, one will have the same kind of keywords as on their organic ad campaign. The majority of a paid search engine is run on a commercial search engine such as Google, Yahoo, Bing, etc. Paid search work on the pay-per-click model, in which marketers will only pay when someone clicks on their ad. The search engine algorithm will determine the rank of the advertiser's ad based on their bid and quality score. Many advertisers prefer paid search rather than SEO in short term due to its ability to give a faster result.

E. Online Display Advertising

In traditional marketing, there is a poster or billboard of any company on both sides of the road or an ad in a magazine/newspaper to promote their product or service. Online display advertising is a digital version of that. Today, a marketer can use online display advertising to achieve the same thing. There are different types of display advertising such as video ads, banner ads, interactive ads, and rich media, etc. Display advertising is great for catching the eye due to graphic ads.

An online display advertising marketer can target an audience based on website content, geography, gender, age, device type, etc. So the marketer can show a suitable ad to the relevant customer which helps in decreasing the budget and increasing sales.

COMPARISON BETWEEN TRADITIONAL AND DIGITAL MARKETING

Customary promoting is the most conspicuous type of showcasing. The vast majority are utilized to conventional advertising because of its life span. A few instances of customary promoting incorporate unmistakable things like advertisements in a paper or magazine. It additionally incorporates a bulletin, handout, business on TV or radio, banner, and so forth It is a non-advanced method of advertising. Though advanced advertising utilizes different computerized channels to arrive at clients. A few examinations are given beneath:

IMPORTANCE OF DIGITAL MARKETING

Internet promoting is interminably more moderate than any disconnected showcasing strategies. It can contact a more extensive crowd without any problem. In advanced promoting results can be followed and checked effectively with the assistance of different following programming. Maybe than driving exorbitant customer research, associations can quickly see customer response rates and measure the accomplishment of their advancing exertion ceaselessly, enabling them to plan even more sufficiently for the accompanying one.

Collecting input from clients is not difficult to contrast with customary advertising mediums like TV, radio, or announcement. They can undoubtedly give input on any item utilizing a site in internet showcasing which assists a financial specialist with updating themselves in their particular area.

It helps in propelling a business through the online medium like web or compact in this manner showing up at an enormous number of customers in a second. Various little and enormous associations are following the philosophies of electronic displaying to guarantee themselves universally.

Digit publicists screen things like what is being seen, how often and for how long, what substance works and doesn't work, etc. While the web is perhaps, the channel most immovably associated with computerized promoting, others join remote substance educating, compact applications, progressed TV, and radio channels. Digital promoting is moderate, targetable, and quantifiable and subsequently associations do it and sponsors love it.

ADVANTAGES OF DIGITAL MARKETING

In the year 2020 shoppers can get to the web any time from any spot on the planet. Also, because of the computerized idea of advanced promoting buyers can remain refreshed with regards to any item or administration all day, every day times.

Due to the web shopper can do different exercises like going to the organization's site, perusing the data, purchasing items, and so forth This has expanded shopper's commitment and worked on their experience.

In conventional showcasing, there is a little possibility that buyers can be misguided by salesmen, however in computerized advertising, buyers get clear and exact data about any item or administration. What's more, the web gives thorough thing information that customers can rely upon and make a purchase decision.

Many various organizations advance their item through computerized promoting, so it becomes helpful to think about results of the various organizations for the shopper. They don't have to visit different retail locations to analyze items. Internet is accessible the entire day so there is no limitation on schedule and clients can purchase the item any time. Due to the computerized medium watcher can share data and qualities about items or administrations with others.

The association shows the expenses of things through computerized channels and this makes costs amazingly comprehended and clear for the purchaser. The organization likewise changes its costs on any occasion or celebration to give a rebate and is extremely straightforward to the shopper

VI. DISADVANTAGES OF DIGITAL MARKETING

A contender can without much of a stretch duplicate the computerized advertising effort of others. Brand names or logos can be used to cheat clients.

If the web association is slow or there is some issue with sites, then, at that point sites might require some investment to open and the client won't stand by a lot and leave.

In customary promoting, clients can genuinely contact the items to confirm yet it is beyond the realm of imagination in web based business.

Though India is digitalizing, numerous clients actually don't confide in the online installment framework or don't have a clue.

The nonattendance of trust of the customers because of the huge number of fakes concerning virtual progressions. Reasonable associations may be impacted since their image and the reputation of significant worth can suffer hurt.

There are many situations when clients requested items through pay on conveyance technique without really any aim of purchasing utilizing a phony id. This shows a proviso of pay on the conveyance technique.

Digital promoting is significantly dependent on the web/innovation which can be leaned to botches.

Digital advertising isn't yet gotten a handle on by all people: a few clients, particularly more settled ones don't have confidence in a modernized environment, needing to use the customary techniques.

CHALLENGES FACING DIGITAL MARKETERS

Consumers utilize diverse advanced gadgets and different computerized channels and those gadgets have different computerized channels that lead to the expansion of advanced channels. Furthermore, Marketers face trouble in picking medium and crowd. Digital promoting is incredibly modest contrast with conventional showcasing and it covers each private venture which prompts serious rivalry. Consumers leave behind a tremendous measure of information in advanced channels each time they visit the channel. It's incredibly difficult to see such information, similarly as find the right information inside exploding data volumes that can help you to settle on the ideal decisions.

OBJECTIVES OF THE STUDY

- ✓ To Understand the various channels of digital marketing
- ✓ To understand the Comparison of traditional marketing and digital marketing
- ✓ To know the Importance of digital marketing
- ✓ To study the Advantages and disadvantages of digital marketing
- ✓ To find the Challenges for digital marketer

FINDINGS

- ✓ 95% of the people are using digital marketing for their convenience
- ✓ It is the time saving method.
- ✓ Easy and very convenient to all.
- ✓ We can receive the products to our location itself.
- ✓ We can use computerized promoting.

SUGGESTIONS

- ✓ The study helps us to know the process of digital marketing.
- ✓ It provides shoppers to improve their showcase the products to the customer.
- ✓ Digital marketing provides various channels to improve their promoting ideas.
- ✓ Digital marketers and shoppers are can use various websites to sale their products.
- ✓ Users of the digital marketers should be aware about the fraudulent and threats.

CONCLUSION

It can't be dismissed that the world is rapidly moving from easy to the computerized world. People are putting more in online substance and organizations that think that it is difficult to process this reality in their promoting system need to change rapidly. The additional time people spend on the web each year; the more computerized stage they use play a consistently creating capacity in their lives. The fundamental point of advanced India is to advance computerized medium. Since individuals can utilize computerized stage any time anyplace from the world organizations needs to change their advertising system from conventional to advanced. In case the organizations don't use the computerized stage to promote their item and administrations then they can't rival contests and will ultimately close down.

At the point when clients need to purchase any item on the web, they can undoubtedly get item data and can contrast and different items without visiting any retail location or shopping center. It shows that shoppers are more disposed towards internet purchasing instead of visiting a retail location. As purchaser's purchasing conduct is changing organizations likewise need to change their promoting system and embrace computerized stages for showcasing.

References

1. A. Yasmin. S. Tasneem and K. Fatema, “Effectiveness of digital marketing in the challenging age: an empirical study,” Journal of International Business Research and Marketing, vol. 1, no. 5, Apr., pp. 69-80, 2015.
2. [https://blog.hubspot.com/marketing/what-is-digital-marketing#:~: text= Digital %20marketing%20encompasses%20all%20marketing,with%20current%20and%20prospective%20customers.](https://blog.hubspot.com/marketing/what-is-digital-marketing#:~:text=Digital%20marketing%20encompasses%20all%20marketing,with%20current%20and%20prospective%20customers.)
3. P. Ankita and K. Naveen, “Review of digital marketing with latest tools and its effect on business models,” International Journal for Research in Applied Science and Engineering Technology, vol. 8, no. 5, May., pp. 2321-9653, 2020. [Online Serial]. Available: <https://www.ijraset.com/files/serve.php?FID=29098>. [Accessed September 25, 2020].

SECURITY AND PRIVACY BY DESIGN FOR AHA-IOT APPLICATIONS AND SERVICES

*Ms.P Monisha,
Assistant professor of computer Application,
MKJC,TN, India.*

ABSTRACT:

The chapter aims at describing the cybersecurity and privacy methodologies and solutions that the architecture defined within the ACTIVAGE Large-Scale Pilot, and therefore the corresponding implementation in nine Deployment sites should follow to secure the IoT system and protect the private data from potential malicious cyber-attacks and threats. It further presents common definitions, methods and repeatable processes to analyse and address all potential threats interms of cybersecurity and privacy that might occur during the exploitation phase of the project. The Internet of Things (IoT) is an emerging technology of the 21st century. the most important law concerning the protection of private data in Malaysia, the private Data Protection Act (PDPA) 2010 might be used as a benchmark for assessing the adequacy of data protection law in the country. Thus, this paper attempts to shed light on data protection challenges within the IoT era then assess the adequacy of this Act in handling those challenges. The paper employs a legal doctrinal method to research the legal frameworks relevant to non-public data protection

1.Introduction:

The Internet of Things (IoT) is an emerging technology that plays an important role within the modern age. At present time, the IoT technology are often found almost altogether things surrounding people like cars, houses, wearable devices and such like. The basic function of IoT revolves around connecting things (persons, animals, cars, trees, etc.) to the web and enabling them to speak then process (send, receive, generate, etc.) information about themselves and therefore the things they're attached to. The billions of “things” are often the target of intrusions and interferences which may dramatically jeopardize personal privacy. • Technological – a secure large-scale deployment of connected objects. • Societal – related to the project context, which is to create a smart environment for ageing well of elderly people allowing the collection of sensitive personal data. As in ACTIVAGE, the experimentations will involve around 7,000 users across the consortium features a great concern when it involves the safety and privacy related challenges and an opportunity to resolve these issues with the assistance of large-scale validation and testing. Platforms using public communication infrastructure will interconnect many IoT devices, which are inherently weakly secured. Since

Marudhar Kesari Jain College For women,vaniyambadi

IoT is seen as a key enabler for creating new services and improving overall quality of life, consumers need to have trust and confidence about their data being secured and protected, therefore, making the cybersecurity of IoT systems an essential part the “Active and Healthy Ageing” (AHA) domain producing evidence of the IoT value on fostering the deployment of AHA solutions in Europe, through the mixing of advanced IoT technologies across the worth chain, demonstrating multiple AHA-IoT applications at large-scale during a usage context, in real operational conditions. IoT for the AHA domain may be a strategic element for the creation of dynamic ecosystems to answer and stop the challenges faced by health and social care systems. Differently from other sectors, “AHA-IoT” services are provided to persons taken individually and it takes place across all domains, as persons sleep in houses, neighbourhoods, cities, rural areas, mountains and valleys, access to move systems, drive cars, attend shopping centres, airports, theatres, etc.

2. SECURITY APPROACH

Methodology to realize the objectives defined above, variety of activities are performed to get down the safety and privacy policies within the context of ACTIVAGE project. For the aim of security, activities include:

- Perform a reference risk analysis within the ACTIVAGE IoT environment so as to spot the overall ACTIVAGE security requirements, which depend upon the criticality of applications or services.
- Countermeasures to mitigate risks are identified at this stage.
- Create and elaborate the ACTIVAGE security questionnaire.
- Analyse questionnaires’ responses and perform assessments for the DS’ security requirements.
- Define the security cartography and recommendations for each deployment site. The elaboration of the safety questionnaire considered the subsequent aspects:
- Collect relevant information allowing the identification of missing mechanisms to ensure full end-to-end cyber security and privacy for each of the DSs.
- Make it easy for the DSs security managers to reply. The DS security manager is in charge of the security and privacy aspects related to this DS.
- Make the DSs security managers aware of cybersecurity and privacy issues that have not yet been identified and support the other stakeholders to realize the high importance of these aspects that are critical considering the nature of the project, which includes data confidentiality,

higher vulnerability by connecting “smart objects” to the system, etc. Assets identification and description An Assets list was established as a guideline to be carefully analysed, completed (if needed) and used for each DS. It includes all data within the system, services, pieces of hardware, software, communication links and should be extended to property, brand reputation, buildings etc. The most important items during this list are given hereafter. Data assets include application and management data. The typesets and formats should be defined in the data model. Application data describe the elements

- Data describing all entities producing or consuming data (Identifiers and attributes of individuals, stakeholders, sensors). Data that are monitored and analysed by the IoT system in order to ensure the expected service (raw measurements, processed data elements). Decisions of the system that influence the subject’s environment (guidance or prescriptions for individuals, environmental instructions for smart sensors, configuration instructions for devices). Management data relate to system operation. They include, for example: Procedure, action plan descriptions (definition of all the planned actions in case of occurrence of an extreme event). Data storage organization definition (for example, a Grading Table, Detail Description predefines categories for data storage, such as Medical information, Medical report, Wellness information, Service, etc.). Access Rights Table, defining the access rights for each stakeholder profile. Transaction registers, logging the History of all operated transactions (communication channel, data, data user, time, etc.).

Methodology to Perform Privacy Analysis and Recommendations shows the Privacy methodology proposed in order to perform risk privacy analysis on an IoT system. This is the methodology we've utilized in ACTIVAGE for this purpose. The expected outcomes are the identification of the countermeasures/recommendations for this IoT system to minimize the risks of privacy threats: data theft, data misuse or any other malicious usage. This methodology is addressed to any non-professional data protection manager to facilitate, him/her, the implementation of the GDPR regulation.

3. What is the GDPR?

Architecture and topology where the Data will be generated, stored, processed and exploited (and by whom) to identify security rights? In order to get the answers to these questions, the following documents are available. Identify personal data flow and storage – For any IoT system, it is required to know its complete and detailed architecture and topology as

discussed in Section. “Easily” the identification of assets, data flows, data storage, process units, users, etc. and their location. Perform Data Impact Performance Assessment – (DPIA) This step is key in the methodology. The importance of this step and therefore the thanks to develop it are described with more details within the next paragraph. Provide Privacy Impact Analysis and Recommendations – This step provides the DPIA analysis results of the IoT system under study and the recommendations proposed to deploy the system with good Privacy Data Protection Impact Assessment (DPIA)⁴ GDPR introduces the concept of a Data Protection Impact Assessment (DPIA)⁵ [20] and strongly recommend carrying out one for each system concerned. This paragraph addresses the subsequent questions: what's a DPIA?, when a DPIA is mandatory and the way to hold it?, and what are the main elements containing a DPIA?

4. What is a DPIA?

“A DPIA is a process designed to describe the processing assess The necessity and proportionality of a processing and to help managing the risks to the rights and freedoms of natural persons resulting from the processing of personal data. DPIAs are important tools for accountability, as they assist controllers not only to suits requirements of the GDPR, but also to this information contained during this paragraph was extracted from [20].⁵ The term “Privacy Impact Assessment (PIA) is often used in other contexts to refer to the same concept”,

5. Privacy Approach 121 demonstrate that appropriate measures are taken to make sure compliance with the Regulation. In other words, a DPIA may be a process for building and demonstrating compliance”.

6. Conclusion

In this chapter, two complementary methodologies were presented one for security and the other for privacy in order to address the challenges presented in the previous paragraphs. They were developed to help the IoT System developers of ACTIVAGE to secure their systems and implement correctly personal data protection to cope with the GDPR requirements. These methodologies follow a twofold approach a top down and an end-to-end. These approaches concern from one side the safety risk analysis to spot beforehand potential threats and find the countermeasures to mitigate/avoid them. From the opposite side, a privacy approach to place in

situ the GDPR following a DPIA analysis to spot the system characteristics and evaluate the risks associated with the personal data and its protection. This work, developed in the frame of the ACTIVAGE project, can be also reused for any other IoT system considering the high constraints on security and privacy required by AHA applications. Finally, the solutions presented give a good overview of the possibilities offered by the use of the Secure element component to secure IoT devices (Gateways and Sensor nodes) and the Blockchain technology in AHA applications. Both technologies will take a crucial place within the implementation and validation of the safety and privacy requirements of the ACTIVAGE's Deployment sites to supply secure IoT systems with a high level of private data protection and thus to extend the users' trust. Future work will put in situ and validate these methodologies and therefore the potentials solutions to secure the 9 Deployment sites of ACTIVAGE project also because the protection of the personal data of each of the seven thousands of patients "elderly people" participating in the project.

References

135

Acknowledgement This research project has received funding from the European Union's Horizon 2020 research and innovation programme ACTIVAGE under grant agreement No 732679. The activities concerning the Secure Gateway has received funding from the French National Research Agency within the framework of the "Investissements d'avenir" program (ANR-10-AIRT-05)".

[1] Internet Security Threat Report ISTR Ramsonware 2017 An ISTR Special Report July 2017. [2] Proposal for a Regulation OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on ENISA, the "EU Cybersecurity Agency", and repealing Regulation (EU) 526/2013, and on Information and Communication Technology cybersecurity certification ("Cybersecurityact"), online at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1505290611859&uri=COM:2017:477:FIN>

[3] Internet of Things European Large-Scale Pilots Programme, online at: <https://european-iot-pilots.eu>

[4] ACTIVAGE Large-Scale Pilot project, online at: <https://www.activageproject.eu/>

[5] General Data Protection Regulation (GDPR) REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC Official journal of the European Union, online at: <https://gdpr-info.eu/>

- [6] NIST FIPS 199, Standards for Security Categorization of Federal Information and Information Systems (February 2004), online at: <https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.199.pdf>
- [7] NIST SP 800-53A, NIST Special Publication 800-53A. Revision 4. Assessing Security and Privacy. Controls in Federal Information. Systems and Organizations (December 2014), online at: <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-53Ar4.pdf>
- [8] Mahmoud Elkhodr, Seyed Shahrestani and Hon Cheung, The Internet of Things: New Interoperability, Management and Security Challenges
- [9] End-to-end Security and Privacy by Design for AHA-IoT Applications
- [10] International Journal of Network Security & Its Applications (IJNSA) Vol. 8, No. 2, March 2016.
- [11]. Sadeghi, A.-R.; Wachsmann, C.; Waidner, M. Security and privacy challenges in industrial internet of things. In Proceedings of the 2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC), San Francisco, CA, USA, 8–12 June 2015; pp. 1–6.
12. Izzat, A.; Chuck, E.; Lo'ai, T. The NICE Cyber Security Framework, Cyber Security Management; Springer: Basel, Switzerland, 2020; ISBN 978-3-030-41987-5.
13. Tawalbeh, L.A.; Tawalbeh, H. Lightweight crypto and security. In Security and Privacy in Cyber-Physical Systems: Foundations, Principles, and Applications; Wiley: West Sussex, UK, 2017; pp. 243–261.

NETWORK SECURITY AND CRYPTOGRAPHY

*Ms.S.Bhavaniv¹, V.Kavi Bharathil², L.Sathiya³
Dept of Computer Science ,
MKJC,TN, India.*

ABSTRACT

Network Security & Cryptography is a concept to protect network and data transmission over wireless network. Data Security is the main aspect of secure data transmission over unreliable network. Network security involves the authorization of access to data in a network, which is controlled by the network administrator. Users choose or are assigned an ID and password or other authenticating information that allows them access to information and programs within their authority. Network security covers a variety of computer networks, both public and private, that are used in everyday jobs conducting transactions and communications among businesses, government agencies and individuals. Networks can be private, such as within a company, and others which might be open to public access. Network security is involved in organizations, enterprises, and other types of institutions. In this paper we also studied cryptography along with its principles. Cryptographic systems with ciphers are described.

KEYWORD : Network security, problems and attacks, cryptography, principles and services, encryption and decryption, cipher, algorithm, advantages and disadvantages, web security.

INTRODUCTION

Network Security is the most vital component in information security because it is responsible for securing all information passed through networked computers. Network Security refers to all hardware and software functions, characteristics, features, operational procedures, accountability, measures, access control, and administrative and management policy required to provide an acceptable level of protection for Hardware and Software, and information in a network. Network security problems can be divided roughly into four closely intertwined areas: secrecy, authentication, nonrepudiation, and integrity control. Secrecy, also called confidentiality, has to do with keeping information out of the hands of unauthorized users. This is what usually comes to mind when people think about network security. Authentication deals with determining whom you are talking to before revealing sensitive information or entering into a business deal. Nonrepudiation deals with signatures. Message Integrity: Even if the sender and receiver are able to authenticate each other, they also want to insure that the content of their communication is not altered, either maliciously or by accident,

in transmission. Extensions to the checksumming techniques that we encountered in reliable transport and data link protocols. Cryptography is an emerging technology, which is important for network security. The widespread use of computerised data storage, processing and transmission makes sensitive, valuable and personal information vulnerable to unauthorised access while in storage or transmission. Due to continuing advancements in communications and eavesdropping technologies, business organisations and private individuals are beginning to protect their information in computer systems and networks using cryptographic techniques, which, until very recently, were exclusively used by the military and diplomatic communities.

NETWORK SECURITY

Network security is a broad term that covers a multitude of technologies, devices and processes. In its simplest term, it is a set of rules and configurations designed to protect the integrity, confidentiality and accessibility of computer networks and data using both software and hardware technologies.

PROBLEMS AND ATTACKS

Problem #1: Unknown Assets on the Network

There are many businesses that don't have a complete inventory of all of the IT assets that they have tied into their network. This is a massive problem. If you don't know what all of the assets are on your network.

Problem #2: Abuse of User Account Privileges

According to data cited by the Harvard Business Review, for the year of 2016, "60% of all attacks were carried out by insiders."

Problem #3: Unpatched Security Vulnerabilities

Many businesses are concerned with "zero day" exploits. These exploits are those unknown issues with security in programs and systems that have yet to be used against anyone. However, zero day vulnerabilities aren't the problem—unlatched known vulnerabilities are the problem

ATTACKS

Endpoint attacks—gaining unauthorized access to user devices, servers or other endpoints, typically compromising them by infecting them with malware. Malware attacks—infecting IT resources with malware, allowing attackers to compromise systems, steal data and do damage. These also include ransom ware attacks.

CRYPTOGRAPHY

Cryptography is a method of protecting information and communications through the use of codes, so that only those for whom the information is intended can read and process it. The prefix "crypt-" means "hidden" or "vault" -- and the suffix "-graphy" stands for "writing."

In computer science, cryptography refers to secure information and communication techniques derived from mathematical concepts and a set of rule-based calculations called algorithms, to transform messages in ways that are hard to decipher. These deterministic algorithms are used for cryptographic key generation, digital signing, verification to protect data privacy, web browsing on the internet, and confidential communications such as credit card transactions and email.

PRINCIPLES AND SERVICES

In present day scenario security of the system is the sole priority of any organisation. The main aim of any organisation is to protect their data from attackers. In cryptography, attacks are of two types such as Passive attacks and Active attacks.

Passive attacks are those that retrieve information from the system without affecting the system resources while active attacks are those that retrieve system information and make changes to the system resources and their operations.

SERVICES

This article summarizes the basic cryptographic security services that can be used to protect information (or as a supporting protective mechanism) against attacks, as described in the NIST Special Publication 800-57 (1, rev.4) for Key Management.

The publication describes the following basic security services as confidentiality, integrity, authentication, source authentication, authorization and non-repudiation.

A range of cryptographic and non-cryptographic tools may be used to support these services. While a single cryptographic mechanism could provide more than one service, it cannot provide all services

ENCRYPTION AND DECRYPTION

S.NO	Encryption	Decryption
1.	Encryption is the process of converting normal message into meaningless message.	While decryption is the process of converting meaningless message into its original form.

2.	Encryption is the process which take place at sender's end.	While decryption is the process which take place at receiver's end.
3.	Its major task is to convert the plain text into cipher text.	While its main task is to convert the cipher text into plain text.
4.	Any message can be encrypted with either secret key or public key.	Whereas the encrypted message can be decrypted with either secret key or private key.
5.	In encryption process, sender sends the data to receiver after encrypted it.	Whereas in decryption process, receiver receives the information(Cipher text) and convert into plain text.

CIPHER

Ciphers, also called encryption algorithms, are systems for encrypting and decrypting data. A cipher converts the original message, called plaintext, into ciphertext using a key to determine how it is done.

Ciphers are generally categorized according to how they work and by how their key is used for encryption and decryption. Block ciphers accumulate symbols in a message of a fixed size (the block), and stream ciphers work on a continuous stream of symbols. When a cipher uses the same key for encryption and decryption, they are known as symmetric key algorithms or ciphers. Asymmetric key algorithms or ciphers use a different key for encryption/decryption.

Example:

“The government of ancient Rome was among the first civilizations to use ciphers to transmit sensitive information such as military conversations.”

ALGORITHMS

RSA is a public-key encryption algorithm and the standard for encrypting data sent over the internet. It also happens to be one of the methods used in PGP and GPG programs. Unlike Triple DES, RSA is considered an asymmetric encryption algorithm because it uses a pair of keys.

ADVANTAGES AND DISADVANTAGES

As discussed, network security keeps a check on unauthorized access. A network contains a lot of confidential data like the personal client data. Anybody who breaks into the network may hamper these sensitive data. So, network security should be there in place to protect them.

Prevent cyber ADVANTAGES AND DISADVANTAGES

As discussed, network security keeps a check on unauthorized access. A network contains a lot of confidential data like the personal client data. Anybody who breaks into the network may hamper these sensitive data. So, network security should be there in place to protect them.

Prevents cyber attack

Most of the attack on the network comes from internet. There are hackers who are experts in this and then there are virus attacks. If careless, they can play with a lot of information available in the network. The network security can prevent these attacks from harming the computers.

DISADVANTAGES

Network security is a real boon to the users to ensure the security of their data. While it has many advantages, it has lesser disadvantages. Let us discuss some of them.

Costly setup

The set up of a network security system can be a bit expensive. Purchasing the software, installing it etc can become costly especially for smaller networks. Here we are not talking about a single computer, but a network of computers storing massive data. So, the security being of prime importance will definitely cost more. It cannot be ignored at any cost!

attack

Most of the attack on the network comes from internet. There are hackers who are experts in this and then there are virus attacks. If careless, they can play with a lot of information available in the network. The network security can prevent these attacks from harming the computers.

DISADVANTAGES

Network security is a real boon to the users to ensure the security of their data. While it has many advantages, it has lesser disadvantages. Let us discuss some of them.

Costly setup

The set up of a network security system can be a bit expensive. Purchasing the software, installing it etc can become costly especially for smaller networks. Here we are not talking

about a single computer, but a network of computers storing massive data. So, the security being of prime importance will definitely cost more. It cannot be ignored at any cost!

WEB SECURITY

In general, web security refers to the protective measures and protocols that organizations adopt to protect the organization from, cyber criminals and threats that use the web channel. Web security is critical to business continuity and to protecting data, users and companies from risk.

REFERENCES

1. Menezes, van Oorschot and Vanstone, Handbook of Applied Cryptography (1996, CRC Press; 2001 with corrections), free online for personal use.
2. Keith M. Martin, Everyday Cryptography (2017, 2/e; Oxford University Press).
Kaufman, Perlman and Speciner, Network Security: Private Communications in a Public World, second edition (2003, Prentice Hall).
3. William Stallings, Cryptography and Network Security: Principles and Practice, 5/e (2010, Prentice Hall). Relative to this book's 4th edition, the network security components and an extra chapter on SNMP are also packaged as Stallings' Network Security Essentials: Applications and Standards, 3/e (2007, Prentice Hall).

NETWORKING & SECURITY*Ms.J.Varshini¹, Ms.S.Soniya², MS.U.Kayalvizhi³**Dept of Computer Application,
MKJC,TN, India***ABSTRACT**

The Security of computer networks plays a important role in modern computer systems . In order to execute high protection levels against hatful attack, a number of software tools have been currently developed. Intrusion Detection System has recently become a heated research topic due to its capability of noticing and preventing the attacks from malicious network users. A pattern matching IDS for network security has been suggest in this paper. Many network security applications depend on pattern matching to remove the threat from network traffic. The increase in network speed and traffic may make extant algorithms to become a performance bottleneck. Therefore it is very necessary to develop faster and more efficient pattern matching algorithm in order to overcome the troubles on performance.

INTRODUCTION:**NetworkSecurity :**

Network and computer security is captions to the financial health of every organization. Over the past few years ,Internet-enabled business,ore-business, hasseverely improved efficiency and revenue growth. E-business applications such as e-commerce,supply-chain management, and remote access allow companies to streamline processes, loweroperating costs, and increase customer satisfaction. Such applications require mission-critical networks that accommodate voice, video, and data traffic, and these networks must be scalable to support increasing numbers of users and the need for greater capacity and performance. However,asnetworksenablemore and more applications and are available to more andmore users, they become ever more vulnerable to a widerrange of security threats. To combat those threats and ensurethat e-business transactions are not compromised, securitytechnologymust playa majorrole in today's networks.

**Necessity:**

Security incidents are rising at an alarming rate every year. As the complexity of the threats increases, so do the security measures required to protect networks. Data center operators,

network administrators, and other data center professionals need to behold the basics of security in order to safely dispose and manage networks today.

Objectives :

As time goes on, too many new technology will be developed to further improve the efficiency of business and communications. At the same time, development in technology will provide even greater network security, therefore, greater piece of mind to operate in cutting edge business environments. Provided that enterprises stay on top of this emerging technology, as well as the latest security threats and dangers, the welfare of networks will most certainly outweigh the risks.

Importance of Network Security :

To protect company assets: One of the primary goals of computer and network security is the protection of company information that is housed on a company's computers and networks.

To gain a competitive advantage: Developing and maintaining successful security measures can provide an organization with a competitive advantage over its competition. Network security is particularly important in the arena of Internet financial services and e-commerce.

TYPES OF NETWORK SECURITY:-

We have talked about the different types of network security controls. There are some of the different ways we can secure our network.

Network Access Control :

To ensure that potential cannot strike your network, comprehensive access control policies need to be in place for both users and devices. Network access control (NAC) can be set at the most granular level. For example, you could grant administrators full access to the network but refuse access to specific confidential folders or prevent their personal devices from joining the network.

Antivirus and Antimalware Software :

Antivirus and antimalware software safeguard an organization from a range of malicious software, including viruses, ransomware, worms and trojans. The greatest software not only scans files upon entry to the network but continuously scans and tracks files.

Firewall Protection

Firewalls, as their name propose, act as a barrier between the untrusted external networks and your trusted internal network. Administrators generally configure a set of defined rules that blocks or permits traffic onto the network. For example, Forcepoint's Next Generation Firewall

(NGFW) offers seamless and centrally managed control of network traffic, whether it is physical, virtual or in the cloud.

Virtual Private Networks :

Virtual private networks (VPNs) create a connection to the network from another endpoint or site. For example, users working from home would typically connect to the organization's network over a VPN. Data between the two points is encrypted and the user would need to authenticate to allow communication between their device and the network. Forcepoint's Secure Enterprise SD-WAN allows organizations to rapidly create VPNs using drag-and-drop and to safeguard all locations with our Next Generation Firewall solution.

ENEMIES OF NETWORK SECURITY:

Hackers: This generic and often over-romanticized term applies to computer enthusiasts who take delight in gaining access to other people's computers or networks.

Unaware Staff: As employees focus on their particular job duties, they often overlook standard network security rules. Like simple password, use of Virus effecting CD/DVD etc.

Snoops: Employees known as "snoops" part take in corporate espionage; get unauthorized access to confidential data in order to provide competitors with otherwise inaccessible information.

EFFECT OF ENEMIES**Viruses:**

Viruses are the most generally known security threats, because they often garner extensive press coverage. Viruses are computer programs that are written by devious programmers and are designed to replicate themselves and infect computers when triggered by a specific event.

A network can be spoiled by a virus only if the virus enters the network through an outside source—most often through an infected floppy disk or a file downloaded from the Internet.

When one computer on the network becomes infected, the other computers on the network are highly susceptible to contracting the virus.

Trojan horse Programs:

Trojan horse programs, or trojans, are delivery vehicles for destructive code. Trojans seem to be harmless or useful software programs, such as computer games, but they are actually enemies in disguise. Trojans can delete data, mail copies of themselves to e-mail address lists, and open up computers to additional attacks.

Trojans can be contract only by reduce the Trojan horse program to a system, via a disk, downloading from the Internet, or opening an e-mail attachment. Neither trojansnor viruses can be spread through an e-mail message itself—they a respired only through e-mail attachments.

Vandals:

Web sites have come alive through the enlargement of such software applications as ActiveX and Java Applets. These devices enable animation and other special effects to run, making Web sites more attractive and interactive. However, the ease with which these applications can be downloaded and run has provided a new vehicle for inflicting damage. A vandal is a software application or applet that causes destruction of varying degrees. A vandal can demolish just a single file or a major portion of a computer system.

Attacks:

Innumerable types of network attacks have been documented and they are commonly classified in three general categories: Reconnaissance attacks, Access attacks and Denial of service (DoS) attacks. Reconnaissance attacks are essentially information gathering activities by which hackers collect data that is used to later compromise networks. Usually, software tools, such as sniffers and scanners, are used to map out network resources and exploit potential defect in the targeted networks, hosts, and applications. Access attacks are conducted to exploit vulnerabilities in such network are as authentication services and File Transfer Protocol (FTP) functionality in order to get entry to e-mail accounts, databases, and other confidential information.

DoS attacks prevent access to part or all of a computer system. They are usually achieved by sending large amounts of jumbled or otherwise unmanageable data to a machine that is connected to a corporate network or the Internet, blocking legitimate traffic from obtain through. Even more malicious is a Distributed Denial of Service attack (DoS) in which the attacker compromises multiple machines or hosts.

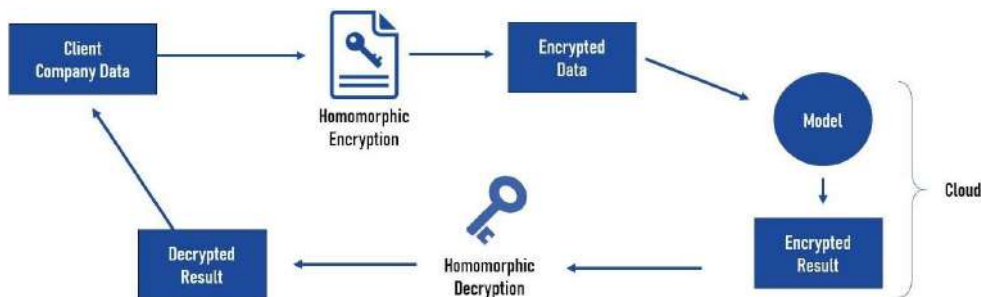
Virus protection software is packaged with most computers and can counter most virus threats if the software is regularly renew and correctly maintained. The anti-virus industry relies on a vast network of users to provide early warnings of new viruses, so that antidotes can be developed and distributed quickly. With thousands of new viruses being generated every month, it is essential that the virus index is kept up to date. The virus database is the recordheldbytheanti-viruspackagethathelpsittoidentifyknownviruses when they attempt to strike. Reputable anti-virus software vendors will publish the latest antidotes on their Web sites, and the software can prompt users to systematically collect new data Network security

policy should impose that all computers on the network are kept up to date and, ideally, are all protected by the same anti-virus package—if only to keep maintenance and update costs to a minimum. It is also fundamental to update the software itself on a regular basis. Virus authors often make getting past the anti-virus packages their first priority.

Security Policies:

When setting up a network, whether it is a local area network (LAN), virtual LAN (VLAN), or wide area network (WAN), it is necessary to initially set the fundamental security policies. Security policies are rules that are electronically programmed and stored within security material to control such areas as access privileges. The policies that are implemented should control who has access to which areas of the network and how unaroused users are going to be prevented from entering restricted areas. The individual or group of people who police and maintain the network and its security must have access to every area of the network.

Encryption: Encryption technology ensures that messages cannot be caught or read by anyone other than the authorized recipient. Encryption is usually deployed to protect data that is transported over a public network and uses advanced mathematical algorithms to “scramble” messages and their attachments.



A network-based intrusion detection system (IDS) provides around-the-clock network surveillance. An IDS evaluates packet data streams within a network, searching for unauthorized activity, such as attacks by hackers, and enabling users to respond to security breaches before systems are compromised. When unauthorized activity is detected, the IDS sends alarms to a management console with details of the activity and can often order other systems, such as routers, to cut off the unaroused sessions. In the physical analogy, an IDS is equivalent to a video camera and motion sensor; detecting unauthorized or suspicious activity and working with automated response systems, such as watch guards to stop the activity.

CONCLUSION

As time goes on, more and more new technology will be developed to further improve the capability of business and communications. At the same time, breakthrough in technology will provide even greater network security, therefore, greater piece of mind to operate in cutting edge business environments. Provided that enterprises stay on top of this emerging technology, as well as the latest security threats and dangers, the benefits of networks will most certainly outweigh the prospect.

REFERENCES

1. http://en.wikipedia.org/wiki/Network_security
2. <http://www.interhack.net/pubs/network-security/>
3. <http://e-articles.info/e/s/s/Network-security/>
4. ijns.femto.com.tw/contents/ijns-v10-n1/ijns-v10-n1.html
5. pnbiit.com/download/JulSep09.pdf

BIG DATA ANALYTICS

*Ms.M. Hemavathi¹, Ms.B. Vedhasri², Ms.B.Harini³
Dept of computer Applications
MKJC,TN,India.*

ABSTRACT

In the information of data have become available on hand to decision makers. Big data is a new driver of the world economic and societal changes. The world's data collection is reaching a tipping point for major technological changes that can bring new ways in decision making, managing our health, cities, finance and education. Big Data Analytics poses a grand challenge on the design of highly scalable algorithms and systems to integrate the data and uncover large hidden values from datasets that are diverse, complex, and of a massive scale. Potential breakthroughs include new algorithms, methodologies, systems and applications in Big Data Analytics that discover useful and hidden knowledge from the Big Data efficiently and effectively. Big Data is a field that treats ways to analyse, systematically extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional data- processing application software. Big data refers to data sets that are not only big, but also high in variety and velocity, which makes them difficult to handle using traditional tools and techniques. Due to the rapid growth of such data, solutions need to be studied and provided in order to handle and extract value and knowledge from these datasets. This paper aims to analyse some innovative techniques methods and tools which can be applied to big data.

INTRODUCTION

Big data analytics is the process of examining large data sets containing a variety of data types that is big data to uncover hidden patterns, unknown correlations, market trends, customer preferences and other useful business data's. Imagine that today's world without a data storage: a place where every information of a person or every details about the organization, transaction of companies or every aspect which can be documented is lost after use. In this case this article provides new opportunities and advantages, that where anything ranging from customers data, to products available and further details and etc, has become necessary for day-to-day connectivity. Data is the building block upon which any organization succeed. The primary goal of big data analytics is to help companies make more informed business decisions by enabling data scientists, predictive modelers and other analytics professionals to analyse large volumes of transactions data, as well as other forms of data that may be untapped by conventional business intelligence programs. Now think of the extent of details and the surge of

data and information provides nowadays through the advancements in technologies and the internet, with the increase in storage capabilities and methods of data collection, huge amount of data have become easily available. The size, variety, and rapid change of such data require a new type of big data analytics as well as different storage and analysis methods. Such amounts of big data need to be correctly analysed and pertaining data should be extracted. Big data can be analysed with software tools commonly which used as part of advanced analytics disciplines such as predictive analytics, data mining, text analytics and statistical analysis. As a result of many organization looking to collect and process and analyse big data have turned to a newer class of technologies that Hadoop and related tools such as YARN, MapReduce, Spark, Hive and pig as well as NoSQL databases.

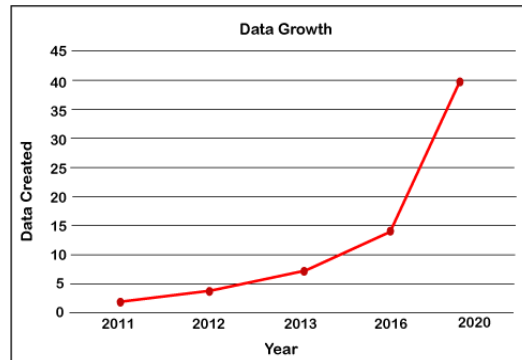
CHARACTERISTICS OF BIG DATA

Big Data contains a large amount of data that is not being processed by traditional data storage or the processing unit. It is used by many Multinational Companies to Process the data and Business of many organizations. The data flow would exceed 150 exabytes per day before replication. There are five v's of Big data



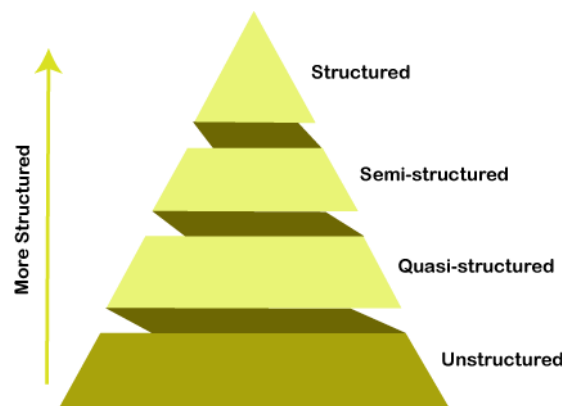
Volume

The name Big data itself related to an enormous size. Big data is a vast 'volumes' of data generated from many sources daily such as Business process, machines, social media and platforms, networks, human interactions, and many more.



Variety

Big Data can be structured, unstructured, and semi-structured that are being collected from different sources. Data will only be collected from databases and sheets in the past, But these days the data will comes in array forms, that are PDFs, Emails, audios, SM posts, photos, videos, etc.

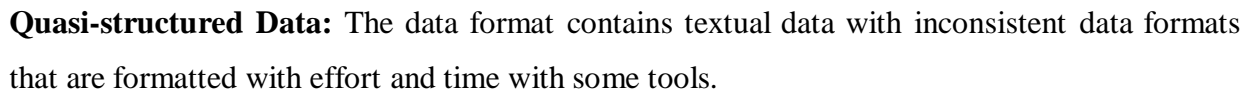


The data is categorized as below:

Structured data: In Structured schema, along with all the required columns. It is in a tabular form. Structured Data is stored in the relational database management system.

Semi-structured: In Semi-structured, the schema is not appropriately defined, e.g., JSON, XML, CSV, TSV, and email. OLTP (Online Transaction Processing) systems are built to work with semi-structured data. It is stored in relations, i.e., tables.

Unstructured Data: All the unstructured files, log files, audio files, and image files are included in the unstructured data. Some organizations have much data available, but they did not know how to derive the value of data since the data is raw.



Veracity means how much the data is reliable. It has many ways to filter or translate the data. Veracity is the process of being able to handle and manage data efficiently. Big Data is also essential in business development.

Value

Value refers to the benefits that your organization derives from the data. Does it match your organization's goals? Does it help your organization enhance itself? It's among the most important big data core characteristics.

Velocity Plays an important role compared to others. Velocity creates the speed by which the data is created in real-time. It contains the linking of incoming data sets speeds, rate of change, and activity bursts. The Primary aspects of Big data is to provide demanding data rapidly. Big data velocity deals with the speed at the data flows from sources like application logs, business processes, networks, and social media sites, sensors, mobile devices, etc.



BIG DATA STORAGE AND MANAGEMENT

Big data storage is a compute and storage architecture that collects and manages large data sets and enables real time analytics. Companies apply big data analytics to greater intelligence from metadata, storage infrastructure that is designed specifically to store and manage and retrieve massive amounts of data or big data. One of the first things that the organization is have to manage is that the dealing with the big data. The traditional methods of structured data storage and retrieval include relational databases, data marts, and data warehouses. The data is uploaded to the storage from operational data stores using Extract, Transform, Load (ETL), or Extract, Load, Transform (ELT), tools which extract the data from outside sources, transform the data to fit operational needs, and finally load the data into the database or data warehouse. Thus the data is cleaned, transformed, and catalogued before being made available for data mining and online analytical functions. Accordingly, several solutions, ranging from distributed systems and massive parallel processing (MPP) databases for providing high query performance and platform scalability, to non-relational or in memory databases have been used for big data. Non relational databases, such as not only SQL (NoSQL), were developed for storing and managing unstructured or Non relational, data. NoSQL databases aim for massive scaling, data model flexibility, and simplified application development and deployment. Country to relational databases rather focus on the high performance scalable data storage and allow data management tasks to be written in the application layer instead of having it written in databases specific languages. On the other hand, in memory databases manage the data in server memory, thus eliminating disk input/output and enabling real time responses from the database. Alternatively, Hadoop is a framework for performing big data analytics which provides reliability, scalability, and manageability by providing an implementation for the Map Reduce paradigm. Hadoop consists of two main components: the HDFS for the big data storage, and Map Reduce for big data analytics. The HDFS storage function provides a redundant and reliable file system, which is optimized for large files, where a single file is split into blocks and distributed across cluster nodes.

BIG DATA ANALYTICS TECHNIQUES

Big data is characterised by the three V's : the major volume of data, the velocity at which its processed, and the wide variety of data. Its because of the second descriptor, velocity, that data analytics has expanded computer based analysis techniques data harnesses, analysis also on the traditional statistical methods. The analytics tools can be applied to both big data and other smaller datasets:

A/B Testing

This data analysis technique involves comparing a control group with a variety of test groups, in order to discern what treatments or change will improve a given objective variable. McKinsey gives the example of analysing what copy, text, images, or layout will improve conversion rates on a e-commerce site. Big data once again fits into this model as it can test huge numbers, however, it can be only be achieved if the groups are of a big enough size to gain meaningful differences.

Data fusion and data integration

By combining a set of techniques that analyse and integrate data from multiple sources and solutions, the insights are more efficient and potentially more accurate than if developed through a single source of data.

Data mining

A common tool used within big data analytics, data mining extracts pattern from large data sets by combining methods from statistics and machine learning, within databases management. An example would be when customer data is mined to determine which segments are most likely to react to an offer.

Natural language processing (NLP)

NLP known as a subspecialty of computer science, artificial intelligence, and linguistics, this data analysis tool uses algorithms to analyse human (natural) language.

APPLICATION OF BIG DATA ANALYTICS

Big data applications can help companies to make better business decisions by analyzing large volumes of data and discovering hidden patterns. These data sets might be from social media, data captured by sensors, website logs, customer feedbacks, etc. Organizations are spending huge amounts on big data applications to discover hidden patterns, unknown associations, market style, consumer preferences, and other valuable business information.

Healthcare

Media and Entertainment

Government

Manufacturing

Big Data in Healthcare

The volume of patient, clinical and insurance records in healthcare generates mountains of data. Big data analytics lets hospitals get important insights out of what would have been an unmanageable amount of data. The ability to extract useful information out of structured

and unstructured data can lead to better outcomes in patient treatment and organizational efficiency.

Big Data in Media and Entertainment

The media and entertainment industries are creating, advertising, and distributing their content using new business models. This is due to customer requirements to view digital content from any location and at any time. The introduction of online TV shows, Netflix channels, etc. is proving that new customers are not only interested in watching TV but are interested in accessing data from any location.

The media houses are targeting audiences by predicting what they would like to see, how to target the ads, content monetization, etc. Big data systems are thus increasing the revenues of such media houses by analyzing viewer patterns.

Big Data in Government

Government agencies face a constant pressure to do more with less resources. Public safety agencies are expected to combat crime and budgets do not always rise in conjunction with crime rates. Big data analytics allows law enforcement to work smarter and more efficiently. It is also used for handling census data. And it allows any government agency to streamline operations and better target resources for maximum results.

Big Data in Manufacturing

Clinical research trials commonly fail, even after using a lot of resources and time. Big data visual analytics provides the insights researchers need to try more trials faster. It allows for automated solutions that affect speed and efficiency.

CONCLUSION

In this research, we have examined the innovative topic of big data, which has recently gained lots of interest due to its perceived unprecedented opportunities and benefits. Big data analytics can be applied to leverage business change and enhance decision making, by applying advanced analytic techniques on big data, and revealing hidden insights and valuable knowledge. Big Data is among the most in-demand technologies currently. Companies of various industries are looking for ways to utilize Big Data to enhance their operations, attract more customers, and get ahead of their competitors. Accordingly, the literature was reviewed in

order to provide an analysis of the big data analytics concepts which are being researched, as well as their importance to decision making. Consequently, big data was discussed, as well as its characteristics and importance. Moreover, some of the big data analytics tools and scope in particular were examined. Thus, big data storage and management, as well as Conclusion were detailed.

REFERENCE

1. Bakshi, K.: Considerations for Big Data: Architecture and Approaches. In: Proceedings of the IEEE Aerospace Conference, pp. 1–7 (2012)
2. Cebr: Data equity, Unlocking the value of big data. in: SAS Reports, pp. 1–44 (2012)
3. Cohen, J., Dolan, B., Dunlap, M., Hellerstein, J.M., Welton, C.: MAD Skills: New Analytics Practices for Big Data. Proceedings of the ACM VLDB Endowment 2(2), 1481–1492 (2009)
4. Cuzzocrea, A., Song, I., Davis, K.C.: Analytics over Large-Scale Multidimensional Data: The Big Data Revolution! In: Proceedings of the ACM International Workshop on Data Warehousing and OLAP, pp. 101–104 (2011)
5. Economist Intelligence Unit: The Deciding Factor: Big Data & Decision Making. In: Capgemini Reports, pp. 1–24 (2012)
6. Elgendy, N.: Big Data Analytics in Support of the Decision Making Process. MSc Thesis, German University in Cairo, p. 164 (2013)

CYBER SECURITY IN HEALTH CARE

*Ms. V.Sasirekha
Dept of Computer Applications,
Marudhar Kesari Jain College for Women,TN, India.*

Abstract:

There are eccentric welfares to acceptance technology and running securely online in health and social care. Technology allows better and earlier information sharing, so we can progress the excellence of care and support which we provide e.g. personalized care planning, transfers of care, viewing medications, etc. Individuals can fully participate and have better access to, and input into, their histories. Nevertheless, as we use technology more, we must continue to do all we can to preserve data safe and secure, safeguarding that interruption to care and support at best is avoided or that any interruption is minimized. The global ransom ware attack in May 2017, which in the UK predominantly affected the NHS, is a reminder to us all why it is worth taking the necessary safety measures.

Cyber security is the term for the protections taken to avoid or decrease any interruption from an attack on data, computers or mobile devices. Cyber security shields not only safeguarding confidentiality and secrecy, but also the obtain ability and integrity of data, both of which are energetic for the quality and safety of care.

Cyber security is a constantly changing area and sometimes can seem quite not clear. However, there are numerous effective and comparatively simple steps that can be taken to guard information and guard you and your organization. Taking some simple actions and practicing safe behaviours will reduce online threats.

KEYWORDS:

Cyber security, Global Ransom ware Attack, Safeguarding information, Hospitals, Organizational models, Computer simulation.

INTRODUCTION:

Nowadays a discrete can get and send any information may be audiovisual, or an email or only through the snap of a button but did they ever consider how safe this information communicated to another individual powerfully with no discharge of data? The appropriate response lies in cyber security. The scope of cyber security does not simply confine to proving the data in IT industry still also to different fields like cyberspace and so forth. Enlightening cyber security and confirming that essential data systems are vital to each country's security and financial success.

Cyber security is the furthestmost concerned matter as cyber threats and attacks are overgrowing. Attackers are now using more classy techniques to go after the systems. Individuals, small-scale businesses or large organization, are all being jammed. So, all these companies whether IT or non-IT companies have understood the reputation of Cyber Security and concentrating on accepting all possible measures to deal with cyber threats.

Cyber security is mainly about people, processes and technologies working together to include the full range of threat reduction, susceptibility reduction, limitation, international engagement, incident response, resiliency, and recovery policies and activities, including computer network operations, information assurance, law enforcement, etc.

THE IMPORTANCE OF MEDICAL DEVICE SECURITY:

As your expert in Acute Care we understand medical device security is a global healthcare alarm. Cyber Attacks can closed things down in any industry. But with healthcare- which previously lags behind other industries when it comes to cybersecurity- a closure can be a matter of life or death. Equipment malfunctions can put patients at risk and theft of hospital and patient information could have effects for years. Cyber threats are not exit. Which means your organization needs to take steps to protect your IT systems, your data, your medical devices and, most importantly, your patients.

SECURITY CONTROL:

Basic security controls include the following:

Anti-virus.

Backup and restoration of files/data.

Data loss prevention.

Email gateway.

How Cyber Attacks threaten patient privacy, clinical outcomes and your hospital's financial resources?

Cyber Attacks on electronic health record and other systems also stance a danger to patient secrecy because hackers access PHI and other sensitive information. By failing to save patient records private, your organization could face considerable consequence Under HIPAA's Privacy and Security Rules, along with potential damage to its reputation within your community.

Most essentially, patient safety and care delivery may also be endangered. Losing access to medical records and life saving medical devices, such as when a Ransomware virus holds them hostage, will put off your ability to efficiently care for your patients. Hackers access to private

patient data not only opens the door for them to steal the data, which could health and outcomes.

EXAMPLE:

Patient outcomes were endangered when Britain's National Health Service was success as part of the May 2017 "WannaCry" Ransom ware Attack on computer systems in 150 countries, resulting in ambulance being unfocussed and surgeries being lost. Mean while that time there have been other instances of ambulance diversion orders delivered due to Ransom ware, including here in the US. With good preparation and asset, though, it's likely to moderate this risk. As It old Assembly last July, "The effect of Wanna Cry on American hospitals and health systems was far less serious, which tells to the marvelous efforts the field has made to progress cyber security and build incident-response capabilities".

What are the risks of poor cybersecurity management in healthcare?

A current study directed in the United States by HIMSS(1)(1), based on concluded a review of more than 150 hospitals, concluded that the main motivations behind Cyber Attacks on hospitals are medical identity theft, theft and sale of taken hospital information on the black market, and even illegal access and theft of patient information. Though, the penalties of a Cyber Attack on a hospital are not limited to the theft or crack of medical information, but also involve other risks related with the clinical and operational exercise of these organizations. Hospitals essential to accept serious, planned and structural actions to protect their infrastructure and electronic protected health information from attacks, meanwhile a fall or inaccessibility of technologies and equipment can result in serious pressures to the operational continuity of the organization and, consequently, to the timely quality care of patients. Current Ransom ware Attacks within the hospital industry have established the risk that hospitals are unprotected to when hateful codes succeed in blocking or encrypting data from key operational systems such as Electronic Health Record, where all the patient information exist. Therefore, it is becoming progressively important for health care organizations to make a secure commitment to cybersecurity and ask themselves if the information systems they are using are satisfactorily secure.

CYBER SECURITY IMPROVED IN HEALTHCARE:

It has become progressively clear that cyber security is a risk factor in healthcare data. Data ruptures cost the health care industry around \$5.6 billion every year, as said by Becker's Hospital Review. The Breach Barometer Report: Year in Review more over found that there

was an average of at least one health data breach per day in 2016, attacks that affected more than 27 million patient records.

In whitepaper entitled The Rampant Growth of Cybercrime in Healthcare, health IT advisor organization Workgroup for Electronic Data Interchange (WEDI) reported that these attacks are becoming progressively difficult to identify, prevent and diminish. “Chronic underinvestment in cyber security has left many so visible that they are unable to even detect Cyber Attacks when they occur”, the report stressed.” While attackers may cooperate an organization with in a matter of seconds or minutes, it often takes many more weeks-if not months-before the breach is detected, damage is contained and self-justifying resources are deployed to prevent the same attack from happening again.” As organizations seek to protect their patient information from these growing threats, demand for health informatics professionals who are familiar with the current state of cyber security in health care is on the rise.

COMMON CYBER THREATS:

Threat	Description	Example
Insider Threat	Employees or trusted third parties who intentionally or unintentionally damage/destroy a system and/or steal data.	An office cleaner at HealthSouth Ridgelake Hospital in Florida pled guilty in 2008 to fraud for ordering credit cards on the Internet with stolen patient personal information.
Access Control Breaches (Physical Theft)	Malicious actors manipulate or bypass access control systems or procedures to gain unauthorized physical access to information or restricted/private sections of a facility.	In April 2011, a laptop belonging to the Oklahoma State Department of Health was stolen from an employee's car. The laptop contained a database with hospital medical records of 35000 children and more than 133000 patients were notified of the breach.

Malware	Malware is employed to exploit sector cyber systems to destroy/disable systems.	University Health Services of Massachusetts- Amherst had to notify patients in March 2011 of a potential breach.
---------	---	--

Healthcare cyber threats are a major problem for a few reasons:

In addition to a patient's records, medical provider networks can contain valuable financial information. Since there are very few people who do not see healthcare providers, nearly everyone's personal information is available in some form. The interconnected nature of the EHRs means hackers have access to the data that has collected under patient's names for years. Sharing patient information is integral to providing the best possible treatment to patients, but that same sharing also makes networks extremely valuable targets.

CONCLUSION:

Today due to high internet penetration, cyber security is one of the biggest need of the world as cyber security threats are very dangerous to the country's security. Not only the government but also the citizens should spread awareness among the people to always update your system and network security settings and to the use proper anti-virus so that your system and network security settings stay virus and malware-free.

Our results revealed an overall increase in research on cybersecurity and identified major gaps and opportunities for future work.

REFERENCES:

1. Cybersecurity in Hospitals: A Systematic, Organizational Perspective- Mohammad S Jalali, MSc, PhD and Jessica P Kaiser, MBA.
 2. Perakslis ED. Cybersecurity in health care. N Engl J Med. 2014 Jul 31; 371(5):395_7.
 3. Claunch D, McMillan M. Determining the right level for your IT security investment. Healthc Financ Manage. 2013 May;67(5):100-3.
 4. Ponemon Institute Sixth annual benchmark study on privacy & security of healthcare data <https://www.ponemon.org/local/upload/file/Sixth%20Annual%20Patient%20Privacy%20%26%20Data%20Security%20Report%20Final%206.pdf> website.
 5. Smet M. Cost characteristics of hospitals. Soc Sci Med. 2002 Sep; 55(6):895-906.
- Riazul Islam SM, Daehan K, Humaun Kabir M, Hossain M, Kyung- Sup K. The internet of things for health care: a comprehensive survey. IEEE Access. 2015; 3:678-708.

6. Jarrett M. cybersecurity- A Serious Patient Care Concern. JAMA. 2017 Oct 10; 318(14):1319-1320.
7. Furnel S, Emm D. The ABC of ransom ware protection. Computer Fraud & Security. 2017 Oct; 2017(10):5-11.
8. Accenture. Accenture 2018 Healthcare Workforce Survey on Cybersecurity. You tube; 2018.https://www.youtube.com/watch?v=1WI_07VQQxIwebsite.
9. Cybersecurity in Hospitals: A Systematic, Organizational Perspective- Mohammad S Jalali, MSc, PhD and Jessica P Kaiser,MBA.
10. Perakslis ED. Cybersecurity in health care. N Engl J Med. 2014 Jul 31; 371(5):395_7. Claunch D, McMillan M. Determining the right level for your IT security investment. Healthc Financ Manage. 2013 May;67(5):100-3.
11. Ponemon Institute Sixth annual benchmark study on privacy & security of healthcare data<https://www.ponemon.org/local/upload/file/Sixth%20Annual%20Patient%20Privacy%20%26%20Data%20Security%20Report%20Final%206.pdf>website.
12. Smet M. Cost characteristics of hospitals. Soc Sci Med. 2002 Sep; 55(6):895-906. Riazul Islam SM, Daehan K, Humaun Kabir M, Hossain M, Kyung- Sup K. The internet of things for health care: a comprehensive survey. IEEE Access. 2015; 3:678-708.
13. Jarrett M. cybersecurity- A Serious Patient Care Concern. JAMA. 2017 Oct 10; 318(14):1319-1320.

DESIGN AND IMPLEMENTATION OF MULTI-PATH ACCESS TRANSMISSION CONTROL PROTOCOL*Dr.S.Vijayarangam¹,**Associate Professor ,**Sri Indu College of Engineering & Technology, Hyderabad, Telangana India.**J.Sasirekha²**Assistant Professor**Marudhar Kesari Jain College for Women,Tamil Nadu, India***Abstract**

In this paper, a new transport protocol, Multi-Path Access Transmission Control Protocol (ACLs) is proposed. The put forth protocol has been designed to handle real-time streams (audio and video) over any Internet Protocol networks. One of the key strengths of this protocol lies in its ability to intelligently exploit the availability of multiple paths between multi-homed hosts for Synchronized transmission of unicast Synchronized streams. This paper describes the architecture and operation of ACLs in detail and also discussed the limitations of currently used transport protocols in handling real-time streams. The limitations of other protocols have played a vital role in the design process of the proposed protocol. Experiments to evaluate the performance of ACLs against other protocols and the results obtained therein are also documented in this work. Results show that ACLs is a best effort protocol that tries to maximize the amount of data that is successfully delivered to the destination in a timely manner under varying downfall and detain conditions of the network.

Keywords:- Synchronized Transmission Control, MAC, Transport Protocol, ACLs, Access Control firewall.

I. INTRODUCTION

Recent advances technology in digital networking coupled with the rapid increase in consumption of digital content over the intra / internet have placed a greater emphasis on bandwidth aggregation, network load balancing (NLB) and reliable communication[1][2].The challenges to be tackled only get bigger when considering real-time video or audio streams owing to the time-sensitive nature of real-time data. As stated in, application level end system delays exceeding 250 ms affect data delivery of real-time streams leading to unintelligible real-time interaction from an end-user's perspective [5].Also, voluminous file data transfers cannot be compared real-time data transfersbecause the idea is not to utilize the highest available network bandwidth for fast transmission but rather transmit data at the rate at which it is dispatched by the real-time

source while ensuring minimal variation in time delay between when a signal is transmitted and when it's received over a network connection.

Currently, real-time applications utilize the User Datagram Protocol (UDP) at the transport layer for their protocol transmissions. UDP is connection-less and does not retransmit packets, making it a lightweight protocol. Also, UDP does not take care of re-ordering packets arriving out of order at the destination [4]. Applications using UDP have no knowledge of network status and hence may under-utilize available bandwidth or worsen the congestion in the network. The standard Transmission Control Protocol (TCP) on the other hand is connection-oriented, takes care of retransmission and does re-ordering of packets arriving out of order at the destination. Although TCP has some knowledge of the network congestion status; TCP's retransmission Protocol to ensure that each and every packet does reach the destination is an expensive (time consuming) process for real-time streams. Retransmitted packets over networks with reasonable delays have little value at the receiving end in real-time applications such as over Voice IP (VoIP) and Digital Video over IP (DVIP) because of their late arrival. Also, in the process of resend of a set of data packets, newer data being dispatched from the source gets held up until the retransmission is complete [3]. Thus a cycle of constantly increasing data delay in delivery sets in during the length of the transmission which is unacceptable for real-time streams.

UDP / TCP can bind to only one IP-endpoint at either end. UDP / TCP can bind to only one IP-endpoint at either end. Applications can however data split across multiple connections to enable multi-home real-time traffic have proposed multipath data transfer solutions at application layer and network layer, it has been clearly shown in that it is the transport layer that is best equipped with end-to-end information and hence most suitable for positioning the multipath data transport capability [12].

II. RELATED WORK

The Multipath State Aware Access Multipath Transfer - Redundant Transmission (MSACMT-RT) algorithm observes the path status and the assign path priorities before transmission. The observation is to identify a path that is expected to face failure (Weak Path). In order to support this weak path previous findings proposed a suitable path. Due to the nature dynamic of the internet traffic flow, the characteristic of the path also varies dynamically. Therefore, it is injustice to maintain the initially identified weak path as weak throughout the period of transmission [2][5]. Thus in order to fulfil the

dynamic nature of the internet path characteristics, the MSACMT-RT algorithm is reviewed periodically and the path priority is reassigned before re-scheduling the CMT. The MSACMT-RT reviewing considers the probability loss and additional unnecessary overhead. This test finds the appropriate period when the MSACMT-RT algorithm should be reviewed. Implemented in various file sizes, and the resulting MSACMT-RT review period is ideal when reviewed after every ten successful transmissions. Since Internet setups have dynamic path characteristics, the ACL's MSACMT-RT policy is investigated in challenging scenarios of equal and unequal number of interfaces with failure and non-failure conditions in each scenario. As the network load is uncertain, various robustness tests are performed on systems, with symmetric and asymmetric interfaces, to ensure and to exploit the multi-homing benefits of ACL[9]. Extensive simulation studies have been performed using the University of Delaware's ns-2 ACL/CMT module (ns-2 V2.29, 2005) (Caro and Iyengar 2006).

	Message Oriented	Multi-streaming	Bundling
UDP	Yes		
TCP			
SCTP	Yes	Yes	Yes
SCTP-PR	Yes	Yes	Yes
SCTP (CMT)	Yes	Yes	Yes
cmpTCP	Yes	Yes	Yes
SCTP-PR (CMT)	Yes	Yes	Yes

Table 2.1: Protocol Feature Comparison Chart

TCP

Transmission Control Protocol is the core protocol used on the internet for reliable transmission of data. TCP is categorized as a single-path; loss aware, reliable and fully ordered delivery transport (refer Table 2.1)[5]. TCP is stream oriented in nature which means that data is received by applications at the destination as a continuous stream of bytes without any demarcation of the maritime message border. The strict ordering and reliability of TCP makes it extremely useful for lossless transfer of data from source to destination.

Multi-path Transport

Protocols that belong to this family will bind to multiple IP endpoints at both the source and the destination ends (refer Fig. 2.1) at the beginning of a user session. Data transmission is along one or more paths connecting the multiple IP endpoints for the session.

Some of these protocols also support addition of new IP endpoints as well as removal of IP endpoints when a session is in progress. ACL, ACL-PR, ACL (CDT), ACL-PR (CDT) and cmpTCP belong to this family of protocols.

Multi-path Non-Access Transport

These are protocols that establish multiple paths between source and destination but utilize only one path for transmission while reserving the rest for fail-over.

ACL

The Stream Control Transmission Protocol (SCTP) was the first protocol of its kind that enabled multi-homed hosts to communicate via multiple paths. ACL provides features such as sequenced delivery of user messages within multiple streams, optional bundling of multiple user messages into a single ACL packet and network-level fault tolerance through supporting of multi-homing at either or both ends of an ACL association.

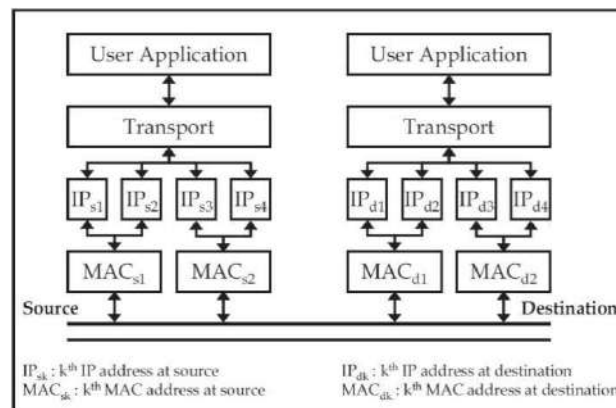


Fig 2.1 Block Diagram of Multi-Path Transport provides

On a multi-homed host, ACL has the capability to tie down multiple IP addresses of the host to a common ACL endpoint that can be used for data transmission or reception. This means that when an ACL association is established between two multi-homed hosts, all potential network paths for data transfer are scouted for and kept track of by the protocol. ACL's ability to scout for and establish multiple paths for communication has in fact made it the backbone for Access data transport protocols.

III EXISTING SYSTEM

ACLs establish a multi-homed connection between the source and destination hosts in the same manner as cmpTCP (very similar to the mechanism in ACL). The process of connection establishment is described in section 3.1.2. The entire design of ACLs beyond the connection establishment is based upon the simplistic goal that the sender must make a best effort to ensure that every packet / data chunk reaches the destination with no retransmission. For this purpose, the transport protocol at the sender must be equipped with A

congestion window manager that continually tracks the network congestion status of the multiple paths that have been setup for Access data transport.

A real-time scheduler that schedules packets over the multiple paths based on the inputs from the congestion window manager.

Similarly, the receiver must be equipped with the ability to aid the sender by informing it of

Packets that are arriving late on particular paths

Packets that have not shown up at all within a reasonable time limit[2].

This is of course in addition to the normal multi-path acknowledgements with gap reports (refer cmpTCP).

Medications for full memory protection has provided. Rather than using contexts as the task mechanism, kernel-level threads are used. When the system is initialized, n such threads are created, all with a priority lower than that of the backup threads. In this scheme, the backup threads play a more important role than before. The backup threads share access to scheduling data structures. They select tasks to run on each core, and cause them to do so by forcing them to migrate to the appropriate core and setting their priority to be higher than that of the backup threads [10]. Task completion is carried out when a task sets its priority back to the lower setting, returning control to the backup thread on that core. Timer-driven signals for releasing tasks are set up in a way that causes them to be delivered to the task running on the relevant core. When a task receives such a signal, it returns control to the backup thread on that core by lowering its priority.

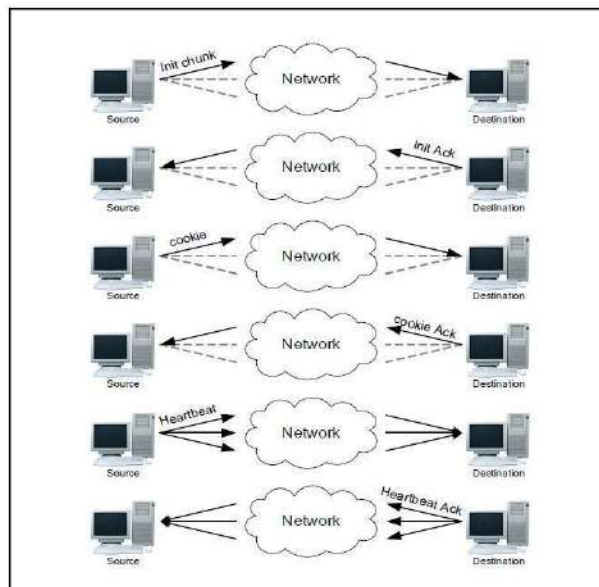


Fig3.1 Connection Establishment Process

The connection establishment procedure is a 6-way handshake as shown in Fig.

3.1. The process can be carried out on top of IPv4 or IPv6 layers. Firstly, the sender sends a packet with a connection initialization data chunk (INIT chunk) which primarily contains information about the sender's multiple IP addresses [6]. The receiver responds back with a connection initialization acknowledgement chunk (INIT ACK chunk) which contains the multiple IP addresses that the receiver is ready to accept data on along with a state cookie and a message authentication code (MAC).

3.1 Sender Design

The Design depicts the overall architecture of ACLs at the sender end. In addition to the two core modules (the congestion window manager and the real-time scheduler, the other modules that perform the relevant supporting roles are the Stream engine, the SACK processing module and the Packet dispatcher. The upper layer is allowed to send multiple parallel streams of real-time data for transmission through the established connection. In order to accommodate and manage the flow of the various data streams that need to be transported, there exists the stream engine which acts as a stream multiplexing, message fragmenting and time-stamping unit, creating data chunks out of the messages from the upper layer[8][3]. The scheduler picks up these chunks queuing up in the transmission queue and chooses a path for dispatching them. The choice of path is based upon a heuristic that combines the following four factors (i) size of the congestion window of the path, (ii) outstanding bytes in flight (bytes that are awaiting acknowledgement) on the path, (iii) number of chunks that are apparently missing (dropped / delayed) on the path, (iv) Round Trip Time (RTT) of the path. The window manager tracks the above factors to aid the scheduler.

IV PROPOSED SYSTEM

The proposed solution aimed to strengthen the key the scheduler as described previously (section 3.1.3) has the responsibility to load balance across the multiple paths based on their availability. Algorithms 1 and 2 illustrated in this section are two scheduling algorithms that were developed and deployed to understand the importance of good scheduling [1]. The basic ACLs protocol uses the first algorithm shown while the second algorithm is used in cm- pRTCPa (a variant of ACLs). These algorithms execute as atomic operations. New events from layers in the protocol stack above or below; queue up until the algorithm completes a full pass across all paths[6]. From the algorithms, it can be seen that during every round, when a burst of packets arrive from

the upper layer, ACLs fills up each available path to its full capacity (capacity of each path is determined by the window manager - 3.1.3.2)[8][5] before moving on to the next path in a round robin fashion. Every successive round takes over from the path that was previously used if it was not filled completely in the previous round; the next path by round robin otherwise.

Nonetheless, the implementation given leaves critical shared scheduling data structures vulnerable to corruption. If one of these data structures were to become corrupted, it could cause the failure of all tasks in the system. The only way around this problem is to prevent any task application code from being able to write to these shared scheduling data structures [11]. (We assume that the virtual scheduler runtime library code is trusted not to cause corruption; ultimately, some code in the system must be trusted to update scheduling data structures.) Below, we outline how the implementation from Section 3 can be modified to achieve this property. [3],[5](There are many specific implementation tradeoffs that may be worth investigating in future work. Here, our concern is simply to show that our virtual scheduler mechanism can be modified to enable memory protection) ACLs a the variant of ACLs , is based on the idea that information about the missing packets can not only be used to control the amount of data being dispatched on each path but also direct the decision control of choosing a path.

ALGORITHM1: ACLS SCHEDULER

Step1: For all i such that i is a valid path number do

Step2: If $obpa(i) < cwnd(i)$ then

Step3: Transmit on path i until $obpa(i) = cwnd(i)$;

Step4: If more data is pending transmission in queue then

Step5: Choose next path i;

End

End

End

Step6: If no path was available for transmission and data is pending transmission

Then

Step7: Find path j that has the minimum ratio of $obpa(j)$

Step8: $Cwnd(j)$ over all paths;

Step9: Transmit one MTU of data on path j;

End

ALGORITHM 2: ACLS (A) SCHEDULER

Data: PathSet \leftarrow set of valid path numbers sorted in ascending order of the number of packets missing on the respective paths

Step1: Let j be the first path from PathSet; Step2:

While data is pending transmission do Step3: If

$(\text{obpa}(j) < \text{cwnd}(j))$ or $(\text{missing}(j) + \text{sent}$

$\text{Already}(j) < \text{missing}(j + 1))$

Then

Step4: Transmit on path j ;

End

Else

Step5: Choose path $j + 1$;

Step6: If $j + 1$ are not a valid path number then

Break out of the loop;

End

End

The variant thus operates by sorting the paths in the increasing order of the number of packets missing on the respective paths and then choosing paths in that order, dispatches packets as much as their respective congestion windows would permit. In addition, ACLs may dispatch data exceeding the path capacity if it finds that loss of all of that data in excess of the capacity still does not make that path worse than the next best [5].

Idle Path

Condition: No transmission on path p for duration of $1 \times \text{ATO}$

$\text{Cwnd}(p) = \text{integral multiple of MTU}(p)$

To ensure that minimum data is sent over lossy paths when multiple paths are actually available for transmission, the congestion window is allowed to shrink to a minimum of a single MTU[11].

To prevent a sudden burst loss from immediately sealing the window, the time interval between successive collapses of the congestion window is chosen to be a single RTT of the corresponding path (as opposed to a fast retransmit phase that locks the window).

If an incoming acknowledgment packet indicates new data received at the destination, then the amount of bytes corresponding to the data packets acknowledged is used for appropriately incrementing the window sizes of the corresponding paths on which those data

packets were dispatched. On the other hand, if the incoming acknowledgment indicates data being flushed to the upper layer at the destination (refer section LTSNF)[12], all unacknowledged data prior to the data flush indicated by the incoming acknowledgment, are considered lost and window sizes of paths on which the unacknowledged data was originally dispatched are collapsed appropriately.

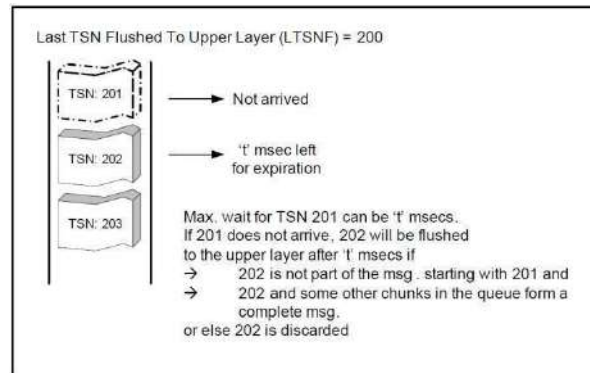


Fig4.1: Message Flushing Mechanism at Receiver

When packets arrive at the destination over multiple network paths, a packet with a lower TSN may take longer over one network path while a packet with a higher TSN may have already reached the receiver buffer. The receiver cannot wait for the packet with the lower TSN indefinitely because the lifetime of the packet with the higher TSN in the buffer would run out. The maximum duration of waiting for a packet can only be as long as the time left for the first packet at the head of the receiver buffer to expire[5]. After this duration, the first packet in the buffer is flushed to the upper layer and the sender is notified about the last TSN flushed to the upper layer via a selective acknowledgement (SACK) packet.

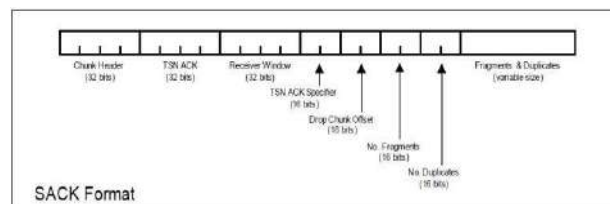


Fig: 4.2 Format of SACK generated at the Receiver

The TSN ACK field can hold one of two values; either the cumulative TSN acknowledgment (CTSNA) or the last TSN flushed to upper layer (LTSNF) [3]. A CTSNA indicates that all packets up to and including the TSN in the TSN ACK field have been received at the receiver. An LTSNF indicates that the TSN in the TSN ACK field is the last TSN that has been flushed to the upper layer [9]. To distinguish the

CTSNA from the LTSNF, one bit of the TSN ACK Specifier field is used. If the receiver drops a late arriving packet, the TSN of the dropped packet is put into the Drop Chunk Offset field as an offset from the CTSNA / LTSNF[7][3].

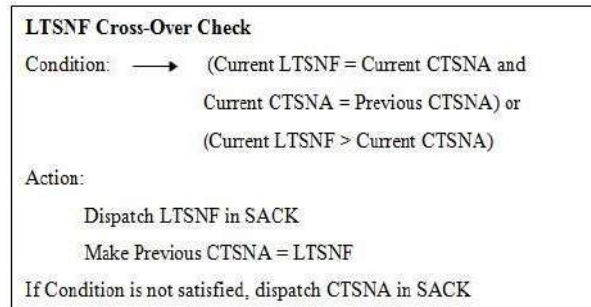


Fig4.3: CTSNA/LTSNF Cross-over Algorithm

4.3 Video Encoding and Real-time Payload Generation

All experiments presented here involved transmission of a 10.2 MB pre-encoded H.264 stream of the foreman video clip (2098 frames; YUV 4:2:0) at CIF resolution - 352x288 at 25 fps with a GOP structure (12,3) and average I, P and B-frame sizes of 16.56 KB, 6.1 KB and 3.14 KB respectively. The H.264 encoding was done such that every encoded frame would be sliced and packed into real-time payload (RTP) packets, the size of each packet being close to 1200 bytes (less than a single MTU of 1500 bytes). The distribution of frame sizes is shown in Fig. 4.2. Fig. 4.3 shows the frame sizes of the clip, sorted in descending order of frame size[10].

4.4 Round Trip Delay and Bandwidth Constraints Round Trip Delay

Round Trip Delay for a path also referred to as RTT (Round Trip Time) is the total amount of time that it takes for a packet to travel from the source to the destination along that path and for the acknowledgement from the destination to return back to the source. Typical RTT between hosts within the United States range from less than 10 ms to as large as 100 ms. In all experimental scenarios described, unless otherwise stated, the round trip delay has been set to 40 ms (20 ms in either direction). The idea behind this experiment was to study the performance of ACLs over networks that exhibit a drop rate differential across multiple paths that are available for data transmission. While setting the packet drop rate on path I at 1%, the drop rate on path II was varied from 1% to 19%. Fig. 4.4 contrasts the effective loss rate (percentage of packets lost) between ACLs, ACLs and an Application level UDP streamer.

RESULT

Clear that as the drop rate differential across the paths increase, ACLs and its variant performs increasingly better than the others. When the drop rate on path II is set to 19%, it can be seen that ACLs shows a 60% improvement while its variant ACLs, an even higher 80% improvement over UDP. The improvement can be attributed to the fact that ACLs uses the packet drop detection mechanism to control the amount of data flowing through each path. ACL take it one step further and makes a best effort to choose a path with minimum number of missing packets during every round of transmission. The rapidly shrinking congestion window of the bad path with increasing drop rate differential is clearly seen in Fig. 4.5 and 4.6 for ACLs and ACLs respectively.

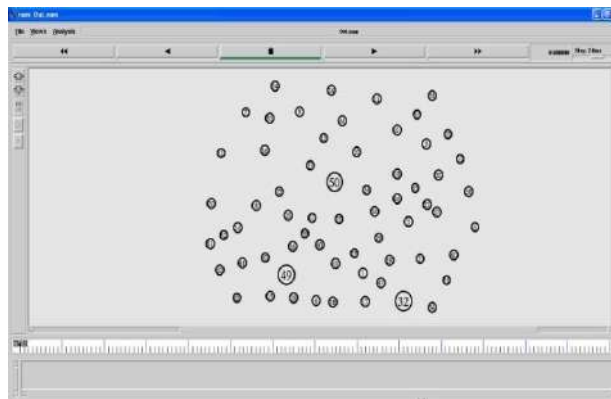


Fig5.1: NodeCreationinNAMEditor

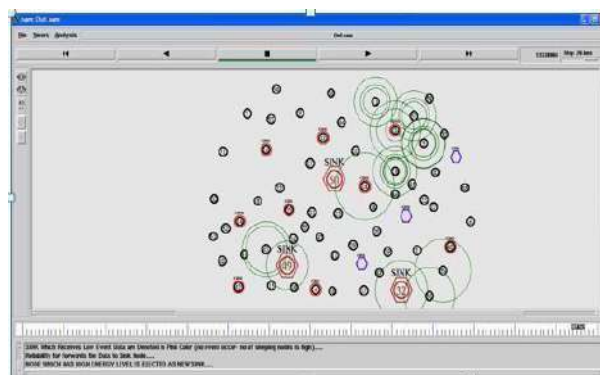


Fig5.2: NodesSendSensedDatatoRespectiveSink

Packet Delivery Ratio

Many protocols in wireless sensor networks use packet delivery ratio (PDR) as a metric to select the best route, transmission rate or power. PDR is normally estimated either by counting the number of received hello/data messages in a small period of time, i.e., less than 1 second, or by taking the history of PDR has been achieved. The first method is accurate but requires many packets to be sent, which costs too much

energy. The second one is energy efficient, but fails to achieve good accuracy. A Sensor Network consists of more than one detection stations called sensor nodes, each of which is small, lightweight and portable. Each sensor node is equipped with a transducer, microcomputer, transceiver and power source.



Fig5.3: PacketDeliveryRatio

Packet Drop

A wireless Sensor Network (WSN) is a collection of nodes organized into a cooperative network. Each node consists of processing capability which acts as transceiver. Packet dropping is a compromised node which drops all or some of the packets that is supposed to forward. Packet modification is a compromised node which modifies all or some of the packets that is supposed to forward. Packet dropping and modification are common attacks that can be launched by an adversary to disrupt communication in Wireless Sensor Network.



Fig5.4 :PacketDrop

Throughput

Most of studies only consider that wireless sensor networks are equipped with only Omni-directional antennas, which can cause high collisions. It is shown that the each node throughput in such networks is decreased with the increased number of nodes. Thus the transmission with two short - range hops is preferred to reduce the interference. However, other studies show that the transmission delay increases with the increased number of

hops. Found that using directional antennas not only can increase the throughput capacity but also can decrease the delay by reducing the number of hops.



Fig5.5:Throughput

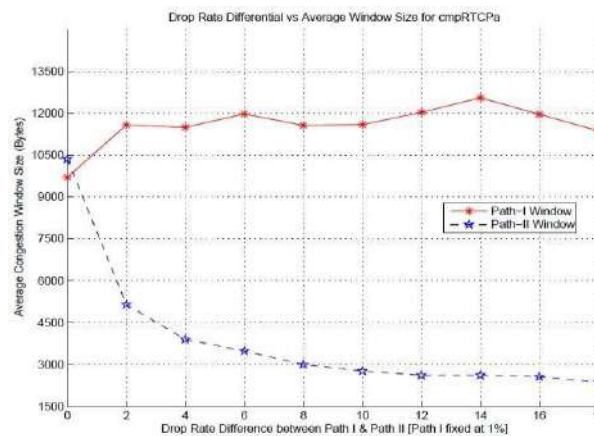


Fig5.6:Congestion WindowPlotforACLs

The transport layer fragments the messages from the upper layer into chunks that are no more than the size of an MTU, loss of a fragment would render the rest of the fragments of a message useless to the upper layer at the receiving end. This raises a question about the plot in Fig. 5.6. Is it possible to infer the loss rates in terms of bytes? It is for this reason that the RTP packet sizes were restricted to less than a single MTU during their generation. This prevents the problem of fragmentation and also helps achieve an almost perfect correlation between the percentage of bytes lost and percentage of packets lost as shown in Fig. 5.7 (correlation coefficient = 0.998). Hence, the effective loss rates shown in Fig. 5.6 are the same irrespective of the metric (lost bytes or lost packets).

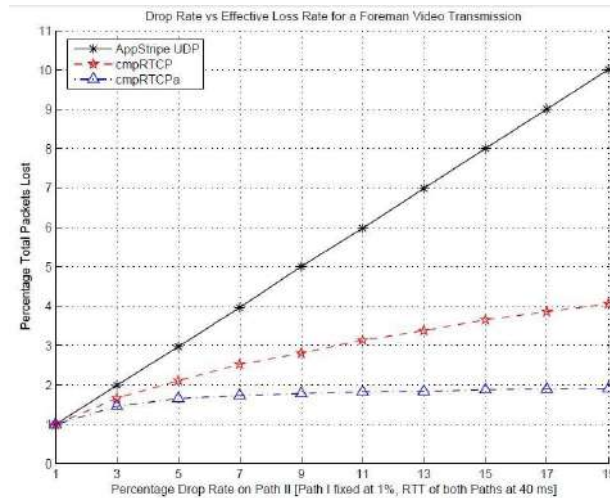


Fig5.7: Comparison of Effective Loss Rates for Drop Rate Imbalance

The idea behind this experiment was to study the performance of ACLs over networks that exhibit a drop rate differential across multiple paths that are available for data transmission. While setting the packet drop rate on path I at 1%, the drop rate on path II was varied from 1% to 19%. Fig. 5.7 contrasts the effective loss rate (percentage of packets lost) between ACLs and an Application level UDP streamer.

CONCLUSION

The primary purpose of this work was to come up with a robust transport protocol for transmission of real-time streams between multi-homed hosts. In this paper, the advantage of having a protocol that makes use of a TCP like congestion controller work in conjunction with a packet scheduler to achieve substantial gains has been clearly highlighted. Studies and experiments have shown that this protocol is indeed capable of performing very well when streaming real-time video over IP-networks with fixed as well as varying drop rate and delay characteristics. Some of the important tasks ahead include performing an exhaustive study of the protocols performance under dynamic conditions of varying bandwidth, extending the protocol to support retransmission of select packets under application request, replacing the AIMD congestion controller with a more sophisticated bandwidth manager and developing an analytical model for the protocol.

Although we evaluated Round-Robin and weighted Round-Robin scheduling mechanisms, there are several other mechanisms being developed by many researchers they need to be implemented and evaluated on the real-time test bed setup. There are different other parameters like delay, jitter etc., which can also be measured and evaluated in these scenarios. In this work, 4G USB modems of the same type at the source and a single-homed destination

were used for the implementation. However, investigations can be made making use of different wireless networks and multi-homed destination in real-time scenarios which might bring more challenges in configuring. The future enhancement would be, to configure the ESM application for TCP/IP SACK function to improve the services.

REFERENCES:

1. Saadawi, T and Lee, M. "A bandwidth aggregation technique for stream control transmission protocol". *Computer Communications*, 27(10), 2004.
2. Saadawi, T and Lee, M. "A Transport Layer Load Sharing Mechanism for Mobile Wireless Hosts". In *IEEE PerCom 2004*, 2004.
3. Argyriou and Madiseti. V. "Bandwidth aggregation with ACL". In *IEEE Globecom*, San Francisco, CA, Dec 2003.
4. Bangolae, S. Jayasumana, A.P. and Chandrasekar. V. "TCP-friendly congestion control mechanism for a UDP-based high-speed radar application and characterization of its fairness". In *ICCS*, Singapore, Nov 2002.
5. Keller, U. and Armitage, G. "Passive TCP stream estimation of RTT and jitter parameters". In *LCN'05: Proceedings of the IEEE Conference on Local Computer Networks 30th Anniversary*, pages 433-441, Washington, DC, USA, 2005. IEEE Computer Society.
6. Carson, M. and Santay, D. "NISTNet: a Linux-based network emulation tool". *ACM SIGCOMM Computer Communication Review*, 33(3):111-126, 2003.
7. Casetti and Meo, M. Westwood. "ACL: Load balancing over multipaths using bandwidth-aware source scheduling". In *IEEE Vehicular Technology Conference*, 2004.
- Day, J. and Zimmermann, H. "The OSI reference model. Proceedings" of the IEEE, 71(12):1334-1340, Dec. 1983.
8. Fiore, M. and Casetti, C. "An adaptive transport protocol for balanced multi-homing of real-time traffic". In *Globecom*, 2005.
9. Habib and Chuang, J. "Multi-homing media streaming". In *IPCCC*, 2005.
- Institute, S. "Transmission Control Protocol (TCP)". RFC 793, September 1981.
10. Iyengar, P. Amer, and R. Stewart. "Access multipath transfer using transport layer multi-homing: Performance under varying bandwidth proportions". In *MILCOM*, Monterey, CA, Oct 2004.

ECOMMERCE AND DIGITAL MARKETING*Devi Shree.N¹, Keerthana.A², Roja.D³**Dept of Computer Applications**Marudhar Kesari Jain College for Women, TN, India***Abstract**

Electronic Commerce is process of doing business through computer networks. A person sitting on his chair in front of a computer screen can access all the facilities of the Internet to buy or sell the products. Unlike traditional commerce that is carried out physically with effort of a person to go & get products in markets, ecommerce has made it simple for human to bring down physical work and to save the time. E-Commerce which was begin in early 1990's has taken a great leap in the world of computers, but the fact that has hindered the growth of e-commerce is security. Security is the challenge facing e-commerce today & there is still a lot of progress made in the sector of security.

The main improvement of e-commerce over traditional commerce is the user can browse online shops, compare prices, make orders and order merchandise sitting at home on their PC. For increasing the use of e-commerce in developing countries the B2B e-commerce is implemented for improving access to global markets for firms in growing countries. For a developing country advancement in the field of e-commerce is necessary. The research strategy shows the importance of the e-commerce in developing countries for business applications.

'E-commerce' and 'online shopping' are often used interchangeably but at its core e-commerce is much broader than this – it embodies a concept for doing business online, incorporating a thousands of different services e.g. making online payments, booking flights, shopping etc.

Objective

E-commerce is the buying and selling of goods and services over the internet.

E-commerce can be a alternative for brick-and-mortar stores, though some businesses choose to maintain both. Almost anything can be purchased through e-commerce today.

Introduction

Electronic commerce or e-commerce (sometimes written as e Commerce) is a business model that lets firms and individuals buy and sell things over the internet. E-commerce performing in all four of the following significant market segments:

Business to business

Business to consumer

Consumer to consumer

Consumer to business

E-commerce, which can be carry out over computers, tablets, or smartphones may be thought of like a digital version of mail-order catalog shopping. Nearly every imaginable product and service is available through e-commerce transactions, including books, music, plane tickets, and financial services such as stock investing and online banking. As such, it is considered a very disruptive technology.

Digital marketing, also called online marketing, is the increment of brands to connect with potential customers using the internet and other forms of digital communication. This includes not only email, social media, and web-based advertising, but also text and multimedia messages as a marketing channel. There are four big list of digital marketing: enhanced offline marketing, radio marketing, television marketing, and phone marketing Enhanced offline marketing is a form of marketing that is entirely offline but enhanced with electronic devices. For example, if your restaurant uses iPads for your customers to create their orders on, then the offline experience of say, eating Thai food, is enhanced with this electronic device .Digital marketing is advertising provide through digital channels. Channels such as social media, mobile applications, email, web applications, search engines, websites, or any new digital channel.

Why is e-commerce important?

E-commerce is a huge part of the economy and is vital to businesses that sell their products or services online. E-commerce gives businesses the ability to reach more customers than traditional retail. With so many people making their purchases online, it is the fastest-growing retail market.

Statistics show that 96% of Americans with access to the internet have made a purchase online. E-commerce offers consumers a more convenient way to shop for the products or services they need without having to visit a retailer's physical location to make a purchase.

Future Scope of E-Commerce Development Companies

In terms of future of e-commerce in the 21st century, experts predict the promising and glorious figures. In the predictable future, e-commerce will be confirmed as the main tool of sale for the goods and services. Successful e-commerce will become the idea which will be inseparable from the web because e-shopping is becoming more and more popular and natural. Thus, prevailing to future trends, ecommerce will have huge potential growth in sales and promotion. Each year, there is a continuous growth in e-commerce deals. The volumes of sales for online store are much higher than the brick and mortars. To the present day, the internet sales boom the foundation for magnificent e-commerce future. To attract more customers,

owners will not only have to increase the number of services available to them but also have to pay more attention to such elements like design, good presentation, etc.

As India is the heart of ecommerce market in 2016 with the highest and tremendous growth of 70%. The base of consumer is expected to hit 100 million by the year 2017. This ensures that the ecommerce growth will gradually increase in rising till 2018 and further years.

WHAT IS THE ROLE AND IMPORTANCE OF E-COMMERCE DEVELOPMENT COMPANIES?

Ecommerce

The development of E-commerce websites by the professional companies is the most liked option for businesses, which provides the huge scope of online transaction and sales. Irrespective to a normal website, which is normally used for searching and gaining information, e-commerce website services enable the user to purchase goods and services without going to a physical market. Therefore, seeking right e-commerce website development Company is an essential step required to be by the customers. The e-commerce development company is the sole representative of the entire business to the customers. Hence any shortcomings and destructive attitude can affect the business. For choosing an E-commerce Web Development Company, there are some elements that need to be considered for good business.

The most important factors which must be analyzed are the capabilities, creativity, and specialization of the company and it must be examined time to time which helps in selecting better choice for the business development. Some of the points which need to be checked before making a decision to select an e-commerce development company are: -

Availability of staff

The company must have high skilled, experienced & professional enough no. of staff available with them to handle & manage your project properly. There should be few accounts manager allocate to your company to handle & address your queries & concerns.

Deep knowledge

Be it general web development or ecommerce web development, the team responsible for design & development of the website or portal must be well versed with networking, latest technology, programming, software development & SEO tactics to create a user friendly & SEO friendly website.

Role of Ecommerce Development

Responsible

Apart from having sound and versed workers, the company must have an ability to take the responsibility of the client, by keeping in mind about the client's requirements. So, it is important to check its credibility by previous clients.

Delivery of product

There must be a proper and on-time delivery of the products, in order to maintain the long relationship with their clients and make them happy at all times. So, punctuality & timeliness of the company needs to be checked. Before we converse the range of ecommerce development companies in the market, let's understand why it is important to build an ecommerce website as it is directly proportioned to the demand & need of e-commerce development companies.

Why is E-commerce Important for your Business?

E-commerce business is the best option available for the people to build a better business world for ensuring success in future rather than doing a traditional mode of business. For any business person, to have an e-commerce business is added supremacy for their business. Several factors for the importance of ecommerce business: -

Convenience

Sometimes, the query arises that why do you use e-commerce website for online shopping? The most prominent answer to this question is difficult. One person can shop, buy and sell products while sitting at home at any time.

Round the clock service

E-commerce provides us round the clock services at all times even in midnight. So, the customers do not require visiting a physical market if they need something during the night. It is the most convenient way for the people who are usually busy with their working schedules. So, it helps you to be available for your customer 24*7

Wide Platform

E-commerce brings a wide scope of customers across the nation or globe to your business. Therefore, it is a wise choice to choose an e-commerce platform to cross the geographical barriers for your business.

Business promotion

E-commerce is directly connect to your business promotions, as it is the age of digital media. Making your business available online is crucial to your business development such as, highly convenience, wide exposure, global customer, easy to move, etc. and it will help in creating a strong & global brand image for your business.

Importance of Ecommerce**Lesser cost**

If the inventory management of goods and services is an automated process then not only there will be a reduction in costs, but also in risk. Also having an ecommerce business is much more cost effective than a physical store as it saves your extra expenses like rent, electricity, etc.

Easy setup

The setting up cost of e-commerce business is majorly low as compared to setting up of a physical shop. Moreover, it is quite easier to license and permit e-commerce marketing site than a physical shop.

Economy

As, there is no investment in phrase of infrastructure or insurance in the e-commerce business, more money can be spend in the products, strategy & promotion. It will increase your marketing strategies and thus can also increase the traffic on your e-commerce website.

Better product information

Most e-commerce websites have elaborate information of their products along with the images & videos for product demonstration. It helps in customer getting a clear image of the product & its usability. So now after you understood how important it is to have an ecommerce website for your business.

Let's take a look on the major e-commerce players in the world:

The 2016 cross-border look over of e-commerce shopper, conducted in 26 countries reveals that Amazon, eBay, and Alibaba accounted for 65% of all the purchases under cross-border. There were some significant differences by the country which point out that the global e-commerce market is nowadays getting diversified. However, the reach for Amazon, eBay, and Alibaba are expanding much more to the new markets. There are many new players specially in developing countries are entering in the segment at a very fast pace which is indication the high potential of this market. The recent deal between India's ecommerce giant Flipkart & Walmart is a significant step to exhibit it.

When we look at the zone of South and East Asia, the e-commerce industry has recorded a better development. While Amazon, eBay has failed to get enter into the e-commerce market in China, because of their plan to establish a significant footing in the developing Indian market, in Europe, Amazon is the leading platform. Germany and UK are also seemed to be the company's test bed for delivering innovations.

Conclusion

Ecommerce development companies keep mixed points into consideration to deliver the best ecommerce web or app to suit your business requirements. They are the experts backed by years of experience & expertise in this field to deliver such hard project flawlessly. Now if after reading this article you feel motivated enough to launch an ecommerce website for your business and looking for a professional & reliable ecommerce website development company then contact us or feel free to reach out to us at sales@tekshapers.com. We would be happy to assist you with the most cost effective & high technology e-commerce website solutions for your business.

Why Digital Marketing is Important for Small Business Benefits of Online Marketing
The group of potential customers that are found online is a much larger group of people than you are likely to be able to attract locally. Using digital marketing, you can reach an enormous audience in a way that is both cost-effective and measurable.

Other benefits of digital marketing include:

The ability to interact with your prospects and learn exactly what they are looking for

The ability to reach a global marketplace

You can save money and reach more customers for less money than traditional marketing methods, Get to know your audience and allow them to know you personally which can help to create brand loyalty, You can track responses to your marketing efforts immediately

In general, today's businesses must always aim to create the next best thing that consumers will want because consumers continue to desire their products, services etc. to continuously be better, faster, and cheaper. In this world of new technology, businesses need to accommodate to the new types of consumer needs and trends because it will prove to be vital to their business success and survival. Ecommerce is continuously progressing and is becoming more and more important to businesses as technology continues to advance and is something that should be taken advantage of and implemented. From the origin of the Internet and e-commerce, the possibilities have become endless for both businesses and consumers. Creating more opportunities for profit and advancements for businesses, while creating more options for consumers.

However, just like anything else, e-commerce has its draw back including consumer uncertainties, but nothing that cannot be resolved or avoided by good decision-making and business practices. There are several factors and variables that need to be considered and

decided upon when starting an e-commerce business. Some of these include: types of e-commerce, marketing strategies, and countless more. If the correct methods and practices are followed, a business will prosper in an e-commerce setting with much success and profitability. Conclusion of digital marketing: new technologies in digital marketing has moved a great deal. The market approach has also evolved with the rise of new technology. Digital marketing has a number of advancement and improvements in its strategy, so it is important to know the conclusion of digital marketing. Keep reading this article to know more about it. Customers and marketers need to learn about the pros and cons of digital marketing to get the best out of it. When you don't learn the pros and cons of digital marketing, you won't reap the full benefits.

Reference

1. Gawer, Annabelle, and Michael A. Cusumano. 2014.
2. Industry Platforms and Ecosystem Innovation. Journal of Product Innovation Management. Felix, Reto, Philipp A. Rauschnabel, and Chris Hinsch. 2017.
3. Elements of strategic social media marketing: A holistic framework. Journal of Business Research.
4. Van den Bulte, Christophe, and Yogesh V. Joshi. 2007.
5. New product diffusion with Influentials and imitators. Marketing Science.
6. Hanna, Richard, Andrew Rohm, and Victoria L. Crittenden. 2011.

ARTIFICIAL SIGHT FOR THE VISUALLY IMPAIRED PEOPLE (WALKING STICK).

Ms. L. Sumi¹, Ms. J.Sasirekha²

*Assistant Professor, Department of Computer Applications,
Marudhar Kesari Jain College for Women,TN,India*

ABSTRACT

God gifted sense of vision to the human being is an important aspect of our life. Visually impaired person find themselves challenging to go out independently. Blind people have big problem when they walk on the street or stairs using white cane, but they have sharp haptic sensitivity. In this project, we propose a gifted navigating device for visually impaired or blind people. The Walking stick will help the blind person by providing more convenient means of life. In the proposed system, we develop a walking stick named as Walking stick. A walking stick is equipped with sensor called ultrasonic sensor and it is used to detect the obstacles in front of the user. GPS technology is used for outdoor navigation purpose. GPS and GSM both are used to pin point the location of the visually challenged person. Solar panel is placed into the walking stick for power consumption which can be used when the batteries are drained, even it can be used to charge mobile phones. The main aim of the project is to contribute our knowledge and services to the people of blind and vision to India.

Keywords

Stick, Arduino board, Ultrasonic sensor, GPS Modem, GSM Modem, Buzzer, Solar Panel, Battery.

INTRODUCTION

According to estimates from world health organization (WHO) prevention of blindness there are, 285 million people are visually impaired worldwide and 39 million are blind and 246 million have low vision.[10] In India the WHO has estimated nearly 8 million blind and 54 million are visually impaired in India. We can see several numbers of people around us which are visually impaired, and among them millions of people are blind and there are thousands of people those who are irreversibly blind.[5] For blind people, performing daily activities is a difficult task since vision plays a central role in almost every activity of ours. It is not possible to search an object in an unknown place or surroundings without having the eye sight.

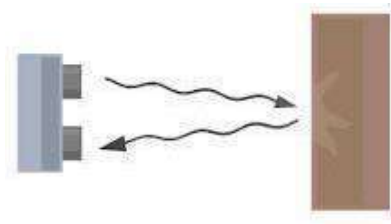
There can be found different technologies such as ultrasonic sensor, GPS and GSM based technologies which was useful for the blind persons. [2] Our walking stick helps the blind people and visually impaired people in moving and allowing them to perform their work easily and comfortably [1].

PROPOSED METHOD

We are developing a gifted tool Walking stick for the blind or visually impaired people to migrate easily and independently without any collision in the busy world. In this design we are interfacing a sensor which is called ultrasonic sensor to detect the obstacles or objects in front of the blind people. [11] This sensor has the eminent feature of finding an obstacle and it is helpful for the partially sighted people to get avoid from the accidents or injuries. Moreover we are integrating a GPS and GSM technology into the stick to find out the location of the blind person. GPS technology are gathering an information from the satellite of latitude and longitude and it send the information to the microcontroller. [4] Microcontroller will collect all the data from the GPS and it gives the gathered data to the GSM. In GSM already integrated a SIM card which stores the information of the location and it sends the latitude and longitude number message to the authorized mobile number. [14] By using this latitude and longitude number the authorized mobile number will receive a SMS to every 15 minutes and they can able to find their exact location of the blind and visually challenged peoples. An additionally we are attaching a solar panel into the stick to observe the sunlight and it saves the electrical energy. This electrical energy will be useful when the battery gets drained and this solar energy is useful for emergency purpose, even we use this energy to charge the mobile by using USB cable.

ULTRASONIC SENSOR

An ultrasonic sensor is a device that can measure the distance to an object by using sound waves. Says it measures distance by sending out a sound wave at a specific frequency and listening for that sound wave to bounce back. By recording the elapsed time between the sound eave being generated and the sound wave bouncing back, it is possible to calculate the distance between the ultrasonic sensor and the object.



Ultrasonic sensor Figure 1.1

ARDUINO UNO

Arduino is a single-board microcontroller meant to make the application more accessible which are interactive objects and its surroundings. [13] The hardware features with an open-source hardware board designed around an 8-bit Atmel AVR microcontroller or a 32-bit Atmel ARM. Current models consist of a USB interface, 6 analog input pins and 14 digital I/O pins that allow the user to attach various extension boards. [7], [15]



Arduino board

Figure 1.2

GPS TECHNOLOGY

A GPS navigation device, GPS receiver, or simply GPS is a device that is capable of receiving information from the GPS satellite and then to calculate the device's geographical position. [8] A GPS tracking unit is a device, normally carried by a moving vehicle or person. That uses the Global Positioning System to track the device movements, at intervals, and to determine its location, and its carrier. Data tracking software is available for smart phones with GPS capability. [12], [9]

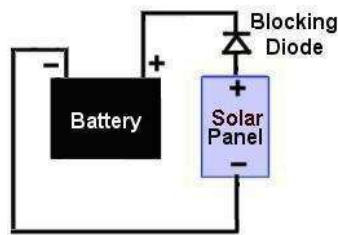
GSM TECHNOLOGY

A GSM modem is a device which can be either a mobile phone or a modem device which can be used to make a computer or any other processor communicate over a network. [3] A GSM modem requires a SIM card to be operated and operates over a network range

subscribed by the network operator. It can be connected to a computer through serial, USB or Bluetooth connection. [6] A GSM modem can also be a standard GSM mobile phone with the appropriate cable and software driver to connect to a serial port or USB port on your computer.

SOLAR PANEL

Solar panels are those devices which are used to absorb the sun's rays and convert them into electricity or heat. A solar panel is actually a collection of solar (or photovoltaic) cells, which can be used to generate electricity through photovoltaic effect. These cells are arranged in a grid-like pattern on the surface of solar panels. Solar panels are used in wide-ranging electronic equipment like calculators, which work as long as sunlight is available.



Circuit diagram

Figure 1.3

EXPERIMENTAL RESULT Working of experimental result



Figure 1.4

The ultrasonic sensor detects an obstacle by following the information of microcontroller. The GPS receiver receives the information from the satellite and it sends the data to the microcontroller. Then the microcontroller will send the data to the GSM. This GSM pretends like a mobile phone; it sends the SMS to the predefined mobile number. And the solar panel will store the energy from the sunlight for the emergency use.

CONCLUSION

The main objective of the paper is to provide a talkative assistance to blind people. We are going to develop an intelligent system that works efficiently good in outdoor. This focuses on designing a device for visually impaired people that help them to travel independently also it must be comfortable to use. The proposed device is used for guiding individuals who are blind or partially sighted. Moreover, it provides the alert to avoid obstacles based on ultrasonic sensors. The GPS and GSM technology is added into the system. A Solar panel can be placed into blind person walking stick. The whole device is designed to be small and it is used for the visionless people and this can be helpful as a artificial sight for the blind and visually impaired people.

REFERENCE

1. Nitish Ojha¹., Pravin Kumar Pradhan²., Prof. M.V.Patil³. "Obstacle Sensing Walking Stick for Visually Impaired", International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue: 04 , Apr -2017
2. Babita Gurjar, "Smart Stick for Visually Impairment Person", SSRG International Journal of Computer Science and Engineering (SSRG-IJCSE) , volume 3 , Issue 8, August 2016
3. Samleo L. Joseph, Jizhong Xiao. June 2015. Being Aware of the World: Toward Using Social Media to Support the Blind with Navigation. IEEE transactions on human-machine systems. 45(3).
4. KherChaitrali S., DabhadeYogita A., Kadam Snehal K., Dhamdhare Swati D., Deshpande Aarti V. "An intelligent walking stick for the blind", International Journal of Engineering Research and General Science Volume 3, Issue 1, Jan-Feb, 2014.
5. Amjed S. Al-Fahoum, Heba B. Al-Hmoud, and Ausaila A. Al-Fraihat "A Smart Infrared Microcontroller-Based Blind
6. Guidance System" :Active and Passive Electronic Components Volume 2013. Harshad Girish Lele, Viten Vilas Lonkar, Varun Vasant Marathe, Mrunmayi Mohan Modak. "electronic path guidance for visually impaired people." The International Journal Of Engineering And Science (IJES), 09-14, 2013.
7. Koley, S. & Mishra, R., 2012. Voice operated outdoor navigation System for visuallyimpaired persons.

8. International Journal of Engineering Trends and Technology Mohd Helmy Abd Wahab, Amirul A. Talib, Herdawatie A. Kadir, Ayob Johari, A.Noraziah, Roslina M. Sidek, Ariffin A. "Smart Cane: Assistive Cane for Visually-impaired People." IJCSI

9. International Journal of Computer Science Issues, pp. 21-27, July 2011.

Dynamic environment exploration using a virtual white cane, in Computer Vision and Pattern Recognition, 2005. CVPR2005. IEEE Computer Society Conference.

ROBOTIC TRANSPOSE INTENT PARTITION FOR EXPLICITLY TRANSPOSE CAMERAS

Ms.Prithi.M¹ , Dr. Tamizharasi. K²

*Dept of Computer Science,
Periyar University,TN, India.*

Abstract

This paper proposes a brand-new shifting item partition set of rules for explicitly transpose cameras which may be very not unusual place for the out of doors Surveillance device, the automobile construct-in surveillance device, and the robot navigation device. A -layer based totally absolutely affine transformation model optimization method is proposed for virtual digital camera compensation purpose, in which the outer layer era is used to clean out out the non- records characteristic elements, and the inner layer era is used to estimate a fashionable affine model based totally absolutely on the RANSAC method. Then the characteristic elements are categorised into foreground and records in keeping with the detected motion facts. A geodesic primarily based totally graph reduce set of rules Is then employed to extract the transferring foreground based totally completely on the labeled functions. Unlike the prevailing worldwide optimization or the long term characteristic issue tracking based totally completely technique, our set of regulations only performs on successive frames to phase the transferring foreground, which makes it suitable for the web video processing packages. The check consequences screen the effectiveness of our set of regulations in every of the immoderate accuracy and the fast speed.

1. Introduction

Moving object detection and segmentation is a number one approach for lots packages collectively with clever video surveillance, clever transportation device, video content material fabric evaluation, video event detection, and video semantic annotation. In those sort of packages, the cameras taking images the films might not be static. For example, digital digicam of an outside surveillance device can be barely shaking due to robust winds, and the video used for content material evaluation or occasion detection can be captured via way of means of a A hand-held virtual digicam. Thus a transferring object detection and segmentation set of policies that can deal with the freely transferring cameras is crucial for the ones cases. However, on one hand most of the winning transferring object detection and segmentation algorithms are most

effective designed for the static cameras, together with Gaussian Mixture Models proposed via way of means of Stauffer and Grimson [1], Kernel density estimation (KDE) applied in [2]. Although many strategies have been proposed to beautify the ones styles of algorithms, collectively with Sun et al. [3] who proposed to hire graph [4] set of rules to enhance the accuracy of the segmentation effects and Patwardhan et al.[5] who built a layer version for the scene To beautify the robustness of foreground detection and segmentation, none of these strategies can be immediately extended for the freely transferring cameras. In modern day years, severa transferring object detection and segmentation algorithms for freely transferring cameras have been proposed. Liu and Gleicher [6] proposed to examine a transferring object model thru manner of approach of collecting the sparse and insufficient motion statistics for the duration of the video. They first stumble on the transferring patches of the foreground object, and then combine the transferring patches of many frames to examine a color model of the foreground object this is used for segmentation. However, this shape of method can only be used to method video sequences offline and cannot be applied for the internet cameras.

To first estimate the dense depth map for each body, and then withinside the segmentation step, a international optimization is applied to multiframe to extract the transferring item. The depth map estimation and item segmentation step may be run iteratively for numerous instances a good way to acquire correct effects. This technique may be very time eating and may most effective be used for offline video sequences. Several algorithms using factor trajectories to section the shifting items are proposed in latest years. The instinct of those types of techniques is that the movement resulting from the digital digicam motion is limited via way of means of a few geometric constraints, even as the movement resulting from the item motion isn't. Thus the shifting item may be detected and segmented via way of means of reading the long time trajectories of the important thing factors. However those techniques typically want to calculate the dense optical glide over long term frames, which can be too time eating to run in actual time. And it additionally updates the factor trajectories and movement/look fashions on-line, in order that this set of rules may be used for the web video segmentation scenario. However, the excessive computational value remains a problem. In this paper, we endorse a singular shifting item detection and segmentation set of rules for the freely shifting cameras.

Compared to the prevailing shifting item segmentation algorithms for freely shifting cameras, our set of rules has the subsequent characteristics.

(1) Unlike maximum of the prevailing algorithms, our set of rules does not no longer rent the worldwide optimization or long time function factor monitoring. It most effectively makes use of successive frames to extract the shifting item, which makes it appropriate for the web video processing task.

(2) A layer generation primarily based totally digital digicam movement repayment technique is proposed, in which the outer layer generation is used to replace the foreground and history function units in step with the cutting-edge parameters of the digital digicam movement repayment fashions, and the internal layer generation employs a RANSAC technique to estimate the parameters of the digital digicam movement repayment version primarily based totally at the cutting-edge history function set. This -layer generation primarily based totally technique makes the digital digicam movement repayment extra sturdy and correct.

(3) A function class and filtering set of rules primarily based totally on GMM colour version is proposed, and the categorized function factors are used because the enter of the geodesic distance primarily based totally graph reduce set of rules, that can go back a totally correct segmentation end result.

2. Algorithm Overview

The set of rules has three steps.

(1) Camera Motion Compensation. Since the digital digicam motion among successive frames may be very small in maximum cases, we will truly anticipate that the history among the previous body and the cutting-edge body most effectively has the interpretation and the rotation motion. Thus an affine transformation version may be hired to simulate the motion of the history. When estimating the affine transformation parameters, the corresponding function factors are first located via way of means of a ahead and backward optical glide set of rules, after which a -generation primarily based totally technique is proposed to estimate the parameters.

(2) Feature Extraction and Classification. The part and the nook functions [15] are extracted after which categorized into the shifting foreground functions (denoted as purple factors) and the history functions (denoted as blue factors) in step with the detected movement areas. The foreground and history function units are then filtered via way of means of GMM colour fashions.

(3) Foreground Extraction with Geodesic Distance Based Graph Cut. After the foreground and history function units are received, the geodesic distance from different **pixels to**

the function factors are calculated, and a geodesic self assurance map is generated. By incorporating the geodesic distance and the geodesic self assurance map with the graph reduce set of rules, correct foreground item may be segmented.

3. Details of Our Algorithm

Camera Motion Compensation.

Motion repayment is an algorithmic method used to expect a body in a video, given the preceding and/or destiny frames via way of means of accounting for movement of the digital digicam and/or items withinside the video. It is hired withinside the encoding of video information for video compression, as an example withinside the era of MPEG-2 files. Motion repayment describes a photo in phrases of the transformation of a reference photo to the cutting-edge photo. The reference photo can be preceding in time or maybe from the destiny. When pix may be as it should be synthesized from formerly transmitted/saved pix, the compression performance may be improved. Motion repayment is one of the key video compression strategies utilized in video coding standards, in conjunction with the discrete cosine rework (DCT). Most video coding standards, together with the H.26x and MPEG formats, generally use movement-compensated DCT hybrid coding,[1][2] called block movement repayment (BMC) or movement-compensated DCT (MC DCT). Functionality Motion repayment exploits the truth that, often, for plenty frames of a movie, the most effective distinction among one body and some other is the end result of both the digital digicam shifting or an item withinside the body shifting. In connection with a video file, this indicates a great deal of the facts that represents one body can be similar to the facts used withinside the subsequent body. Using movement repayment, a video circulation will comprise a few full (reference) frames; then the most effective facts saved for the frames in among will be the facts had to rework the preceding body into the following body. In this model, it is just assumed that the displacement vector $u = (u, v)$ of pixel (x, y) can be written as an affine function of the coordinate (x, y) .

where R is the rotation matrix with parameters $R = [\begin{matrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{matrix}]$. T is translation matrix with $T = [\begin{matrix} \alpha_1 & \alpha_2 \end{matrix}]$. Here the rotation matrix R also contained the scale change α_1 and α_4 ; thus this model can handle the scale changes of the background scene, such as the video captured by a forward or backward moving camera. Since the camera motion and the

foreground motion are distinct, this means that the foreground motion is not appropriate to be modeled by the affine transformation model. Thus in ideal, the pixels used to estimate the affine parameters should only contain the background pixels. This can be achieved by our two-layer iteration based method. The outer layer iteration is used to update the fore- and background feature points according to the motion regions detected by the current affine parameters. The RANSAC procedure is used to estimate the affine parameters primarily based totally at the up to date history functions. The function factors utilized in our paper are the brink and nook factors which may be detected the use of the technique defined in [15]. In order to estimate the affine version parameters, the corresponding function factors of the 2 successive frames have to be detected. We rent the ahead and backward optical glide estimation to gain this goal. For the cutting-edge body $\delta = \frac{1}{4} \delta_i$, we first extract its function factors (denoted as $F_{\delta} = \{f_{\delta,1}, f_{\delta,2}, \dots, f_{\delta,N}\}$, where N is the number of the feature points) and then use the pyramid Lucas Kanade optical flow [16] to tune those functions to the following body $\delta+1$.

Initialization: Algorithm 1: Two-layer iteration based camera motion compensation.

The background feature point set is initialized as $S_{\delta} = S_{\delta,0}$;

Step 1. Inner-layer iteration, employs the RANSAC algorithm to estimate the affine parameters based on the current feature set S_{δ}

Step 2. Moving region detection, finds the moving regions based on the current affine parameters;

Step 3. Update the background feature set according to the detected moving regions;

Step 4. Jump to Step 1 to start a new outer-layer iteration until it converges.

That is, the feature points in S_{δ} are stable.

Feature Extraction and Classification.

After obtaining the final affine parameters, we can obtain the frame difference using (4) and then classify the feature points $F_{\delta+1}$ of the current frame into the foreground and background feature sets $F_{\delta+1}^f$ and $F_{\delta+1}^b$. $F_{\delta+1}^f$ and $F_{\delta+1}^b$ cannot be directly used for graph cut algorithm in the following step to extract the foreground object, because there usually exist some classification errors. As pointed out by the green ellipses some foreground feature points are misclassified into background. This is because the moving regions detected by the frame difference are composed of both the real foreground regions and the false foreground regions. These false foreground regions are actually background regions occluded by the moving object.

In order to eliminate these misclassifications, we further perform a refining process in our algorithm. Since we already have an initial classification of the feature points, we can build two Gaussian mixture models (GMMs) for F_1 and F_2 and then use these two models to reestimate the probability of each feature point belonging to the foreground. The feature points in F_1 or F_2 are first classified into K clusters, respectively, by a farthest point clustering algorithm and then the mean and variance (μ_k, σ_k) for each cluster are calculated to construct GMMs. So the probability of feature points belonging to the foreground can be estimated as

$$p_f(\mathbf{c}) = K \sum_{k=1}^K p_f(\mathbf{c}) \prod_{d=1}^3 \Phi(\mathbf{c}_d, \mu_{kd}, \sigma_{kd}),$$

$$p_b(\mathbf{c}) = K \sum_{k=1}^K p_b(\mathbf{c}) \prod_{d=1}^3 \Phi(\mathbf{c}_d, \mu_{kd}, \sigma_{kd}),$$

$$P_f(\mathbf{c}) = p_f(\mathbf{c}) p_{\mathbf{c}}(\mathbf{c}) + p_b(\mathbf{c}), P_b(\mathbf{c}) = 1 - P_f(\mathbf{c}),$$

where \mathbf{c} is the color vector of one feature point to be estimated, $P_{\mathbf{c}}(\mathbf{c})$ and $P_{\mathbf{c}}(\mathbf{c})$ are the prior probability of the feature points in this Gaussian component and can be calculated as the ratio between the number of feature points in this component and the number of feature points in the whole GMM, and \mathbf{c}_d denotes the color channel and Φ denotes the Gaussian kernel. Then for the feature points in F_1 or F_2 , if $P_f \geq \tau$, this point will be removed from F_1 . Similarly, for the feature points

Foreground Extraction with Geodesic Graph Cut.

Till now, we have obtained the foreground and background key points. This means we have labeled partial pixels as foreground and background. Starting from the initial labeling, we can obtain a complete foreground segmentation by employing a geodesic graph cut algorithm [18], where we use the geodesic distance and color models to calculate the energy function of the graph cut algorithm where we use the geodesic distance and color models to calculate the energy function of the graph cut algorithm, which is defined as where $L = (L_{\mathbf{c}})$ is a binary vector and $L_{\mathbf{c}}$ is the label F or B for pixel \mathbf{c} . $\psi(L_{\mathbf{c}})$ is a unary term and $\lambda(\mathbf{c}, \mathbf{c}') |L_{\mathbf{c}} - L_{\mathbf{c}}'|$ is the pairwise term of

the energy function. λ is a weight to balance the unary and pairwise term. The unary term is defined as follows:

where $\lambda L_{\mathbf{c}}$ is a constraint for the foreground and background feature points:

where ΩL_{\diamond} indicates the foreground and background features and L_{\diamond} denotes the label opposite L_{\diamond} (i.e., if $L_{\diamond} = F$, then $L_{\diamond} = B$). $\diamond\diamond L_{\diamond}$ ($\diamond\diamond$) is computed by normalizing the relative foreground/background geodesic distances.

4. Conclusion

This paper proposed a actual time on-line shifting item detection and segmentation set of rules for the video captured via way of means of freely shifting cameras which most effective use successive frames to section the shifting item. A -layer generation set of rules is proposed to as it should be estimate the affine transformation parameters among successive frames. A function factor detection and filtering set of rules is proposed to dispose of the mistake foreground and history function factors. The item is subsequently extracted via way of means of a geodesic graph reduce set of rules. This set of rules is confirmed to be very green for plenty films. Compared to the prevailing long time key factor trajectory-primarily based totally set of rules, our set of rules now no longer most effective can carry out in on-line processing mode, however can also run in excessive speed. This makes our set of rules very realistic in lots of packages.

References

- [1] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99), pp. 246–252, June 1999.
- [2] A. Elgammal, D. Hanvood, and L. S. Davis, "Nonparametric model for background subtraction," in European Conference on Computer Vision, pp. 751–767, 2000.
- [3] Y. Sun, B. Li, B. Yuan, Z. Miao, and C. Wan, "Better foreground segmentation for static cameras via new energy form and dynamic graph-cut," in Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06), pp. 49–52, August 2006.
- [4] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 9, pp. 1124–1137, 2004.
- [5] K. Patwardhan, G. Sapiro, and V. Morellas, "Robust foreground detection in video using pixel layers," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 4, pp. 746–751, 2008.

[6] F. Liu and M. Gleicher, “Learning color and locality cues for moving object detection and segmentation,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09), pp. 320–327, June 2009.

[7] A. Kundu, K. M. Krishna, and J. Sivaswamy, “Moving object detection by multi-view geometric techniques from a single camera mounted robot,” in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '09), pp. 4306–4312, October 2009.

[8] G. Zhang, J. Jia, W. Xiong, T. Wong, P. Heng, and H. Bao, “Moving object extraction with a hand-held camera,” in Proceedings of the IEEE 11th International Conference on Computer Vision (ICCV '07), pp. 1–8, October 2007.

[9] P. Ochs and T. Brox, “Object segmentation in video: a hierarchical variational approach for turning point trajectories into dense regions,” in Proceedings of the IEEE International Conference on Computer Vision (ICCV '11), pp. 1583–1590, November 2011.

[10] Y. Sheikh, O. Javed, and T. Kanade, “Background subtraction for freely moving cameras,” in Proceedings of the 12th International Conference on Computer Vision (ICCV '09), pp. 1219–1225, October 2009.

[11] P. Sand and S. Teller, “Particle video: long-range motion estimation using point trajectories,” International Journal of Computer Vision, vol. 80, no. 1, pp. 72–91, 2008.

THERAPEUTIC EFFECT OF APPLE CIDER VINEGAR ON TYPE II DIABETES MELLITUS

*Dr.R.Shobanadevi,
Assistant Professor, Department of ND,
Marudhar Kesari Jain College for Women, TN, India*

ABSTRACT

Type 2 diabetes is the most common metabolic disorder worldwide and its prevalence is growing at an alarming rate in both developed and developing countries. It is characterized by abnormalities in carbohydrate, lipid and lipoprotein metabolism, which lead to hyperglycemia and many complications such as hyperlipidemia, hyperinsulinemia, hypertension and atherosclerosis. In order to prevent diabetes, in addition to oral hypoglycaemic drugs, the dietary component such as Apple cider vinegar seems to be promising for glycemic control in patients with type 2 diabetes as well as for diabetes related medical conditions. Apple cider vinegar is fermented juice from crushed apples. Acetic acid in vinegar seems to suppress disaccharidase activity and increase glucose-6-phosphate levels in skeletal muscle. Thus an attempt is made in the present study to find out the impact of apple cider vinegar in patients with type 2 diabetes. Apple cider vinegar was purchased from local market and 15ml was given before meals– twice a day for one month. The study was conducted for one month with 20 individuals with type 2 diabetes and divided randomly into two equal groups as vinegar group (n=10) and control group (n=10). Information pertaining to the socio economic status, anthropometry assessment, biochemical assessment, dietary pattern, health status and personal habits were collected from the selected subjects. There was a significant decrease in BMI, WHR, fasting blood sugar, post prandial blood sugar levels and HbA1c in vinegar group after supplementation of apple cider vinegar (p value <0.05). No such differences were found among the control group during study period. On conclusion, the above results revealed that apple cider vinegar has got an exclusive antidiabetic property and help in preventing diabetic complications. Vinegar is inexpensive, readily available, and a flavor enhancer. Apple cider vinegar was most effective to decrease glucose, total cholesterol, triglycerides, LDL and increases HDL because of its higher concentration of organic acids and phenolic compounds.

INTRODUCTION

Type 2 diabetes is the most common metabolic disorder worldwide and its prevalence is growing at an alarming rate in both developed and developing countries. It is characterized by abnormalities in carbohydrate, lipid and lipoprotein metabolism, which lead to hyperglycemia and many complications such as hyperlipidemia, hyperinsulinemia, hypertension and atherosclerosis. All forms of diabetes increase the risk of long-term complications. These typically develop after many years (10–20), but may be the first symptom in those who have otherwise not received a diagnosis before that time. The major long-term complications relate to damage to bloodvessels. In order to prevent diabetes, in addition to oral hypoglycaemic drugs, the dietary components such as Apple cider vinegar seems to be promising for glycemic control in patients with type 2 diabetes as well as for diabetes related medical conditions. (Khan et al, 2003, Anderson et al 1999 and Soltan SA et al, 2012).

Apple cider vinegar is fermented juice from crushed apples. Like apple juice, it likely contains some pectin; vitamins B1, B2, and B6; biotin; folic acid; niacin; pantothenic acid; and vitamin C. There is interest in using apple cider vinegar for diabetes and cardiovascular diseases. Feeding apple cider vinegar to animals with experimentally induced diabetes significantly reduces hemoglobin A1C (HbA1C), lowers low-density lipoprotein (LDL) cholesterol and triglycerides, and increases high-density lipoprotein (HDL) cholesterol. In another animal model, apple cider vinegar decreased triglycerides and very low-density lipoprotein (VLDL) cholesterol. Preliminary clinical research suggests that vinegar can lower postprandial glucose levels in healthy volunteers. Vinegar is thought to affect glucose levels by delaying the gastric emptying rating. Acetic acid in vinegar also seems to suppress disaccharidase activity and increase glucose-6-phosphate levels in skeletal muscle.

Many medicinal components that are good for health have been reported in natural vinegar, such as carbohydrates, organic acid (acetic, formic, lactic, malic, citric, succinic and tartaric), alcohols and amino acids and peptides (Cocchia et al., 2006), vitamins and mineral salts, polyphenolic compounds (Gallic acid, catechin, caffeic, ferulic acid). Different types of vinegar are available in market. Traditional vinegar is produced from regional foods according to well established customs. The balsamic vinegar of Modena, Italy is made from the local white Trebbiano grapes. Traditional rice wine vinegar is produced in Asia, coconut and cane vinegar is common in India and Phillippines and date vinegars are popular in the Middle East. Some

scientific investigation clearly states the benefits of vinegar such as antimicrobial properties (Vijayakumar and Wolf Hall., 2002.), prevent inflammation and hypertension (Murooka and Yamshita, 2008), lower serum cholesterol (Fushimi et al., 2006), reduction in systolic blood pressure (Kondo et al., 2000), enhanced calcium absorption and retention (Kishi et al., 1999), decrease the glycemic index of carbohydrate food for people with and without diabetes (Sugiyama et al., 2003; Johnston et al., 2004). Antiglycemic effects of vinegar have been known for more than a century and have been demonstrated in animal as well as human studies (Salbe et al., 2009).

According to Soltan SA et al., 2012, Different types of vinegar have protective effect on pancreas and stomach with 15% concentration for 6 weeks. So that using vinegar has a beneficial effect on diabetic patients. Among all types of vinegar, apple vinegar was most effective to decrease glucose, total cholesterol, triglycerides, LDL and increases HDL followed by grape, sugarcane, coconut, artificial and palm vinegar. These results support a therapeutic effect for Apple cider vinegar in individuals at risk for type 2 diabetes. Thus an attempt is made to supplement Apple cider vinegar in patients with type 2 diabetes.

MATERIALS AND METHODS:

The present study was carried out in M.G. Diabetes Hospital, Salem. This hospital was selected because of the easy accessibility of the diabetic subjects and convenience of the investigator. A total of 40 patients with type 2 diabetes were selected to participate in the study. From the 40 non insulin dependent diabetic subjects, 20 NIDDM subjects aged 30-65 years of both sexes with the range of Fasting blood glucose levels between 120-190 mg/dl, HbA1c 6-8%, post prandial blood glucose levels between 140-200mg/dl were selected for the study. Exclusion criteria were alcohol consumption, pregnancy and lactation women, allergy to vinegar, existing with liver or renal or thyroid diseases and haemolytic anaemia. Patients were informed and oriented about the study and written consent form was taken from them.

Apple cider vinegar was purchased from local market and 15ml was given before meals – twice a day for one month. The study was conducted for one month with 20 individuals with type 2 diabetes and divided randomly into two equal groups. The experiment groups are Group I (n=10) : Vinegar group - Consumption of 15 ml of apple cider vinegar before meals twice a day for one month (Before breakfast and dinner). Group II (n=10) : Control group – no intervention.

All medications were continued as usual and subjects were advised to maintain their

normal diets and continue their habitual physical activity throughout the study. On days 0 and 30 approximately 10ml of fasting blood was collected from each subject and analysed for Fasting blood glucose levels and HbA1c. On the same day approximately 5ml of blood sample was collected 2hours after breakfast to analyse post prandial blood sugar level.

A detailed interview schedule was developed by the investigator in order to elicit information pertaining to the socio economic status, dietary pattern, health status and personal habits of the selected subjects. The general information such as age, sex, educational status, work pattern, income level, type and composition of family were collected from the selected subjects. The following techniques were employed to carry out the assessment of nutritional status. Anthropometric parameters like height, weight, Body Mass Index (BMI) and Waist Hip Ratio (WHR) were calculated for all the subjects. Precise information on food consumption pattern was collected through 24 hour recall method and food frequency method. The biochemical parameters such as Fasting blood glucose, post prandial blood glucose levels and Glycosylated haemoglobin (HbA1c) were analysed for the selected subjects. Clinical data regarding family history of diabetes, duration and complications were collected from selected diabetic subjects. Data collected by Interview schedule were consolidated.

RESULTS AND DISCUSSION

Table 1: Baseline Characteristics of the participants

Particulars	Vinegar group (n=10)		Control group (n=10)	
	0 day	30 th day	0 day	30 th day
Gender, F/M	5/5	-	4/6	-
Age in years	42.3±10.0	-	50.1±10.6	-
BMI (kg/m ²)	27.6±2.9	26.6±2.0	28.1±2.4	27.9±2.6
WHR	1.05±0.5	0.99±0.4	1.24±0.23	1.25±0.28

Baseline characteristics of the participants were shown in table 1. Significant decrease in the BMI was observed in vinegar group when compared to control group. No weight reduction was noticed in the control group and they tend to have more or less same weight at the end of the study. Waist Hip Ratio (WHR) is considered as another important factor in assessing nutritional status of an individual. Like BMI, Vinegar group shows the decrease in the WHR at the end of

the trial when compared to that of Control group. Hence there is a significant difference in the BMI and WHR in vinegar group before and after supplementation (p value <0.05). This result reveals that apple cider vinegar helps in weight reduction.

Table 2: Clinical data of the participants

Particulars	Vinegar Group (n=10)	Control Group (n=10)
Family history of Diabetes, Yes/No	4/6	7/3
Duration of diabetes	5.83±2.63	5±2.82
Complications	Only Diabetic – 7 Diabetes + Hypertension – 3	Only Diabetic – 5 Diabetes + Hypertension – 5

Table 2 indicates the clinical data of the participants. Out of 10 participants in vinegar and control group, 4 and 7 participants respectively were found to have family history of diabetes. Duration of diabetes in both vinegar and control group vary from 2 to 10 years. Among 10 participants 7 subjects in vinegar group and 5 subjects in control group were only diabetic. Whereas the remaining 3 subjects in vinegar group and 5 subjects in control group were both diabetic and hypertensive.

Table 3: Biochemical parameters of the participants before and after supplementation

Biochemical parameters	Vinegar group (n=10)		Control group (n=10)	
	0 day	30th day	0 day	30th day
Fasting Blood glucose (mg/ml)	169.6±18.5	155.8±22.5	184.6±10.5	173.8±15.4
Post Prandial Blood Glucose (mg/ml)	190.5±9.3	165.5±12.6	188.3±7.31	203.7±34.8
HbA1c (%)	8±0.2	6.08±0.22	7.98±0.19	8.86±0.62

Table 3 shows the results of biochemical parameters such as fasting blood glucose levels, postprandial blood glucose levels and HbA1c levels of the participants during the study. Applecider vinegar is thought to affect R Durgadevi and P Nazni (2012). This report supports Thecurrent study findings. In the present study, there was a significant decrease in the fasting blood sugar, post prandial blood sugarand HbA1c levels in vinegar group aftersupplementation of apple cidar vinegar (p value <0.05). Whereas in control group no significant changes was observed. Several mechanisms to account for these effects have beenproposed, including interference with enzymaticdigestion of complex carbohydrates, delayedgastric emptying and enhanced peripheral glucose uptake and conversion to glycogen. Theseresults strongly prove

the antidiabetic activity of the apple cidervinegar.

CONCLUSION

These data reveals that daily consumption of apple cider vinegar favorably influences fasting glucose concentrations, post prandial blood sugar levels and HbA1c levels in diabetic patients. Vinegar is inexpensive, readily available, and a flavor enhancer. In conclusion, Apple cider vinegar has got an exclusive anti- diabetic property and help in preventing diabetic complications.

REFERENCES

- Abduo H: Cinnamon use for diabetes management. Drug info. Volume 2, issue 10, Jan 2014.
- Anderson RA, Jarvill-Taylor KJ, Graves DJ. Role of Psyllium husk in treating diabetes mellitus. J Am Coll Nut. 2001;20:327-36.
- Anderson W, Allgood D, Turner J, Oeltgen R, Daggy P: Effects of psyllium on glucose and serum lipid responses in men with type 2 diabetes and hypercholesterolemia. Amjcn. 1999;70: 466-473.
- Cocchia M, Durantea C, Grandia M, Lambertinic P, Manzinib D, and Marchetti A: Simultaneous determination of sugars and organic acids in aged vinegar and chemometric data analysis. Atlanta. 2006: 69: 1166- 1175.
- Frati-Munari AC, Fernandez-Harp JA, Becerril M, Chavez-Negrete M, Bañales-Ham M. (1983). Decrease in serum lipids, glycemia and body weight by *Plantago psyllium* in obese and diabetic patients. *Arch. Invest. Med. (Mex.)*, 14: 259-268.MEDLINE
- Fushimi T, Suruga K, Oshima Y, Fukiharu M, Tsukamoto Y, and Goda T: Dietary acetic reduced serum cholesterol and triacylglycerols in rats feed a cholesterol rich diet. British Journal Of Nutrition, 2006: 95(5):916- 924.
- Fushimi T, Tayama K, Fukaya M, Kitakoshi K, Nakai N, Tsukamoto Y, Sato Y: Acetic acid feeding enhances glycogen repletion in liver and skeletal muscle of rats. Journal of Nutrition. 2001: 131:1973-1979.
- Khan A, Safdar M, Ali khan MM, Anderson RA: Cinnamon improves glucose and lipids of people with type 2 diabetes. Diabetes Care 26:3215-3218,2003.
- Kishi M, Fukaya M, Tsukamoto Y, Nagasaw T, Takenhana K, Nishizawa N: Enhancing effect of dietary vinegar on the intestinal absorption of calcium in ovariectomized rats. Bioscience, Biotechnology and Biochemistry. 1999: 63:905-910.
- Konda, S., Tayama K, Tsukamoto Y, Ikeda K, Yamori Y: Antihypertensive effects of acetic acid and vinegar on spontaneously hypertensive rats. Bioscience, Biotechnology and Biochemistry.2000:

65:2690-2694.

- Mang B, Wolters M, Schmitt B, Kelb K, Litchtinghagen R, Stichtenoth DO, et al: Effects of a cinnamon extract on plasma glucose, HbA1c, and serum lipids in diabetes mellitus type 2. *Euro J Clin Investig.* 2006;36:340-4.
- Murooka Y, Yamshita M: Traditional healthful fermented products of Japan. *Journal Industrial Microbiology Biotechnology.* 2008: 35:791-798.
- Nishidai S, Nakamura Y, Torikai K, Yamamoto M, Ishihhara N, and Mori H: Kurusu, a traditional vinegar produced from unpolished rice, suppresses lipid peroxidation vitro and in mouse skin. *Bioscience, Biotechnology and Biochemistry.* 2000;64:1909-1914.
- Salbe AD, Johnston CS, Buyukbese AM, Tsitouros PD, and Harman SM: Vinegar lacks antiglycemic action on enterol carbohydrate absorption in human subject. *Nutrition Research.* 2009: 29: 846-849.
- Shimoji Y, Tamura Y, Nakamura Y, Nanda K, Nishidai S and Nishikawa Y: Isolation and identification of DPPH radical scavenging compounds in Kurusu (Japanese unpolished rice vinegar). *Journal of Agricultural and Food Chemistry.* 2002: 50:6501-6503.
- Soltan SA, Shehata M: Antidiabetic and Hypocholesterolemic effect of different types of vinegar in rats. *Life Science Journal* 2012;9(4).
- Sugiyama M, Tang AC, Wakaki Y and Koyama W: Glycemic index of single and mixed meal foods among common Japanese foods with white rice as a reference food. *European Journal Clinical Nutrition.* 2003: 57: 743- 752.
- Vijayakumar C, and Wolf Hall CE: Evaluation of household sanitizers for reducing levels of *Escherichia coli* on iceberg lettuce. *Journal of Food protection.* 2002: 65:1646-1650.
- R Durgadevi and P Nazni, Comparative study of processed Amaranth grains on glycemic indices in NIDDM subjects, (2012) *International Journal of Pharma Medicine and Biological Sciences*, Vol. 1, No. 2, October 2012, pp: 194-205, ISSN 2278 –5221.

COVID-19 DETECTION USING MACHINE LEARNING

*Ms.Asma Farhana. M.N¹ , Nehlath Harmain. H.E² ·Nandhini. V³
Dept of Computer Application,
MKJC,TN,India*

ABSTRACT:

The Global Pandemic situation caused due to the novel Corona Virus impacted the entire World. The detection of the Corona Virus is usually done using the nasal RT-PCR Swab test. However, the Chest CT scans are also more efficient in the diagnosis of Covid-19 as they help in early detection of the virus and its level of intensity. Despite the efficiency of these methods, there comes an unfortunate situation of the delay in the reports using these methods. To overcome this drawback, this Project has conducted a research on the early detection of the Virus to address the delay. The Project involves the use of image classification of chest x-rays and CT scans obtained from patients using Machine learning Algorithms. The Machine is trained using a set of data to identify and predict the early signs of the Virus. The main goal of this project is to develop an improved method of machine learning models that help in the accurate and precise detection of the presence of the Corona Virus in patients. This method also assists the radiologists in early screening and obtaining the results more quickly.

INTRODUCTION:

The major outbreak of the Corona Virus began in the Wuhan City of China. It grew so drastically that it has impacted the entire world to a huge extent. As the number of positive cases increase, there is a huge threat to the world population. Over the past 2 years, there is a substantial amount of increase in the Recovery rates of patients. However, the death rates have also grown in accordance with the reported number of positive cases. The SARS-COV2 or more popularly known as the Covid-19 majorly affects the respiratory tract of the patients infected. The major modes of transmission of the Corona virus is through the infected droplets of patients which is spread while coughing, Sneezing or touching the infected surfaces, etc. The WHO has specified certain guidelines as precautionary measures to prevent the transmission of the deadly virus. No proper treatment methods are available to treat the virus effectively. The treatment is usually done for the existing Symptoms which include Headache, fever, shortness of breath, Coughing, runny nose, etc. After coming into contact, the virus can spread and grow within 2-14 days based on the incubation period. While most of the patients recover quickly, this

condition has turned to be fatal for patients with underlying medical conditions like Heart disease, diabetes, asthmatic patients, etc.

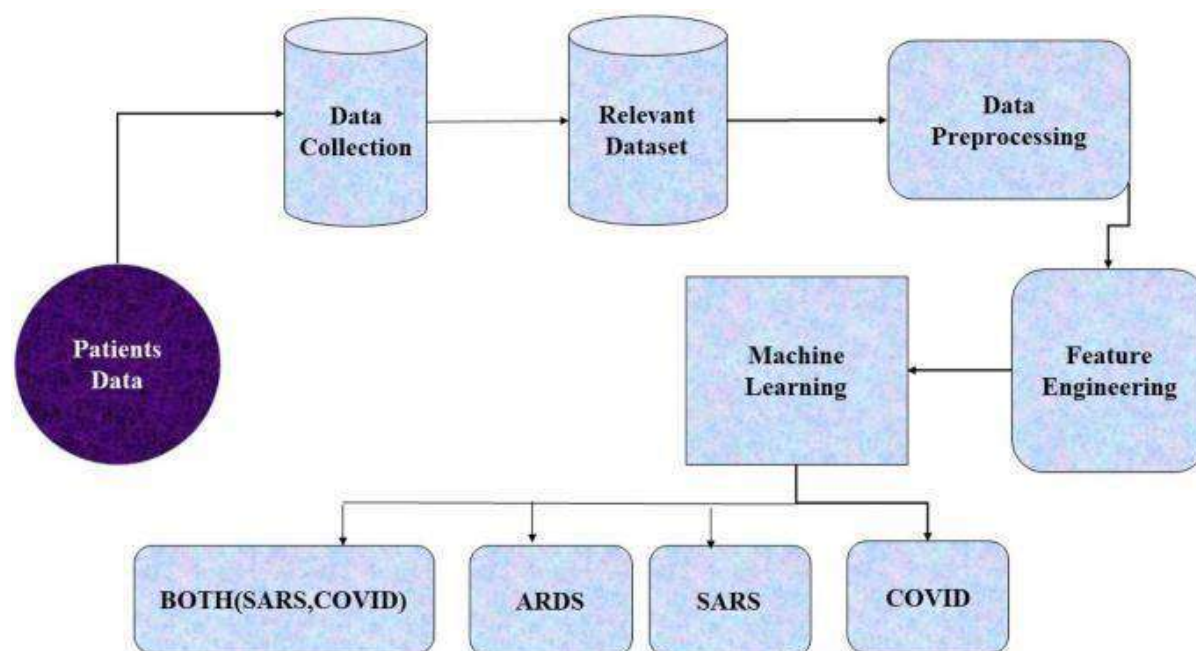
Vaccination drugs have been developed and administered to people. This has helped to curb the spread of the virus to some extent. Early detection of the virus is extremely necessary so that the patients can be treated at the earliest at the initial symptoms. The regular methods available for the diagnosis include the nasal PCR swab test and the CT scans. However, these methods take a lot of time to produce the results. The samples are collected first and sent to lab and then the results are obtained. This causes a delay in obtaining reports. This Project overcomes this drawback by providing an effective and accurate solution through the use of Machine Learning Algorithms. This method can be beneficial to diagnose and predict COVID-19. Early diagnosis of COVID-19 can save millions of lives, thereby preventing the delay in the treatment. These machine learning algorithms are trained with the set of data to produce accurate predictive results. The efficiency of this method is around 96.5% .

OBJECTIVE:

The main goal of this project is to provide quick and effective detection of Covid-19 through Machine Learning Algorithms based on the Clinical data and X-ray and CT images.

RESEARCH METHODOLOGY AND TOOLS:

The proposed research methodology consists of the following steps: collection of data is the first step and the second step involves refining the data, the third step provides an overview of preprocessing, and the next step provides a mechanism for feature extraction. The diagrammatic representation of the proposed methodology is shown in Figure below.



Data collection

The researchers and hospitals give open access to the data regarding this pandemic. We have collected from an open-source data repository GitHub. In which about 2000 patients data is stored that includes symptoms of corona virus and other viruses. Data consists of details such as patient id, offset, sex, age, finding, survival, needed_supplemental_O₂, temperature, pO₂ saturation, leukocyte_count, neutrophil count, lymphocyte count, modality, date, location, folder, filename, DOI, URL. License. Clinical notes and other notes.

Relevant dataset

Since our work is regarding text mining so we extracted clinical notes and findings. Clinical notes consist of text while as the attribute finding consist label of the corresponding text. About 2000 reports were used and their length was calculated.

Preprocessing

The data is unstructured so it needed to be refined such that machine learning can be done. Various steps are being followed in this phase; the data is obtained by removing unnecessary text. Punctuation and lemmatisation are being done such that the data is refined in a better way.

Clinical_notes	Finding	Report_Length	Punctuation	Lemmatisation	Stop_Word Removal
infiltrate in the upper lobe	COVID	45	infiltrate in the upper	infiltrate in the upper lobe	infiltrate upper lobe leave lung
progressive infiltrate and	COVID	40	progressive infiltrate	progressive infiltrate and	progressive infiltrate consolidation
progressive infiltrate and	COVID	40	progressive infiltrate	progressive infiltrate and	progressive infiltrate consolidation
progressive infiltrate and	COVID	40	progressive infiltrate	progressive infiltrate and	progressive infiltrate consolidation
diffuse infiltrates in the bil	COVID	48	diffuse infiltrates in th	diffuse infiltrate in the bila	diffuse infiltrate bilateral lower lungs
progressive diffuse inters	COVID	115	progressive diffuse int	progressive diffuse interst	progressive diffuse interstitial opacities consolidation
Severe ARDS. Person is in	ARDS	53	severe ards person is	severe ards person be intu	severe ards person intubate og place
Case 2: chest x-ray obtain	COVID	563	case 2 chest x-ray obt	case 2 chest x-ray obtain o	case 2 chest x-ray obtain jan 6 (2a) brightness lungs
Case 2: chest x-ray obtain	COVID	563	case 2 chest x-ray obt	case 2 chest x-ray obtain o	case 2 chest x-ray obtain jan 6 (2a) brightness lungs
SARS in a 74-year-old mar	SARS	71	sars in a 74-year-old r	sars in a 74-year-old man	sars 74-year-old man develop symptoms 4 days exp
SARS in a 74-year-old mar	SARS	71	sars in a 74-year-old r	sars in a 74-year-old man	sars 74-year-old man develop symptoms 4 days exp
SARS in a 74-year-old mar	SARS	71	sars in a 74-year-old r	sars in a 74-year-old man	sars 74-year-old man develop symptoms 4 days exp
SARS in a 29-year-old wor	SARS	378	sars in a 29-year-old v	sars in a 29-year-old wom	sars 29-year-old woman present 7 days exposure ()
SARS in a 29-year-old wor	SARS	378	sars in a 29-year-old v	sars in a 29-year-old wom	sars 29-year-old woman present 7 days exposure ()
SARS in a 42-year-old wor	SARS	145	sars in a 42-year-old v	sars in a 42-year-old wom	sars 42-year-old woman present 9 days exposure pr

Feature engineering

From the preprocessed clinical reports, various features are extracted as per the semantics and are converted into probabilistic values. We use TF//IDF technique for extracting relevant features. By giving the corresponding weight to the feature and the same input is being supplied to machine learning algorithms.

lungs	chest	patient	multiple	peripheral	bilateral	lower	lung	leave	image	lob	opacities	ct	right	lobe	air	pneum	glass	opacities	history
0.379	0	0	0	0	0.34539	0.373	0.379	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0.612	0	0	0	0	0	0	0	0	0.45	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.31916	0.241498	0.13287	0	0	0.24	0	0	0.119823	0	0	0	0	0.16	0.256674173	0.223	0
0	0	0	0	0	0.45603	0	0	0	0.573	0	0	0	0	0	0	0	0	0	0
0.342	0	0	0	0	0	0.336	0.342	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.50247	0.298721	0	0	0	0	0	0	0	0	0	0	0	0	0.393524911	0	0
0	0	0	0.26141	0	0	0	0	0	0	0	0	0	0.3	0.357	0.3	0.3	0	0	0
0	0	0	0	0	0	0	0	0	0	0.41	0	0	0	0	0	0	0	0	0
0	0	0.2225	0	0.340237	0	0	0	0.2	0	0.17	0	0	0	0	0	0	0	0	0.314
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.481
0	0	0	0.28544	0	0	0	0	0	0	0	0	0	0.3	0.39	0.4	0.4	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.317
0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.282	0.3	0.3	0	0	0.238

ML is a subset of AI that consists in the algorithmic modeling culture of statistical models , and only needs a small amount of knowledge to learn how to solve problems.

Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), K-nearest Neighbor (KNN), Adaboost, K-means clustering (KC), Density clustering (DC), Hidden Markov Models (HMM), Support vector machine (SVM), Naive Bayes (NB), Restricted Boltzmann Machines (RBM), and Artificial Neural Network (ANN), such as Recurrent Neural Networks (RNN), including Long-short-term-memory (LSTM), Autoencoder (AE), and Generative Adversarial Network (GAN), are ML techniques.

Traditional machine learning algorithms

Logistic regression

This algorithm predicts the class of numerical variable based on its relationship with the label [15]. The 40 features that have been selected in feature engineering with values are represented in the form of a table and are supplied as an input, as shown in Fig. 6. The algorithm generally calculates the class membership probability. Here we have four classes. $y \in \{0, \dots, 3\}$. The posterior probabilities can be calculated with the help of Eq. 1.

$$P(y=k|x) = \frac{\exp(\theta_k^T x)}{\sum_{k=0}^3 \exp(\theta_k^T x)} \quad \forall k=0,1,2,3$$

Multinomial Naïve Bayes

MNB computes class probabilities of a given text by using Bayes rule. Let C denote the set of classes in our problem we have four classes $C = 0, 1, 2$ and 3 . Moreover, N is the set of features here we have $N = 40$ (40 features are taken using TF/IDF) as shown in Fig. 6. Then MNB assigns test text t_i to the class that has the highest probability $P(c|t_i)$ by using Bayes rule shown in Eq. 2:

$$P(c|t_i) = \frac{P(c)P(t_i|c)}{P(t_i)} \quad c \in C$$

$P(c)$ can be calculated by dividing the number of clinical textual data that is labelled as class c to the total number of clinical textual data. $P(t_i|c)$ is the probability of obtaining a clinical text report like t_i in-class c and is calculated as:

$$P(t_i|c) = \frac{1}{\sum_{n=1}^N f_{ni}} \prod_{n=1}^N P(w_n|c)^{f_{ni}}$$

where f_{ni} is the count of word/term 'n' in our clinical text report t_i and $P(w_n|c)$ is the probability of word/term 'n' given in class c . From the training data, the latter probability is calculated by:

$$P(w_n|c) = \frac{1 + F_{nc}}{N + \sum_{x=1}^N F_{xc}}$$

where F_{xc} is the count of word/term 'x' in all the clinical training reports belonging to the class c . for avoiding zero-frequency problem Laplace estimator is used which assigns value one to each word's count.

Decision trees

An alternative approach for classification it partitions the input space into regions and classifies every region independently [18]. The 40 features that have been selected in feature engineering with values are represented in the form of a table and are supplied as an input, as shown in Fig. 6. It splits the space recursively according to the inputs and classifies at the bottom of the tree. The leaf nodes classify the text into four classes. While building a decision tree, a vital function needs to be considered which is known as the splitting criterion. The function defines how data should be split in order to maximise performance. We have used information gain ratio which is equal to information gain to the intrinsic information and is shown in Eq. 6.

$$IGR(EX,a)=IG/IV \quad IGR(EX,a)=IG/IV$$

where IG = information gain. IV = intrinsic information.

Information gain is calculated with the help of entropy as shown below:

$$IG(EX,a)=H(EX)-\sum_{v \in \text{values}(a)}(|\{x \in EX | \text{value}(x,a)=v\}| |EX|.H(\{x \in EX | \text{value}(x,a)=v\})) \\ H(EX)-\sum_{v \in \text{values}(a)}(|\{x \in EX | \text{value}(x,a)=v\}| |EX|.H(\{x \in EX | \text{value}(x,a)=v\})))$$

where EX = set of training examples and $x \in EX$ which defines the value of a specific example x for feature a . H = entropy and a = features.

Intrinsic information value is calculated by:

$$IV(EX,a)=-\sum_{v \in \text{values}(a)}(|\{x \in EX | \text{value}(x,a)=v\}| |EX|. \log_2((|\{x \in EX | \text{value}(x,a)=v\}| |EX|)) \\ -\sum_{v \in \text{values}(a)}(|\{x \in EX | \text{value}(x,a)=v\}| |EX|. \log_2((|\{x \in EX | \text{value}(x,a)=v\}| |EX|))$$

Ensemble machine learning techniques

Bagging

An ensemble machine learning algorithm which improves the performance of other classification and regression machine learning algorithms. Bagging algorithm helps in avoiding overfitting. Let a training set X of size 'n' is given, by sampling uniformly 'm' new training sets X_i are generated each having size 'n' with replacements. The 40 features that have been selected in feature engineering with values are represented in the form of a table and are supplied as an input, as shown. Due to replacements, some observations could repeat in each X_i . If $m=n$ then

set X_i for large n is expected to have a fraction $(1 - 1/e)$ of the unique examples of X , the rest being duplicates. This sample is known as a bootstrap sample. The m models are fitted using m bootstrap samples and combined by voting.

Stochastic gradient boosting

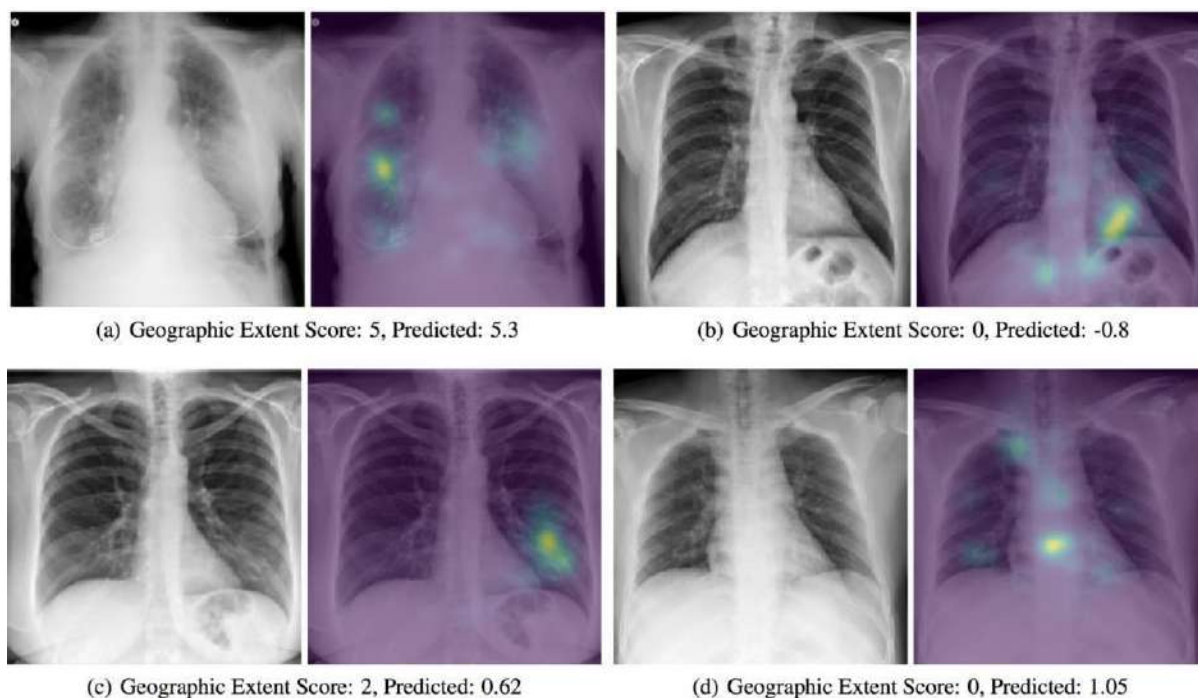
This algorithm allows trees to be greedily created from samples of the training dataset. The 40 features that have been selected in feature engineering with values are represented in the form of a table and are supplied as an input. This is used for reducing the correlation between the trees in gradient boosting. At each iteration, a subsample of the training data is drawn at random without replacements from the full training dataset. The randomly selected subsample is then used instead of the full sample to fit the base learner.

RESULTS:

The statistical computation was done to obtain deeper perspectives about the data. The data is being split into 75:25 ratio where 75% data is used for training the model and 25% is used for testing the model. The clinical text reports are obtained from 2000 patients. The classification was done using machine learning algorithms by supplying them features that were extracted in the feature engineering step. In order to explore the generalization of our model from training data to unseen data and reduce the possibility of overfitting, we split our initial dataset into separate training and test subsets. The tenfold cross-validation strategy was conducted for all algorithms, and we repeat the process several times to avoid the sampling bias introduced by randomly partitioning the dataset in the cross-validation.

The results showed that logistic regression and Multinomial Naïve Bayes Algorithm shows better result than all other algorithms by having precision 94%, recall 96%, F1 score 95% and accuracy 96.2% other algorithms like random forest, gradient boosting also showed good results by having accuracy 94.3% respectively. The model was experimented in two stages. In the first stage, we tested 75% of the available data and it shows less accuracy as compared to the stage in which whole data was used for experimentation. So we can conclude that if more data is supplied to these algorithms, chances of improvement in performance is increased. As it has

become tedious to tackle the deadly virus, our work will somehow help the community by analysing the clinical reports and take necessary actions.



A presentation of a predicted severity scores for COVID-19 chest X-ray scans while using the DenseNet model

Challenges and Future outcomes:

AI-based ML and DL applications in COVID-19 research are currently facing several obstacles, such as legislation, scarcity, and unavailability of large-scale training data, vast noisy data and rumors, limited awareness of the intersection between computer science and medicine, data privacy and security issue, unreliable usability of text data, and more.

CONCLUSION:

With the advancement in technology, Machine Learning techniques can be effectively used to obtain the required efficiency for the early detection of the covid-19. The test results depicted that logistic regression and multinomial Naïve Bayesian classifier gives excellent results by having 94% precision, 96% recall, 95% f1 score and accuracy 96.2%. To curb the transmission of this deadly virus, the advanced technologies like Machine Learning and AI (Artificial Intelligence) methods can be deployed to obtain solutions for the existing problem. More research on this can also be done in the future.

REFERENCES:

- [1] Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, Hu Y, Tao ZW, Tian JH, Pei YY, Yuan ML, Zhang YL, Dai FH, Liu Y, Wang QM, Zheng JJ, Xu L, Holmes EC, Zhang YZ (2020) A new coronavirus associated with human respiratory disease in china. *Nature* 44(59):265–269.
- [2] Chakraborti S, Choudhary A, Singh A et al (2018) A machine learning based method to detect epilepsy. *Int J Inf Technol* 10:257–263. <https://doi.org/10.1007/s41870-018-0088-1>
- [3] Medscape Medical News, The WHO declares public health emergency for novel coronavirus (2020) <https://www.medscape.com/viewarticle/924596>
- [4] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L (2020) Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 395(10223):507–513
- [5] World health organization: Wikipedia coronavirus Pandemic data: https://en.m.wikipedia.org/wiki/Template:2019%E2%80%932020_coronavirus_pandemic_data. Accessed 10 Apr 2020
- [6] Khanday, A.M.U.D., Amin, A., Manzoor, I., & Bashir, R., “Face Recognition Techniques: A Critical Review” 2018
- [7] Kumar A, Dabas V, Hooda P (2018) Text classification algorithms for mining unstructured data: a SWOT analysis. *Int J Inf Technol*. <https://doi.org/10.1007/s41870-017-0072-1>

Cinema and Literature: Two Art Forms That Build And Grow Together

*Dr Hemamalini N,
Assistant Professor, Dept of English,
MKJC, TN, India*

Abstract: The Paper throws literature and cinema with analysis of novel in a movie with critical comments.

Key words : literary novels and cinema

Literary adaptation

The adapting of a literary source (e.g. a novel, short story, poem) to another genre or medium, such as a film, stage play, or video game.

It can also involve adapting the same literary work in the same genre or medium just for different purposes, e.g. to work with a smaller cast, in a smaller venue (or on the road), or for a different demographic group (such as adapting a story for children). Sometimes the editing of these works without the approval of the author can lead to a court case.

It also appeals because it obviously works as a story; it has interesting characters, who say and do interesting things. This is particularly important when adapting to a dramatic work, e.g. film, stage play, teleplay, as dramatic writing is some of the most difficult. To get an original story to function well on all the necessary dimensions—concept, character, story, dialogue, and action—is an extremely rare event performed by a rare talent.

Perhaps most importantly, especially for producers of the screen and stage, an adapted work is more bankable; it represents considerably less risk to investors, and poses the possibilities of huge financial gains. This is because:

- It has already attracted a following.
- It clearly works as a literary piece in appealing to a broad group of people who care.

- Its title, author, characters, etc. may be a franchise in and of themselves already.

Works of literature have been adapted for film from the dawn of the industry. Some of the earliest examples come from the work of Georges Méliès, who pioneered many film techniques. In 1899, he released two adaptations—Cinderella based on the Brothers Grimm story of the same name and King John, the first known film to be based on the works of Shakespeare. The 1900 film Sherlock Holmes Baffled, directed by Arthur Marvin featured Arthur Conan Doyle's detective character Sherlock Holmes intruding upon a pseudo-supernatural burglary. The film, considered the first detective movie, ran for only 30 seconds and was originally intended to be shown in hand-cranked Mutoscope machines.

Méliès' 1902 original science-fiction feature A Trip to the Moon was based loosely on two popular novels of the time: Jules Verne's From the Earth to the Moon (1865) and H. G. Wells' The First Men in the Moon (1901).^[1] The first of many adaptations of the Brothers Grimm tale Snow White was released in 1902 while the earliest surviving copy is the 1916 version. 1903 saw the release of Alice in Wonderland directed by Cecil Hepworth and Percy Stow, the first movie adaptation of Lewis Carroll's 1865 children's book Alice's Adventures in Wonderland.^[2]

The first feature-length film to be shot entirely in Hollywood was Cecil B. DeMille's first assignment, The Squaw Man, in 1914, which was the first of three movie versions (all directed by DeMille) based on Edwin Milton Royle's 1905 play of the same name. Since the early days of the genre, major films have been largely adapted:

- Novels: Gone With the Wind (1939), From Here to Eternity (1953), and The Godfather (1972) were all adapted from novels of the same name.
- Plays: Casablanca (1942), Streetcar Named Desire (1951), and Equus (1977) were all adapted from stage plays.

- Short stories: The Secret Life of Walter Mitty (1947), Breakfast at Tiffany's (1961), The Heart Is a Lonely Hunter (1968), Shawshank Redemption (1994), and Brokeback Mountain (2005) were all made from short stories.

The most celebrated of the early adaptations is Erich von Stroheim's Greed, a 1924 adaptation of the 1899 novel McTeague by naturalist writer Frank Norris. The director intended to film every aspect of the novel in great detail, resulting in a 9½-hour epic feature. At studio insistence, the film was cut down to two hours and was considered a flop upon its theatrical release. It has since been restored to just over four hours^[3] and is considered one of the greatest films ever made.

One book that has been adapted very frequently (in one form or another) is Charles Dickens' 1843 Christmas story A Christmas Carol, which has around 20 film adaptations to date

Difference between literature and novels

Many films are based on novels. However, each of them uses different ways to tell the similar stories. The paper discusses the differences from several aspects in terms of their nature, ways of narration, as well as the effect they bring about. It mainly focuses on the special benefits of using films in English language teaching. Finally, a sample film-based activity which could be used to enhance language classroom is illustrated.

The use of film in teaching literature has become a common instrument in the classroom, perhaps primarily because it helps to understand the materials of the course. Each film uses a variety of ways of conveying the similar story of a literary text that can invigorate the process of education by bringing literature into real life (Bo, 2008). Therefore, using films often provides a clear understanding of difficult concepts, or non-clear themes of the texts may have the capacity to visually communicate knowledge. ...

... As mentioned by Johnson cited in Al-Shalabi (2011), a film can be a way to cover a coffee break, but it should be utilised as an active way of learning. The teacher and students might get involved in it by asking questions to be fully aware of every detail (Bo, 2008). This article discusses the benefits of using films and visual adaptations as a methodology in teaching literary works to English Foreign Language (EFL) learners at Kurdistan universities. use of vocabulary. As well as describing the visual picture and verbal sign by linking film to icons and novel to symbols, Bo (2008) relates to 'the arbitrary connection between two things 'as the term flower can be a flower symbol, while the flower picture is the flower icon showing a 'less arbitrary connection between the two.' He also mentioned that the image of the flower might well be closer than the flower as a word.

Cinema and Literature are two distinct but equally extraordinary works of art. Where literature was a popular form of expression during the 18th and 19th century, cinema has taken its place by the 20th century onwards. Though both these arts have certain connections and differences, both have a similarity of taking its readers/audience to a different world.

Literature has been a way of artistic expression for centuries now. Writers have told tales about gods and goddesses, heroes and their valiant victories, historical epics, romantic tragedies, comic incidents, legendary episodes, and much more. Cinema is by far doing the same thing for quite a few years now. One major strong point in cinema, which is absent in literature, is the advantage of visually showing the whole picture on the screen that helps the audience connect with the moment more closely.

Literature takes its readers on a journey of imagination that is away from the real world while cinema shows such an imaginative world before the audience and they do not have to put much pressure on their minds to delve into their imaginations. They basically view the film through the imagination of the filmmakers.

To put it, in other words, we can say that literature is an art which is developed through writing while cinema brings to life those writings to life through sound, music, visuals, and actors. Literature has all the meanings hidden in itself that are used to develop a film. Though both are somehow interdependent, both need to be studied in order to completely understand a movie based on a piece of literature.

Moreover, literature has always been a great inspiration for cinema all over the world. In India, especially, epics like Mahabharata and Ramayana have been created and recreated on the silver screen several times. Novels of renowned Bengali writers, Gujarati writers, Urdu writers, and English writers are made into films every now and then.

Although it can easily be said that the first step of cinema is literature. Because once a film is in the process of making, it is the script, dialogues, and screenplay that are produced in order to develop it. The production and technical aspects are secondary in the process of filmmaking. Hence, it will not be wrong to say that literature initiated people to move on to cinema.

There is an extraordinary contribution of one art to the other. As history points out that it was all a world of tales and stories from where Aladdin, Ali Baba and the Forty Thieves, Hatim Tai, Cinderella, Snow White, and The Prince and the Pauper originated. And from here cinema took its inspiration and developed films based on these stories or parts of these stories.

There have been several filmmakers who have adapted novels, plays, even poetry into films like J.K. Rowling's Harry Potter Series, Jane Austen's Pride and Prejudice and Sense and Sensibility, Sarat Chandra Chatterjee's Devdas, Homer's Iliad and Odyssey (On which films like Troy and Oh Brother, Where Art Thou? are based).

A Bengali filmmaker, Chidananda Dasgupta, explains about the adaptations of films from literature that certain characters and incidents from the literature may undergo changes, “but the very composition of the elements, the molecular structure if you like, would undergo a transmutation.”

The purpose of the film should not be a mere copy of the literature, rather it must have its own characteristics and techniques that are motivational enough for the audience to enjoy. Though in its literary form many can say that it is in the form of a screenplay of the film but it may not be right to look at a literary piece in such a way. It has an impression of reality even in the written form and once it comes on the silver screen it enhances its qualities to a greater extent than it does in the written form.

Language is another component that differs from a book and a film. Though there may be a similarity in the use of language in both the platforms, there are certain distinct disparities between the usage of language in literature and that in cinema. The relationship between time and space is also quite different from literature to cinema. While in literature an event is described as it has happened, films show as it is happening.

A film must not play the role of the literal visual representation of the book on which it is based. It must be a proper production that has been transformed from the words on paper to the dialogues on the celluloid.

The way both the medium expresses their meaning is where the similarity and the difference lie. Words are the only way to express but while a book has written words on it a film has audio speech which is somehow more powerful and life-like. In a film, a single scene is like a complete sentence or a series of sentences in a book. The power of audio and visual experience has a long lasting effect on the audience over the power of written words.

References:

- Allemand, L. (2015). INDIA'S PARALLEL CINEMA How popular is the Indian alternative movement and how is urban poverty represented?
- Anand, B.K. (2018). Looking at Student-Teacher Bonding through the Lens of Attachment Theories: A Study of Movies Black and Taare Zameen Par. Literary Voice A Peer Reviewed Journal of English Studies. 1(9). 48-55
- Atwal J., (2018). Embodiment of Untouchability: Cinematic Representations of the “Low” Caste Women in India. Open Cultural Studies, 2, 735-745
- Bhugra, D. (2006). Mad Tales from Bollywood. 18
- Butalia, U. (1984). Women in Indian Cinema. Feminist Review, (17), 108-110.
- Chakravarty, S. (1993) National Identity in Indian popular cinema, 1947-1987, University of Texas
- Dasgupta, R.K. (2012). Digital media and the Internet for HIV prevention, capacity building and advocacy among gay, other men who have sex with men (MSM), and transgender (TG): Perspectives from Kolkata, India. Digital Culture and Education, (4) Special Issue: The HIVE.
- Dastidar, S. G. and Elliott, C. (2019). The Indian film industry in a changing international market. Journal of Cultural Economics

PERFORMANCE ANALYSIS OF VARIOUS ENSEMBLE FEATURE SELECTION

*Ms.Sandrilla R¹, Ms.Savitha Devi. M²
Department of Computer Science,
Sacred Heart College, TN, India*

Abstract:

Machine learning modelling encourages the use of multiple models rather than a single model. In general, the most popular method of using multiple models is termed as ensemble where the individual model is grouped together. However, this method has a great impact on feature selection, in which the process of combining the model greatly enhances the better results. Special emphasis is given to elements that are relevant to initial idea. Feature selection is the process of locating and isolating important features while getting rid of outdated or irrelevant features in order to increase classifier accuracy. Hence the paper affords the detailed study for readers that helps them to better understand the basics and new approaches that are required to create an ensemble. We provide a summary of recent advances and predictions, just to lookout for changes in the future. Researchers will fail to discover new paths to excellence if they ignore all of the potential signals of performance in start-up review. By this review the students are grasped to focus on a new subject intensely in order to reach their academic goals in ensemble learning. This study has been conducted to meet the new expectations with substantial number of publications. Once finished reading the content, one will be able to review it and search for easy ways, categories that are made up of distinct metrics. In order to better understand, recognise, and choose features, the work have created a detailed explanation of the feature selection approach. Real-world circumstances are perfect for EFS which may be accomplished using this technique.

Keywords: *Ensemble Feature Selection (EFS), Boosting, Bagging, Stacking, Homogeneous, Heterogeneous*

I. INTRODUCTION

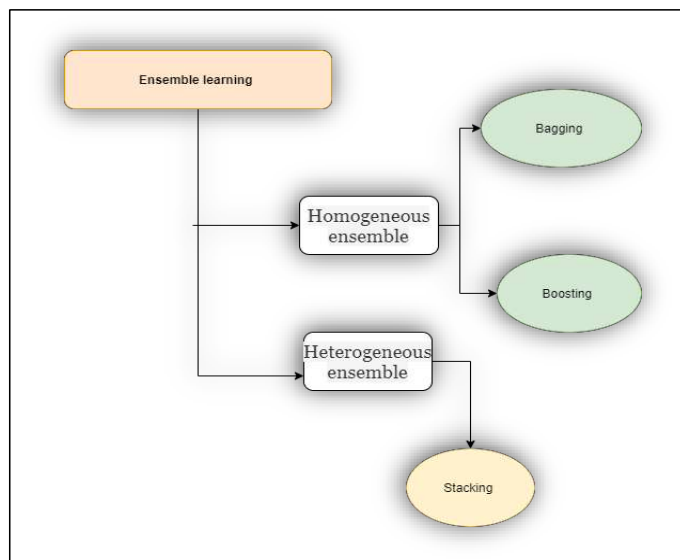
When combining multiple machine learning models, an ensemble model is created where predictions from each model are merged together. Similarly, all ensemble members (or just ensemble members when referring to the entire group of individuals collectively referred to as ensemble members) can be of the same type, but they can also be various types that have been trained on separate training data. When an ensemble member predicts that one of the two possible approaches (i.e., employing the mode or mean) will be used, as well as sophisticated algorithms that learn how much to trust each member and when to trust them, they are making a forecast. The practice of core bagging, boosting, and stacking was introduced in the literature around the year 2000, at the time of ensembles' popularity.

The ensemble group technique is applied, the overall computation complexity increases. Training on many models accounts for the majority of this rise in capability. [6,7] Offering the following advantages in place of using a single model, the ensemble is comprised of multiple models: The resilience of an ensemble is larger than any single model. The ensemble regulates the prediction and model performance as it spreads or diffuses.

Modelling accuracy is increased when ensembles are used. For this circumstance, bias can be used to lower the overall error variance (i.e., as an approach to dealing with the bias-variance trade-off). Additionally, keep in mind that it is also important to have the capacity to enhance the average performance of a model. [13] [14] Despite the unpredictability of the future, we must nevertheless keep machine learning projects' data quality and predictive capability. We favor one or both of a model's attributes from time to time, but it is still important.

II. FEATURE SELECTION METHODS

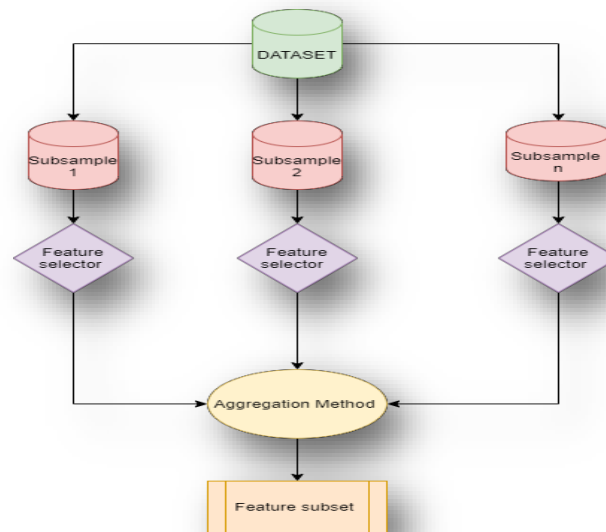
Feature selection procedures can be divided into conservative and experimental approaches. Feature subset selection is a dynamic decision that varies depending on the feature selector (e.g., returning a subset of relevant features or an ordered ranking of all features).



A lower threshold must be set in order to deal with this last circumstance. [17]. A variety of feature selection strategies can be classed as inverse correlation-based and direct correlation-based. [18] Also known as classification methods, filters, which are unconnected to the learning operation, and classifiers, which are connected to the prediction, are both referred to as classification tools. When the system is learning, it is commonly referred to as embedded classification.

Fig.2 Homogeneous Feature Selection Ensembles

Fig.2 Homogeneous Feature Selection Ensembles



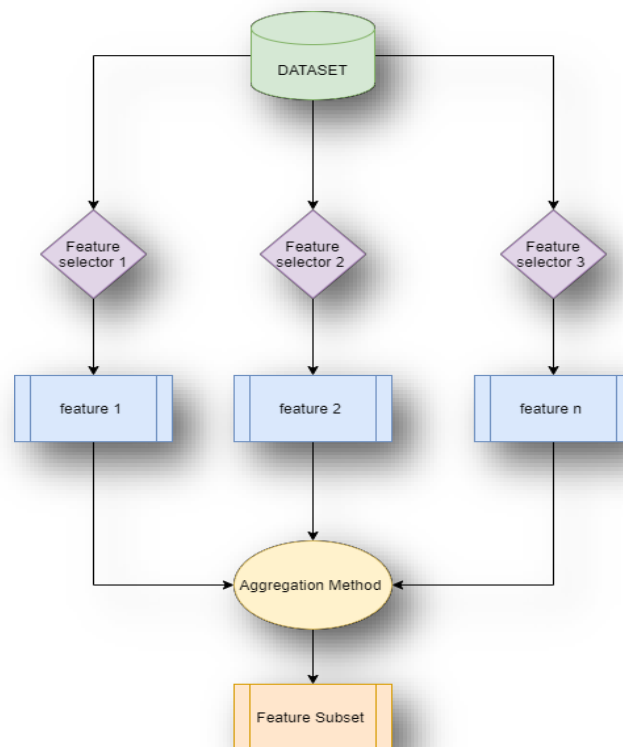
Other methodologies include independent induction processes, as stated in the literature (such as entropy, probability distributions, or information theory). It is imperative that a user be aware of accessible algorithms' properties so a suitable choice may be made. To increase performance, an algorithm compiles a combination of feature selector results into one. Thus, the algorithm allows the user to implement any approach. This study will cover a multitude of ensemble learning approaches.

In general, the ensemble learning considered as the homogeneous and heterogeneous approach described on the figure 1. This study will focus on the most recent examples of both strategies of homogeneous as well as heterogeneous.

Bagging and boosting are the two methods used under homogeneity of ensemble learning. In heterogeneous the stacking has been applied and Feature extraction ensembles are commonly used to do heterogeneous feature selection. In aggregation stage, where the distinct outputs of each feature selector must be joined, is critical to consider as well. The technique picked will dictate whether perform feature subset selection or feature ranking. This research explores other fascinating subjects, like the range and consistency of ensembles, which is covered on the data table 1 with performance measurements [22].

The figure 2 and 3 represents the architecture of homogeneous and heterogeneous feature selection. Two alternative ensemble learning designs were used in the feature selection procedure. In a homogeneous ensemble, the dataset is dispersed among subsamples and various feature selectors are ensembled. In a heterogeneous ensemble, the dataset has been directly worked with various feature selector and the results of different features are aggregated.

To divide characteristics among an ensemble, an independent base selector should be employed. It is an ensemble when there are many types of feature selector working together. It is not ideal if it is homogeneous. The EFS has not always produced the best result of all type of data set. By using the same EFS process, but with various training data subsets, this technique delivers consistent results. Furthermore, as the subsets of data are spread across several nodes, the timing limitations are lowered as well (or divisions). This design is specifically partitioned. Also known as "data variability ensembles," this method uses a number of different datasets to craft new solutions. Additional evidence [26] and [24] help explain situations when different data sets need to be taken into consideration.



As seen above, using multiple feature selection strategies as a design parameter is advised. Heterogeneous feature selection ensembles are also encountered relatively common. Several of these studies utilize both homogeneous and heterogeneous research approaches, as in [26]. In order to acquire either a feature subset or a feature ranking, may use feature selectors and apply an additional threshold step. Multiple aggregation methods may be used to obtain a final best output. This is despite the fact that more recent studies [27] looked into designs where the sequence of combination and thresholding processes is reversed, in which case base feature selectors are rankers [28].

The output of each feature selector is aggregated and returned as the final ensemble result in the second phase. In which a variety of feature selectors are utilized in the first step. Ensemble techniques, similar to supervised learning, could be utilized to improve the robustness of feature selection algorithms. Feature subsets generated by various feature selection algorithms may be considered optimized features in the space of feature subsets, and EFS may provide a better approximation to the optimal subset, ranking, etc. Furthermore, a feature selector's representational power may limit its search space, preventing the finding of optimal subsets. By combining the outputs of many feature pickers, EFS may be able to help solve this problem. This approach was designed for high-dimensional data with few samples, but it may be used with any data.

There has been variety of other methods also be available in order to combine the selected features.

Subset creation

The appropriateness of a group of traits is determined via subset selection. Subset selection algorithms are divided into three categories: wrappers, filters, and embedding strategies. Wrappers use a search algorithm to search over the universe of possible features and evaluate each subset by running a model on it. most widely used subset creation in EFS is union and intersection.

The union creates the subset for all the features selected by the n-feature selector on the ensemble. The intersection considered only common features that available on all the feature selector

Ranking

Before classification or clustering tasks, ranking methods, which are part of the filter method category, undertake feature selection in two steps. In the first, the characteristics are rated using a statistical metric. The features with the highest rankings are then chosen in the second. In ensemble method ranking has been considered for combine the selected features. the min, median and mean are the most widely used ranking method.

III. REVIEW OF THE LITERATURE

TABLE I REVIEW OF VARIOUS ENSEMBLE SELECTION

Reference	year	Method	Performance	
Ref [52]	2021	Fuzzy based Cuckoo Search Ant Colony Optimization	accuracy sensitivity, specificity computational	Efficient than another algorithm

			time.	
Ref [53]	2021	linear regression, random forest, and Xgboost	MSE	0.001
Ref [54]	2021	Information gain SVM-RFE, Lasso/Ridge, Random Forest) from filter, wrapper,	Accuracy f-score	Random forest produces the worst among other
Ref [55]	2021	Pearson's correlation Chi-squared	accuracy, Hamming loss, Jaccard index f-score	Produces best result
Ref. [43]	2020	PCA, GA, C4.3 Serial and parallel ensemble method Intersection Union Multi-Intersection	Classification accuracy Reduction rates	Accuracy GA+PCA=72% Reduction rate 68%
Ref [44]	2020	Sequential forward selection and Support vector machine (SFS-SVM)	accuracy	97.54%, 93.71%
Ref [37]	2020	CUS-GBDT	AUC score ACC score PPV score	0.977497 0.915278 0.996850
Ref [45]	2020	Multi-Rule Based Ensemble Feature Selection	Accuracy	98.81%
Ref [56]	2019	chi-square test, maximum information coefficient and XGBoost,	the arithmetic mean, accuracy	85%

Ref [34]	2019	data perturbation and function perturbation cumulative distribution function	Accuracy	94.6%
Ref [32]	2019	Filter Chi-Square Information Gain (Info Gain) Minimum Redundancy Maximum Relevance • ReliefF Embedded • Recursive Feature Elimination for Support Vector Machines Feature Selection-Perceptron (FS-P)	% average test error	4% with FS
Ref [42]	2019	convolution neural network features of distinct weights	Accuracy	95.50%
Ref [33]	2019	t-test FS, ANOVA-based FS, and FS based on the Pearson correlation coefficient and Gini index robust feature selection RELIEF and FS based on the Fisher criterion	Accuracy F1-score	9.8%
Ref [48]	2019	Cumulative Distribution Function gradient	Accuracy Recall RMSE	94%
Ref [49]	2019	plurality, single-transferable, Borda , weighted-Borda coun	Accuracy Sensitivity	98%
Ref [51]	2019	Radial Basis Function Neural Network Genetic algorithm	Error sensitivity	71%

Ref [47]	2018	InfoGain, GainRatio, ReliefF, Chi-square and Symmetric-Uncertainty	Average Accuracy	95%
Ref [35]	2017	Bi-objective Genetic Algorithm based Feature Select ensemble of multiple such FSGAs	Accuracy Recall Fallout Specificity F1 Score	99 93 1 96
Ref [39]	2017	information gain [53,54], t-statistic [55,56], χ^2 -statistics, gain ratio and correlation	Accuracy Recall Fallout Specificity F1 Score AUC score ACC score PPV score	evaluated with different dataset
Ref [41]	2017	Information Gain (InfoGain) • Minimum Redundancy Maximum Relevance (mRMR) • ReliefF • Recursive Feature Elimination for Support Vector Machines • Feature Selection-Perceptron Fisher discriminant ratio	Average speedup average estimated percentage test errors	50%
Ref [50]	2016	Bayesian logistic regression Ttest and Fisherscore	Feature selection time	45.35%
Ref [40]	2014	mono and bi-objective v PSO, GA and SVM Majority vote, Sum, Weighted sum, support vector machine (SVM), k-nearest neighbour and Naive Bayesian	Accuracy	PSO-based ensembles obtained the best accuracy levels, in both mono and bi-objective versions
Ref [46]	2003	hill-climbing-based refinement cycle	Accuracy f-score	92%

In 2021, ref [52] experimented EFS approach, which combines Fuzzy and ACO Cuckoo Search (FCS) algorithms. The Hybrid ensemble Classification also been experimented with this work. The EEG dataset has been tested with the model. The accuracy, sensitivity reduction, specificity, and computation time has been measured with the current system.

In 2021, ref [53] created a EFS by using heterogeneous way. The model used uses wrapper-based ensemble subset creation for reservoir productivity prediction. The research focused on two crucial components of selector and subset creation. The selector of the EFS model consists 4 model depicted on the table 1. The subset of thickness, depth lateral resistivity difference, gas reserve coefficient, amount of injected liquid, stopped pump pressure, flow-back rate, formation reserve coefficient, and water saturation. Prediction models with the main features outperformed. the proposed work reduces data acquisition costs, model comprehensibility, and computational complexity.

In 2021 ref [54], a Bagging approach with MapReduce EFS framework was evaluated to detect risk indicators. Five models, as well as wrapper and embedding feature selection methods, are included in the framework. Each feature selection model chose ten variables based on the relevance of the characteristics. The models were classified into three groups based on accuracy, F-score, and model properties The voting-based aggregation has been applied to final feature selection on EFS framework. which took into account both feature and model weights. In terms of six interpretability characteristics, we compared our voting approach for picking top-ranked features to two others.

In 2021 ref [55] experimented the student dataset with ensemble EFS. The collection includes both numerical and category data. We chose features using different methods. These methods were chosen because they outperformed other prospective feature selection algorithms in terms of prediction accuracy, or because they are simple and straightforward. After pre-processing, the original dataset comprises a total of 26 characteristics. This was reduced to 18 features using this work. The results show that the 30% feature duplicated.

In 2019 ref [56]. Experimented the EFS framework. To compare the performance of the proposed method to that of existing methods, data pre-processing is utilized to combine 5-feature selection and 3-datasets. The results reveal that the best features are chosen utilizing a sorting integration-based feature selection technique. AUC values are predicted by models based on this attribute. Only a few features, according to this study, can provide useful information to the categorization model. When there are too many features, feature subsets become redundant, lowering the classification model's prediction accuracy.

In 2020 ref [45] deals with EFS Based on ranking algorithms, the authors define two alternative ensemble designs. The work analyzed the best combination for EFS. A new automatic threshold is also built and compared to previous thresholding systems that keep a specific percentage of

attributes. SVM classification accuracy was used to test the performance of these algorithms in three different scenarios.

In 2020 ref [43] used HDLLSS for EFS outperforms single feature selection, the research will determine if parallel or serial combination strategies are superior for EFS. The GA and PCA feature subsets aid the SVM classifier, outperforming those chosen without feature selection. C4.5, the alternative way has the highest rate of feature reduction, resulting in an over selection problem as a result of the elimination of too many features.

In 2020 ref [44] deals with the EFS To save computing time and eliminate disrespectful and noisy features, the paper uses a range of nature categorization learning algorithms to determine which operating engine is best for our model. According to the findings, SVM with ideal features had a training dataset accuracy of 97.54 and a raw dataset accuracy of 93.71. Among the current computational models, the proposed model archives the best results. Both academic and proteomic research will benefit from the new methodology.

In 2020 ref [37], an experimental heterogeneous ensemble technique was used to choose features and generate models. 5 algorithms employed to pick features and the final aggregation the voting has been employed. The model has been experimented with financial crisis prediction. The CUSGBDT and XGBoost heterogeneous ensemble surpasses component learners and single classifiers. In 2019 ref [34] has been focused EFS to identify a suitable subset of characteristics for use in phishing detection. The work adapts to a range of datasets by utilizing patterns in the distribution of filter measure values. The results reveal that the results are competitive when the baseline attributes are paired with the Random Forest classifier.

In 2020, ref [51] deals with the EFS method Sensitivity and ensemble training error are two objectives that are used to influence feature selection. At the same time, the interplay of base classifier learned from feature subsets determined by base selectors is problematic. To find the optimum subsets, The work creates two objective functions: sensitivity and ensemble training error, and use a multi-objective genetic algorithm. The objectives are then minimized to obtain the feature subsets for an ensemble.

In 2019 ref [41] worked with new design on EFS. In a unique feature selection approach dubbed HEFS. As part of both the HEFS and the study develops a general technique that automatically identifies the proper amount of feature. The approach adapts to a range of datasets to identify the filter values. The results reveal that when the baseline attributes are paired with the Random Forest classifier, they are competitive.

In 2019 ref [33] Experimented a lot of EFS algorithms based on voting. have been proposed. Simple FS methods like the Borda-count, STV, or plurality-voting are combined to form complex EFS methods in these methods. More clustering ensemble choices were also offered. The paper evaluated the utility of FS techniques using 3 metrics: FS sensitivity, stability, and

classification accuracy. 5 fake datasets and 10-high-dimensional datasets has been used. When all performance parameters are considered, the current FS technique based on clustered Borda count beats competing alternatives. In terms of Sen rate and prediction efficiency, the ensembles A number of them, though, were having financial troubles. RELIEF outperformed previous techniques in terms of FS sensitivity and stability, but not in terms of prediction efficiency. The T-test FS has an above-average sensitivity, but its stability and prediction are below-average. Fisher produces equal results across all phases, but when compared to ensembles, they were just average.

In 2019, [42] propose new ensemble technique based on RNA and protein sequence knowledge. By deciding whether the convolution neural network is fine-tuned and the features of distinct weights, the technique generates a huge number of data sets. The base classifier which then uses a weighted voting method to select the most likely categories as the final prediction result. These amazing results show that proposed EFS is a reliable and competitive predictor of RNA-protein interactions. The paper will continue to develop the ensemble strategy in the future in order to achieve a higher rate.

In 2018 ref [36] experimented A novel MRFES algorithm was created as a solution. This MCCM paradigm was designed to help a community of co-evolutionary memplexes attain a decision-making agreement. A series of extensive experiments were used to demonstrate MRFES' detailed complexity study, which encompassed aspects of precision, performance, and resilience. The application of proposed work is for human cerebral classification in brain research was established in this study.

In 2018 ref [] published a work of EFS with Mutual-Information (EFS-MI) the work has been experimented using the UCI, network, and gene expression datasets. The overall performance of all of these datasets has been proven to be exceptional. The result considered that the classification accuracy analysis, the proposed EFS-MI addresses the local optimal problem of individual filters in most circumstances, notably for high-dimensional datasets. The theoretical quality of EFS-MI-identified features has also been established in terms of relevance and non-redundancy. To minimize the bias induced by a single classifier, an ensemble of classifiers using a soft computing approach is being developed.

In 2018 ref [38], provides an approach for identifying the most useful spectrum properties for plant phenotyping applications using an EFS methodology. To rate spectral features, the ensemble was formed using six feature selection algorithms. while boosting the precision of discriminating salt-treated vegetation pixels from control pixels by 8.5. the work must first convert a hyperspectral data set to a multispectral data set before converting it to a multispectral data set. In a similar way to the hyperspectral data set, the resulting multispectral data set was used to assess the salt tolerance of the four wheat lines. This shows that the recommended feature selection technique may be used to find the most useful features.

In 2017 ref [32] experimented EFS, instead of using a single technique, an ensemble of filters and embedding algorithms was used. The goal was to rise the constancy while also allowing for variety to take use of the various pickers' strengths and limitations. Depending on how the data was distributed and the feature selection methods utilized, the paper used a Support Vector Machine (SVM) classifier in two different ways. The suitability of the proposed ensembles is demonstrated by experimental validation of the technique using seven different datasets. In a real-world setting, the user does not have to choose which feature selection approach is best for a specific problem, which is a benefit.

In 2017, [39] introduced a novel automated figurative content identification algorithm. To analyze satiric news and sarcastic customer feedback, the ensembled function subset is fed into a variety of binary classifiers. In addition, the inquiry found fascinating characteristics of satiric news and ironic response. As a result, they're satirical or sarcastic. The article also looked at how humor and irony are similar and different.

In 2017 ref [34] Introduced an EFS technique combining a bi-objective evolutionary algorithm and a dynamic aggregation. The recommended genetic algorithm, which can quickly categorize objects and develop the most accurate and informative feature subsets, develops the most accurate and informative feature subsets because objective functions are established using rough set theory and knowledge theory. It's a wonderful match because the suggested method is naturally parallel. The proposed methodology outperforms the majority of existing state-of-the-art methods across all datasets, coming in second only to DSFFC in a few cases. In addition, the experimental findings are statistically analyzed. Although it has many more applications, the system uses a pareto-based approach to solve two objective optimization issues

IV. PERFORMANCE METRICS

True Negatives (TN) - correctly predicted negative values

False Positives (FP) – When actual class is no and predicted class is yes

False Negatives (FN) – When actual class is yes but predicted class is no

$$\frac{\text{TN} + \text{FP}}{\text{TN} + \text{FP} + \text{FN} + \text{TP}} = \frac{\text{TN} + \text{FP}}{\text{TN} + \text{FP} + \text{FN} + \text{TP}} \quad (1)$$

(2)

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$J_{\text{index}} = \frac{TP}{TP + FN} \quad (3)$$

$$J_{\text{index}} = \frac{2 \times (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})} \quad (4)$$

$$P_{\text{index}} = 10 \times \frac{\sum_{i=1}^n \frac{1}{n_i^2}}{n} \quad (5)$$

$$J_{\text{index}} = \frac{1}{\sum_{i=1}^n \sum_{j=1}^n \frac{1}{n_i n_j} [(C_{ij} - C_{ii})(C_{ij} - C_{jj})]} \quad (7)$$

$i=0 \quad j=0$

$$P_{\text{index}} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{1}{n_i^2}} \times 100 \quad (8)$$

The stability and diversity have been the two evaluation measures on the EFS to proof the strength of the selected feature over the classification method.

The stability of the feature preferences it generates when training data is disrupted is referred to as robustness. The reproducibility of a feature selection approach is determined by its stability. High feature selection algorithm stability is equally as important as high classification accuracy when evaluating feature selection performance.

The stability has been measured by various measures the most widely used is J-index and

The equation describes the j-index from the confusion matrix

$$J_{\text{index}} = \frac{P}{P + N} \quad (9)$$

The ensemble diversity can be calculated in a number of different methods. Two of the most widely utilized performance assessments are the correlation coefficient kappa statistic (CCKP).

They're all paired because they can measure variation in classifier predictions. The overall ensemble diversity is the average of all the classifier pairs in the ensemble.

V. DISCUSSION

Modelling with two or more models is an example of employing an ensemble. Predictive modelling is more likely to deliver the most relevant outcomes. While machine learning methods applied to a particular problem may not always be the optimal solution, teams and particular approaches are far more frequently used. Accurate forecasts can be more easily generated using ensemble techniques, and if 're intending to use the ensemble to accomplish the project, it's important that understand how it accomplishes this.

Each machine learning model contributes to the forecasts. Based on training data, members of the same type can be differentiated, however those of a different type may or may not be able to be identified. The estimates of team members can be integrated using complex algorithms that learn which team members to trust and when. The widespread use of collaborative research began in the early 1990s, and that's when collaborative study truly took off. Widely used and well-known processes were disseminated as a result. To a limited extent, at the end of the 21st century, these teams were credited for their machine learning (ML) accomplishments, particularly their strong showings in the Netflix challenge and following Kaggle tournaments. Computations in the research are complicated due to the various approaches utilized. Due to this increase in size, the talent and time required has increased.

Teams apply their approaches to create better prediction results via predictive modelling. Use of a load on the model helps to decrease the model's variability. The greatest benefit of cooperation is that it allows everyone to meet deadlines. It has two noteworthy drawbacks: One is a design defect, and the other is an established feature of the model.

Project demands will determine the modelling techniques we use, and we will choose the one that yields the best results. fresh data projections are then made utilizing the algorithm or pipe when all available data has been factored in. Model error is an estimate of projected performance across different subsets of data.

Combining forecasts, such as regression-average statistics and classification modes, is performed first, and then this data is compared to the model's predictions. In addition, there should be different models so that learning will not be just determined by chance, the makeup of the learning data collection has an impact, and the varied models all play a part. This is implemented to restrict dissemination. It is plausible to predict that, on the basis of these pieces of information, the mean score will be equal to the median, with the lower and upper ends approximately equal to the mean. As a result, the model's actual output will likewise be smoothed. Even if it only takes some of the dispersion out, the team technique at least reduces the spread of outcomes.

Neural classifiers or regressors should learn how to apply mapping functions to integrate input and output data. The data that was not utilized while teaching the mapping, which is referred to as the test data set, is used to demonstrate how to make the mapping. Deviation and variance are

employed in machine learning to represent prior model faults. Postulates: hold true when describing functional data input and output representation Model variance is the variance in the model with respect to various training data. When it comes to depicting how the model's influence on data attributes is accounted for, it is true to say that it illustrates how the model is taken into consideration. Varying input leads to varying outcome. Adding more people should be a clear and straightforward technique to expand the number of participants in the room. Tables are used to boost predictive efficiency, which minimizes error fluctuation.

Forecast error variance is reduced, and this results in better predictions. In a more complex dataset, arrays could allow to correct for a regression line error or improved classification accuracy. Among machine learning competitors who have won events, a machine learning kit is commonly selected. If the performance of this team is on average better than everyone else on the team, the choice should be executed. In other words, it is also necessary to have a separate supporting function. Even if this occurs, the presentation will be a total success. The first condition is in effect when the team has only one remarkable model, and the others have a lower impact. This can also happen if the team achieves a worse outcome than the member's personal best. Great performers could be impeded when a faulty prediction is made by one or more of the other models. This should improve awareness.

VI. CONCLUSION

In order to extract knowledge from high-dimensional data, knowledge discovery is critical. Data mining is a step in the process of knowledge discovery. To improve the quality of data for data mining, pre-processing techniques are used. The mining, categorization, and grouping of connected rules are all data mining tasks. Data mining techniques are widely used to analyze trends in records and find ways to improve them. When creating a predictive model, feature selection entails reducing the number of input variables. In some cases, it is beneficial to lowering the variable to save money on computer modelling while also improving model performance. Statistical selection approaches comprise utilizing statistics to evaluate the relationship between each input variable and the goal variable and then selecting the input variables with the strongest connection to the target variable. Although the choice of statistical measures is dependent on the input and output data types, these methods can be quick and efficient. As a result, while selecting essential features, it may be challenging for a machine learning to find a statistically appropriate measure for a dataset. In this study, the various EFS was examined in conjunction with the performance. The ensemble has a number of options. An ensemble is a collection of two or more models. Predictive modelling is more likely to yield meaningful outcomes. While applying master learning methods to a given problem isn't always the greatest option, teams and specific approaches are used far more frequently. Ensemble techniques make it easier to make precise forecasts, and knowing how to do so is critical if you want to use the ensemble to complete the project.

REFERENCES

- [1]. Sagi, Omer, and Lior Rokach. "Ensemble learning: A survey." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, no. 4 (2018): e1249.
- [2]. Zhang, Cha, and Yunqian Ma, eds. *Ensemble machine learning: methods and applications*. Springer Science & Business Media, 2012.
- [3]. Liu, Yong, and Xin Yao. "Ensemble learning via negative correlation." *Neural networks* 12, no. 10 (1999): 1399-1404.
- [4]. Guan, Donghai, Weiwei Yuan, ng-Koo Lee, Kamran Najeebullah, and Mostofa Kamal Rasel. "A review of ensemble learning based feature selection." *IETE Technical Review* 31, no. 3 (2014): 190-198.
- [5]. Yang, Pengyi, Yee Hwa Yang, Bing B Zhou, and Albert Y Zomaya. "A review of ensemble methods in bioinformatics." *Current Bioinformatics* 5, no. 4 (2010): 296-308.
- [6]. Dey, Ayon. "Machine learning algorithms: a review." *International Journal of Computer Science and Information Technologies* 7, no. 3 (2016): 1174-1179.
- [7]. Dong, Xibin, Zhiwen Yu, Wenming Cao, Yifan Shi, and Qianli Ma. "A survey on ensemble learning." *Frontiers of Computer Science* 14, no. 2 (2020): 241-258.
- [8]. Asmita, Shruti, and K. K. Shukla. "Review on the architecture, algorithm and fusion strategies in ensemble learning." *International Journal of Computer Applications* 108, no. 8 (2014).
- [9]. Krogh, Anders, and Peter Sollich. "Statistical mechanics of ensemble learning." *Physical Review E* 55, no. 1 (1997): 811.
- [10]. G. Brown, *Ensemble learning*, in: *Encyclopedia of Machine Learning*, Springer, 2011, pp. 312–320.
- [11]. V. Bolón-Canedo, N. Sánchez-Marño, A. Alonso-Betanzos, *Feature selection for high-dimensional data*, Springer, 2015.
- [12]. I. Guyon, A. Elisseeff, *An introduction to variable and feature selection*, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
- [13]. V. Bolón-Canedo, N. Sánchez-Marño, A. Alonso-Betanzos, *Feature selection for high-dimensional data*, *Progr. Artif. Intell.* 5 (2) (2016) 65–75.
- [14]. P. Cunningham, J. Carney, *Diversity versus quality in classification ensembles based on feature selection*, in: R. Lpez de Mntaras, E. Plaza (Eds.), *Proc. European Conference on Machine Learning (ECML)*, LNAI 1810, 2000, pp. 109–116.
- [15]. D. Opitz, *Feature selection for ensembles*, in: *Proc. 16th Nat. Conf. on Artificial Intelligence*, AAAI Press, 1999, pp. 379–384.
- [16]. Y. Saeys, T. Abeel, Y. Van der Peer, *Robust feature selection using EFStechniques*, in: W. Daelemans, et al. (Eds.), *Proc. European Conference on Machine Learning (ECML PKDD)*, LNAI 5212, 2008, pp. 313–325.
- [17]. Saeys, Yvan, Thomas Abeel, and Yves Van de Peer. "Robust feature selection using EFStechniques." In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 313-325. Springer, Berlin, Heidelberg, 2008.
- [18]. Tsymbal, Alexey, MykolaPechenizkiy, and Pádraig Cunningham. "Diversity in search strategies for ensemble feature selection." *Information fusion* 6, no. 1 (2005): 83-98.
- [19]. Opitz, David W. "Feature selection for ensembles." *AAAI/IAAI* 379 (1999): 384.
- [20]. Tsymbal, Alexey, Seppo Puuronen, and David W. Patterson. "EFSwith the simple Bayesian classification." *Information fusion* 4, no. 2 (2003): 87-100.
- [21]. Abeel, Thomas, Thibault Helleputte, Yves Van de Peer, Pierre Dupont, and Yvan Saeys. "Robust biomarker identification for cancer diagnosis with EFSmethods." *Bioinformatics* 26, no. 3 (2010): 392-398.
- [22]. Wang, Huanjing, Taghi M. Khoshgoftaar, and Amri Napolitano. "A comparative study of EFStechniques for software defect prediction." In *2010 Ninth International Conference on Machine Learning and Applications*, pp. 135-140. IEEE, 2010.
- [23]. Brahim, Afef Ben, and Mohamed Limam. "EFSfor high dimensional data: a new method and a comparative study." *Advances in Data Analysis and Classification* 12, no. 4 (2018): 937-952.

- [24]. B. Seijo-Pardo, V. Bolón-Canedo, A. Alonso-Betanzos, Testing different ensemble configurations for feature selection, *Neural Process. Lett.* 46 (2017) 857–880.
- [25]. P. Granitto, P. Verdes, H. Ceccatto, Neural network ensembles: evaluation of aggregation algorithms, *Artif. Intell.* 163 (2) (2005) 139–162.
- [26]. S.E. Lacy, M.A. Lones, S.L. Smith, A comparison of evolved linear and non-linear ensemble vote aggregators, in: 2015 IEEE Congress on Evolutionary Computation (CEC), 2015, pp. 758–763, doi:10.1109/CEC.2015.7256967.
- [27]. M. Haque, N. Noman, R. Berretta, P. Moscato, Heterogeneous ensemble combination search using genetic algorithm for class imbalanced data classification, *PLoS ONE* 11 (2016) e0146116.
- [28]. A. Canuto, M. Abreu, L. de Melo Oliveira, J.J. Xavier, A. Santos, Investigating the influence of the choice of the ensemble members in accuracy and diversity of selection-based and fusion-based methods for ensembles, *Pattern Recognit. Lett.* 28 (2007) 472–486.
- [29]. V. Bolón-Canedo, N. Sánchez-Marroño, A. Alonso-Betanzos, An ensemble of filters and classifiers for microarray data classification, *Pattern Recognit.* 45 (2012) 531–539.
- [30]. H. Wang, T.M. Khoshgoftaar, A. Napolitano, A comparative study of EFS techniques for software defect prediction, in: 2010 Ninth International Conference on Machine Learning and Applications, 2010, pp. 135–140, doi:10.1109/ICMLA.2010.27.
- [31]. B. Pes, N. Dess, M. Angioni, Exploiting the ensemble paradigm for stable feature selection: a case study on high-dimensional genomic data, *Inf. Fusion* 35 (2017) 132–147.
- [32]. Seijo-Pardo, Borja, Verónica Bolón-Canedo, and Amparo Alonso-Betanzos. "On developing an automatic threshold applied to feature selection ensembles." *Information Fusion* 45 (2019): 227-245.
- [33]. Drotár, Peter, Matej Gazda, and Liberios Vokorokos. "EFS using election methods and ranker clustering." *Information Sciences* 480 (2019): 365-380.
- [34]. Chiew, Kang Leng, Choon Lin Tan, KokSheik Wong, Kelvin SC Yong, and Wei King Tiong. "A new hybrid EFS framework for machine learning-based phishing detection system." *Information Sciences* 484 (2019): 153-166.
- [35]. Das, Asit K., Sunanda Das, and Arka Ghosh. "EFS using bi-objective genetic algorithm." *Knowledge-Based Systems* 123 (2017): 116-127.
- [36]. Ding, Weiping, Chin-Teng Lin, and Witold Pedrycz. "Multiple relevant feature ensemble selection based on multilayer co-evolutionary consensus MapReduce." *IEEE transactions on cybernetics* 50, no. 2 (2018): 425-439.
- [37]. Du, Xudong, Wei Li, Sumei Ruan, and Li Li. "CUS-heterogeneous ensemble-based financial distress prediction for imbalanced dataset with ensemble feature selection." *Applied Soft Computing* 97 (2020): 106758.
- [38]. Moghimi, Ali, Ce Yang, and Peter M. Marchetto. "EFS for plant phenotyping: A journey from hyperspectral to multispectral imaging." *IEEE Access* 6 (2018): 56870-56884.
- [39]. Ravi, Kumar, and Vadlamani Ravi. "A novel automatic satire and irony detection using ensembled feature selection and data mining." *Knowledge-based systems* 120 (2017): 15-33.
- [40]. Santana, Laura Emmanuella A. dos S., and Anne M. de Paula Canuto. "Filter-based optimization techniques for selection of feature subsets in ensemble systems." *Expert Systems with Applications* 41, no. 4 (2014): 1622-1631.
- [41]. Seijo-Pardo, Borja, Iago Porto-Díaz, Verónica Bolón-Canedo, and Amparo Alonso-Betanzos. "Ensemble feature selection: homogeneous and heterogeneous approaches." *Knowledge-Based Systems* 118 (2017): 124-139.
- [42]. Wang, Lei, Xin Yan, Meng-Lin Liu, Ke-Jian Song, Xiao-Fei Sun, and Wen-Wen Pan. "Prediction of RNA-protein interactions by combining deep convolutional neural network with feature selection ensemble method." *Journal of theoretical biology* 461 (2019): 230-238.
- [43]. Tsai, Chih-Fong, and Ya-Ting Sung. "EFS in high dimension, low sample size datasets: Parallel and serial combination approaches." *Knowledge-Based Systems* 203 (2020): 106097.

- [44]. Ahmad, Ashfaq, Shahid Akbar, Maqsood Hayat, Farman Ali, and Mohammad Sohail. "Identification of antioxidant proteins using a discriminative intelligent model of k-spaced amino acid pairs based descriptors incorporating with ensemble feature selection." *Biocybernetics and Biomedical Engineering* (2020).
- [45]. Sundararajan, Karthik, and Anandhakumar Palanisamy. "Multi-rule based EFS model for sarcasm type detection in twitter." *Computational intelligence and neuroscience* 2020 (2020).
- [46]. Tsymbal, Alexey, Seppo Puuronen, and David W. Patterson. "EFS with the simple Bayesian classification." *Information fusion* 4, no. 2 (2003): 87-100.que
- [47]. Hoque, Nazrul, Mihir Singh, and Dhruba K. Bhattacharyya. "EFS-MI: an EFS method for classification." *Complex & Intelligent Systems* 4, no. 2 (2018): 105-118.
- [48]. Chiew, Kang Leng, Choon Lin Tan, Kok Sheik Wong, Kelvin SC Yong, and Wei King Tiong. "A new hybrid EFS framework for machine learning-based phishing detection system." *Information Sciences* 484 (2019): 153-166.
- [49]. Drotár, P., Gazda, M. and Vokorokos, L., 2019. EFS using election methods and ranker clustering. *Information Sciences*, 480, pp.365-380.
- [50]. Mollaei, Maryam, and Mohammad Hossein Moattar. "A novel feature extraction approach based on EFS and modified discriminant independent component analysis for microarray data classification." *Biocybernetics and Biomedical Engineering* 36, no. 3 (2016): 521-529.
- [51]. Ng, Wing WY, Yuxi Tuo, Jianjun Zhang, and Sam Kwong. "Training error and sensitivity-based ensemble feature selection." *International Journal of Machine Learning and Cybernetics* 11, no. 10 (2020): 2313-2326.
- [52]. Banu, N. Sharmila, and S. Suganya. "EFS (EFS) and ensemble hybrid classifiers (EHCS) for diagnosis of seizure using EEG signals." *ICTACT J. Soft Comput* 11, no. 2 (2021): 2283-2287.
- [53]. Zhou, Changlin, Lang Zhou, Fei Liu, Weihua Chen, Qian Wang, Keliang Liang, Wenqiu Guo, and Liying Zhou. "A Novel Stacking Heterogeneous Ensemble Model with Hybrid Wrapper-Based Feature Selection for Reservoir Productivity Predictions." *Complexity* 2021 (2021).
- [54]. Shi, Xi, Gorana Nikolic, Gorka Epelde, Mónica Arrúe, Joseba Bidaurrezaga Van-Dierdonck, Roberto Bilbao, and Bart De Moor. "An ensemble-based feature selection framework to select risk factors of childhood obesity for policy decision making." *BMC medical informatics and decision making* 21, no. 1 (2021): 1-13.
- [55]. Yekun, Ephrem Admasu, and Abrahaley Teklay Haile. "Student Performance Prediction with Optimum Multilabel Ensemble Model." *Journal of Intelligent Systems* 30, no. 1 (2021): 511-523.
- [56]. Wang, Jie, Jing Xu, Chengan Zhao, Yan Peng, and Hongpeng Wang. "An EFS method for high-dimensional data based on sort aggregation." *Systems Science & Control Engineering* 7, no. 2 (2019): 32-39.

Data Mining – Techniques, Methods and Algorithms: A Review on Tools

Ms. L. Hemalatha

HOD, Dept of Computer Science

MKJC, Vaniyambadi, TN, India

ABSTRACT

Data Mining is the process of extracting the useful data, patterns and trends from a large amount of data by using techniques like clustering, Classification, Association and regression. This techniques provides large number of applications in our daily life. For those techniques various algorithms and tools are available. This paper discusses few of the data mining Tools, Algorithms and some of the data mining tools to be implemented in business and find out the relevant results.

1 INTRODUCTION

Data mining is the process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems. Data mining is an interdisciplinary subfield of computer science and statistics with an overall goal to extract information (with intelligent methods) from a data set and transform the information into a comprehensible structure for further use. Data mining is the analysis step of the "knowledge discovery in databases" process, or KDD. Aside from the raw analysis step, it also involves database and data management aspects, data pre-processing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating. The difference between data analysis and data mining is that data analysis is to summarize the history such as analyzing the effectiveness of a marketing campaign, in contrast, data mining focuses on using specific machine learning and statistical models to predict the future and discover the patterns among data.

Data mining is a field of research that has emerged in the 1990s, and is very popular today, sometimes under different names such as “**big data**” and “**data science**“, which have a similar meaning. To give a short definition of **data mining**, it can be defined as a set of techniques for automatically analyzing data to discover interesting knowledge or patterns in the data. The reasons why data mining has become popular is that storing data electronically has become very cheap and that transferring data can now be done very quickly thanks to the fast computer networks that we have today. Thus, many organizations now have **huge amounts of data** stored in databases, that needs to be analyzed. Having a lot of data in databases is great. However, to really benefit from this data, it is necessary to analyze the data to understand it. Having data that we cannot understand or draw meaningful conclusions from it is useless. So how to analyze the data stored in large databases? Traditionally, data has been analyzed by hand to discover interesting knowledge. However, this is time-consuming, prone to error, doing this may miss some important information, and it is just not realistic to do this on large databases. To address this problem, automatic techniques have been designed to **analyze data and extract interesting patterns, trends or other useful information**. This is the purpose of **data mining**. In general, data mining techniques are designed either to **explain or understand the past** (e.g. why a plane has crashed) or **predict the future** (e.g. predict if there will be an earthquake tomorrow at a given location). Data mining techniques are used to take decisions based on facts rather than intuition.

What is the process for analyzing data?

To perform data mining, a process consisting of seven steps is usually followed. This process is often called the “**Knowledge Discovery in Database**” (KDD) process.

1. **Data cleaning**: This step consists of cleaning the data by removing noise or other inconsistencies that could be a problem for analyzing the data.
2. **Data integration**: This step consists of integrating data from various sources to prepare the data that needs to be analyzed. For example, if the data is stored in multiple databases or file, it may be necessary to integrate the data into a single file or database to analyze it.
3. **Data selection**: This step consists of selecting the relevant data for the analysis to be performed.

4. **Data transformation:** This step consists of transforming the data to a proper format that can be analyzed using data mining techniques. For example, some data mining techniques require that all numerical values are normalized.
5. **Data mining:** This step consists of applying some data mining techniques (algorithms) to analyze the data and discover interesting patterns or extract interesting knowledge from this data.
6. **Evaluating the knowledge that has been discovered:** This step consists of evaluating the knowledge that has been extracted from the data. This can be done in terms of objective and/or subjective measures.
7. **Visualization:** Finally, the last step is to visualize the knowledge that has been extracted from the data.

Of course, there can be variations of the above process. For example, some data mining software are interactive and some of these steps may be performed several times or concurrently.

The actual data mining task is the semi-automatic or automatic analysis of large quantities of data to extract previously unknown, interesting patterns such as groups of data records (cluster analysis), unusual records (anomaly detection), and dependencies (association rule mining, sequential pattern mining). This usually involves using database techniques such as spatial indices. These patterns can then be seen as a kind of summary of the input data, and may be used in further analysis or, for example, in machine learning and predictive analytics. For example, the data mining step might identify multiple groups in the data, which can then be used to obtain more accurate prediction results by a decision support system. Neither the data collection, data preparation, nor result interpretation and reporting is part of the data mining step, but do belong to the overall KDD process as additional steps.

2 OVERVIEW OF DATA MINING

The development of Information Technology has generated large amount of databases and huge data in various areas. The research in databases and information technology has given rise to an approach to store and manipulate this precious data for further decision making. Data mining is a process of extraction of useful information and patterns from huge data. It is also

called as knowledge discovery process, knowledge mining from data, knowledge extraction or data /pattern analysis.

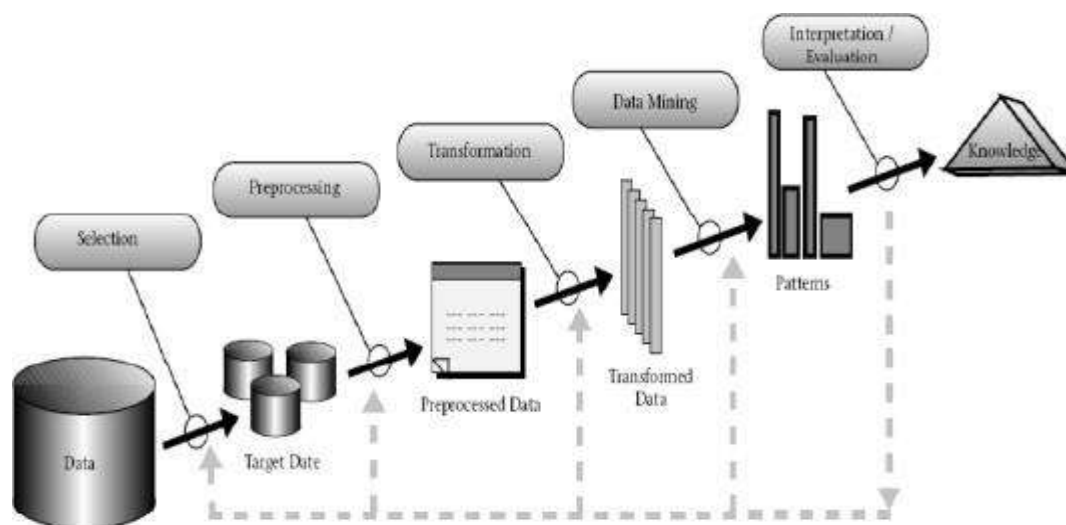


Figure 1. Knowledge discovery Process

Data mining is a logical process that is used to search through large amount of data in order to find useful data. The goal of this technique is to find patterns that were previously unknown. Once these patterns are found they can further be used to make certain decisions for development of their businesses.

Three steps involved are

- Exploration
- Pattern identification
- Deployment

Exploration: In the first step of data exploration data is cleaned and transformed into another form, and important variables and then nature of data based on the problem are determined.

Pattern Identification: Once data is explored, refined and defined for the specific variables the second step is to form pattern identification. Identify and choose the patterns which make the best prediction.

Deployment: Patterns are deployed for desired outcome.

Data Mining is defined as extracting information from huge sets of data. In other words, we can say that data mining is the procedure of mining knowledge from data. The information or knowledge extracted so can be used for any of the following applications,

- Market Analysis
- Fraud Detection
- Customer Retention
- Production Control
- Science Exploration

3 . DATA MINING ALGORITHMS AND TECHNIQUES

Various algorithms and techniques like Classification, Clustering, Regression, Artificial Intelligence, Neural Networks, Association Rules, Decision Trees, Genetic Algorithm, Nearest Neighbor method etc., are used for knowledge discovery from databases.

Classification

Classification is the most commonly applied data mining technique, which employs a set of pre-classified examples to develop a model that can classify the population of records at large. Fraud detection and credit risk applications are particularly well suited to this type of analysis. This approach frequently employs decision tree or neural network-based classification algorithms. The data classification process involves learning and classification. In Learning the training data are analyzed by classification algorithm. In classification test data are used to estimate the accuracy of the classification rules. If the accuracy is acceptable the rules can be applied to the new data tuples. For a fraud detection application, this would include complete

records of both fraudulent and valid activities determined on a record-by-record basis. The classifier-training algorithm uses these pre-classified examples to determine the set of parameters required for proper discrimination. The algorithm then encodes these parameters into a model called a classifier.

Types of classification models:

- Classification by decision tree induction
- Bayesian Classification
- Neural Networks
- Support Vector Machines (SVM)
- Classification Based on Associations

Clustering

Clustering can be said as identification of similar classes of objects. By using clustering techniques we can further identify dense and sparse regions in object space and can discover overall distribution pattern and correlations among data attributes. Classification approach can also be used for effective means of distinguishing groups or classes of object but it becomes costly so clustering can be used as preprocessing approach for attribute subset selection and classification. For example, to form group of customers based on purchasing patterns, to categories genes with similar functionality.

Types of clustering methods

- Partitioning Methods
- Hierarchical Agglomerative (divisive) methods
- Density based methods
- Grid-based methods
- Model-based methods

Prediction

Regression technique can be adapted for predication. Regression analysis can be used to model the relationship between one or more independent variables and dependent variables. In data mining independent variables are attributes already known and response variables are what we want to predict. Unfortunately, many real-world problems are not simply prediction. For instance, sales volumes, stock prices, and product failure rates are all very difficult to predict because they may depend on complex interactions of multiple predictor variables. Therefore, more complex techniques (e.g., logistic regression, decision trees, or neural nets) may be necessary to forecast future values. The same model types can often be used for both regression and classification. For example, the CART (Classification and Regression Trees) decision tree algorithm can be used to build both classification trees (to classify categorical response variables) and regression trees (to forecast continuous response variables). Neural networks too can create both classification and regression models.

Types of regression methods

- Linear Regression
- Multivariate Linear Regression
- Nonlinear Regression
- Multivariate Nonlinear Regression

Association Rule

Association and correlation is usually to find frequent item set findings among large data sets. This type of finding helps businesses to make certain decisions, such as catalogue design, cross marketing and customer shopping behavior analysis. Association Rule algorithms need to be able to generate rules with confidence values less than one. However the number of possible Association Rules for a given dataset is generally very large and a high proportion of the rules are usually of little (if any) value.

Types of association rule

- Multilevel association rule
- Multidimensional association rule
- Quantitative association rule

Neural networks

Neural network is a set of connected input/output units and each connection has a weight present with it. During the learning phase, network learns by adjusting weights so as to be able to predict the correct class labels of the input tuples. Neural networks have the remarkable ability to derive meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. These are well suited for continuous valued inputs and outputs. For example handwritten character reorganization, for training a computer to pronounce English text and many real world business problems and have already been successfully applied in many industries. Neural networks are best at identifying patterns or trends in data and well suited for prediction or forecasting needs.

Types of neural networks

- Back PropagationFeedforward Neural Network – Artificial Neuron
- Radial basis function Neural Network
- Kohonen Self Organizing Neural Network
- Recurrent Neural Network(RNN) – Long Short Term Memory
- Convolutional Neural Network
- Modular Neural Network

4 DATA MINING APPLICATIONS

Data mining is a relatively new technology that has not fully matured. Despite this, there are a number of industries that are already using it on a regular basis. Some of these organizations include retail stores, hospitals, banks, and insurance companies. Many of these organizations are combining data mining with such things as statistics, pattern recognition, and other important tools. Data mining can be used to find patterns and connections that would otherwise be difficult to find. This technology is popular with many businesses because it allows

them to learn more about their customers and make smart marketing decisions. Here is overview of business problems and solutions found using data mining technology.

6 DATA MINING TOOLS

There are various open source tools available for data mining. Some of tools work for clustering, some for classification, regression, association and some for all. We are going to implement our algorithms in the following tools

- Orange
- Scavis
- ELKI
- GraphLab
- Scikit Learn

7. CONCLUSION

Data mining has importance regarding finding the patterns, forecasting, discovery of knowledge etc., in different business domains. Data mining techniques and algorithms such as classification, clustering etc., helps in finding the patterns to decide upon the future trends in businesses to grow. Data mining has wide application domain almost in every industry where the data is generated that's why data mining is considered one of the most important frontiers in database and information systems and one of the most promising interdisciplinary developments in Information Technology. By using various algorithms and tools we are going to illustrate our result.

REFERENCE:

1. hridviRaj MSB. , GuruRao CV (2013) Data mining – past, present and future – a typical survey on data streams. INTER-ENG Procedia Technology 12:255 – 263
2. Srivastava S (2014) Weka: A Tool for Data preprocessing, Classification, Ensemble, Clustering and Association Rule Mining. International Journal of Computer Applications (0975 – 8887) 88:. 10
3. Soni N, Ganatra A (2012) Categorization of Several Clustering Algorithms from Different Perspective: A Review. IJARCSSE
4. Demšar J, Zupan B (2013) Orange: Data Mining Fruitful and Fun - A Historical Perspective. Informatica 37:55–60
5. Jain AK, Murty MN, Flynn PJ (1999) Data Clustering: A Review. ACM Computing Surveys, 31:264-323
6. Han J, Kamber M (2001) Data Mining. Kaufmann Publishers, Morgan Rao IKR (2003) Data Mining and Clustering Techniques DRTC Workshop on Semantic Web, pp. 23-30
7. Mitra S, Pal KS, Mitra P (2002) Data Mining in Soft Computing Framework: A Survey. IEEE, 13: 3-14

ONLINE TEACHING TOOLS USED DURING THE PANDEMIC PERIOD WITH SPECIAL REFERENCE TO SCHOOLS AND COLLEGES IN VELLORE DISTRICT

*M.Ashtalakshmi
Assistant Professor & Head
Department of Commerce CA
Marudhar Kesari Jain College for Women
Vaniyambadi, TN, India.*

Abstract

This paper highlights online teaching tools used by school and college teachers, its benefits and problem faced during the pandemic period in Vellore District of Tamilnadu.

Key words: Online Education, Online Teaching, Online teaching tools

INTRODUCTION

2020 has been a transformative year. Everything has changed, including the way kids and university students go to school. During the pandemic, most of the world insisted social distancing rules, 20 second hand washing practices and preventive mask wearing in public. To say, this affected the way educators and students interact is an understatement. In March millions of children transitioned from face to face classrooms to some kind of virtual education set up at home. During the first months of the pandemic, teachers as to scramble and find the best ways to set up a Virtual classroom that would keep their students engaged. During these months' educators at every level tested tools and programs until they found their favourites.

Online Education

Online education is electronically supported learning that relies on the Internet for teacher/student interaction and the distribution of class materials.

Online Teaching

Online teaching is the process of educating others via the internet. Various methods can be used, such as one-on-one video calls, group video calls, and webinars. You can start teaching from any location (home, coffee shop, co-working space) and enrol students from various backgrounds and geographical areas.

Online Teaching Tools

Online learning tools refer to any program, app, or technology that can be accessed via an internet connection and enhance a teacher's ability to present information and a student's ability to access that information.

From this simple definition comes an almost infinite number of ways to teach and learn outside of traditional classrooms and away from college campuses. With online education, students can turn anywhere with Internet access and electricity into a classroom. It can include audio, video, text, animations, virtual training environments and live chats with professors. It is a rich learning environment with much more flexibility than a traditional classroom.

STATEMENT OF PROBLEM

In the present study, the researcher has attempted to conduct a study on the various online teaching tools and methods used by various schools and colleges for imparting education during the pandemic period. The researcher has found that this pandemic period has been a host of many online teaching tools. The study would be an appropriate one since as long as safety measures are not followed and vaccines are not effective, online teaching tools would be the medium of imparting education to the students.

SCOPE OF THE STUDY

The educational institutions are growing at an alarming rate, so there is tremendous scope of online tools for teaching. Teachers are finding these tools more effective as it eases their teaching and provides the best learning opportunities for the students. The study is carried on among the teaching fraternity of the schools and colleges in Vellore.

OBJECTIVES OF THE STUDY

1. To identify the tools used by school and college teachers during the pandemic period
2. To understand the benefits received by students through online teaching aids
3. To identify the problems faced by the faculties in implementing online teaching tools
4. To provide suggestions for improvement of online teaching tools

HYPOTHESIS OF THE STUDY

H₀: There is no significant relationship between Online teaching tools and Class of students

H₀: There is no significant relationship between Age group of students and Effectiveness on online teaching method

RESEARCH DESIGN

Primary Data Collection

Primary data was collected by primary survey method through a structured questionnaire framed through Google forms. These forms were circulated through e mail ids of the respondents.

Secondary Data Collection

Secondary data was collected from various published and unpublished sources which include reports, books, researcher papers etc.,

Sample Size

It refers to number of respondents selected from the total population taken for the study. Sample size of this study was 91.

Period of Study

The period of study is four months from December 2020 to March 2021.

Sample Area

The area of study is Vellore district. The teachers of schools and colleges such as Auxilium Hr. Sec. School, Government Girls Hr. Sec. School and Holy Angels School and Auxilium College were the respondents of the study.

Sampling Technique

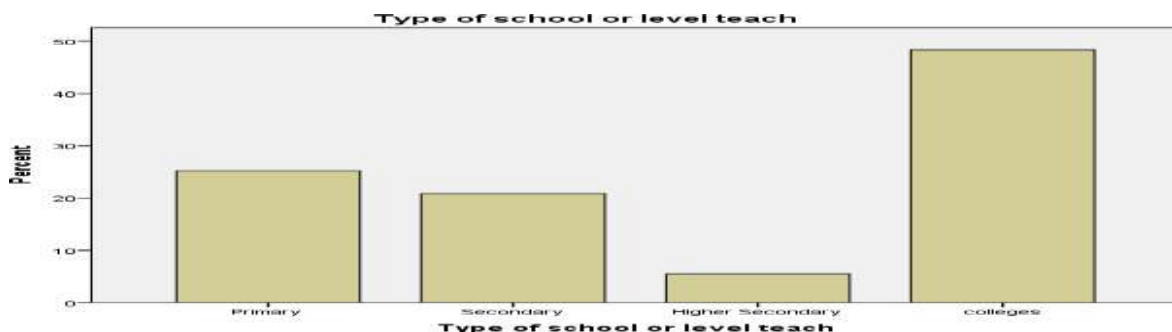
Random Sampling method was used in the study. The researcher had sent Google forms at random to various teachers to gather responses.

Statistical Tools used in the Study

Analytical techniques are used to process information in logical sequence from the data given by respondents through Percentage and Anova

ANALYSIS OF DATA

Types of School or level Teach



Source: Primary data

The above chart states that the sample unit comprises 25.3% of the teachers teaching Primary school, 20.9 % of the teachers teaching Secondary school, 5.5% of the teachers teach Higher Secondary and 48.4% of the teachers teach college students. Majority of the respondents are college teachers.

Satisfaction of the Respondents

Satisfaction	Frequency	Percent	Valid Percent
Yes	46	50.5	50.5
No	45	49.5	49.5
Total	91	100.0	100.0

Source: Primary data

The above table shows that the sample unit comprises of 50.5% of the respondents being satisfied with online teaching tools and 49.5% of the respondents are not satisfied with online teaching. The sample unit comprises of maximum respondents who are satisfied with online teaching tools.

ANOVA

Anova stands for analysis of variance. One – way analysis of variance tells us if there is any statistical difference between the means of three or more independent groups.

Association between Online Teaching Tools Used and Class of Students

H0: There is no significant relationship between Online teaching tools and Class of students

H1: There is the significant relationship between Online Teaching tools and Class of students

ANOVA Table

Apps		Sum of Squares	df	Mean Square	F	Sig.
Zoom	Between Groups	28.774	4	7.194	7.525	.000
	Within Groups	82.215	86	.956		
	Total	110.989	90			
Google Meet	Between Groups	22.617	4	5.654	8.055	.000
	Within Groups	60.372	86	.702		
	Total	82.989	90			
Byju's	Between Groups	7.575	4	1.894	1.300	.277
	Within Groups	125.326	86	1.457		
	Total	132.901	90			
Vedantu	Between Groups	9.071	4	2.268	1.560	.192
	Within Groups	125.061	86	1.454		
	Total	134.132	90			
Go to Meeting	Between Groups	7.986	4	1.997	1.370	.251
	Within Groups	125.311	86	1.457		
	Total	133.297	90			
Next learning platform	Between Groups	35.016	4	8.754	5.450	.001
	Within Groups	138.127	86	1.606		
	Total	173.143	90			

Source: Primary data

From table 4.8, it is found that Zoom ($f=7.525$, $P=0.000$), Google Meet ($f=8.055$, $P=0.000$), Next learning platform ($f=5.015$, $P=0.001$) are significant at 5% level of significance.

This indicates that there is the significant relationship between Online teaching tools used and Class of students with regard to Zoom, Google meet, Next learning platform. As value of significance is more than 0.05, there is no significant relationship between Online teaching tools used and Class of students with regarding Byju's, vedantu, Go to meet.

Association between Age group of students and Effectiveness on online teaching Methods

H0: There is no significant relationship between Age group of students and Effectiveness on online teaching methods

H1: There is the significant relationship between Age group of students and Effectiveness on online teaching methods

ANOVA TABLE

Apps		Sum of Squares	df	Mean Square	F	Sig.
Zoom	Between Groups	3.540	3	1.180	3.223	.026
	Within Groups	31.845	87	.366		
	Total	35.385	90			
Google Meet	Between Groups	2.729	3	.910	3.008	.035
	Within Groups	26.304	87	.302		
	Total	29.033	90			
Byju's	Between Groups	1.224	3	.408	.493	.688
	Within Groups	71.919	87	.827		
	Total	73.143	90			
Vedantu	Between Groups	.228	3	.076	.074	.974
	Within Groups	89.376	87	1.027		
	Total	89.604	90			
Go to meeting	Between Groups	3.907	3	1.302	1.310	.276
	Within Groups	86.510	87	.994		
	Total	90.418	90			
Next learning platform	Between Groups	2.127	3	.709	.666	.575
	Within Groups	92.598	87	1.064		
	Total	94.725	90			

Source: Primary data

From table 4.9 it is exploring that Zoom ($f=3.223$, $P=0.026$), Google Meet ($f=3.008$, $P=0.035$) are significant at 5% level of significance.

This indicates that there is the significant relationship between Age group of students and Effectiveness on online teaching methods with regard to Zoom, Google meet. As value of significant is more than 0.05, there is no significant relationship between Age group of students and Effectiveness on online teaching methods regarding Byjus, Vedantu, Go to Meeting, Next learning platform.

FINDINGS

- It's found that there is the significant relationship between Online teaching tools used and Class of students with regard to Zoom, Google meet, Next learning platform.
- It's found that there is the significant relationship between Age group of students and Effectiveness on online teaching methods with regard to Zoom, Google meet.
- It's found that there is the significant relationship between Students learning paths and Students usually perform tasks.
- It's found that there is the significance relationship between Difficulties faced in online teaching and Types of school with regard to Adaptability struggle.

LIMITATIONS OF THE STUDY

1. The study is limited to Vellore district and therefore the findings of the study cannot be extended to other areas.
2. The sample size is only 91, hence may not fully represent the characteristics of the population.
3. The study is confined to only online teaching tools used during the pandemic period with special reference to schools and college teachers.
4. Analysis is limited to sample not population.

SUGESTIONS

Many students are not well equipped with a high internet connection that is required for online learning. Due to this they face problems in going live for virtual learning and other platforms that require internet connection. They face technical issue as they are not much aware of technology and computer application. A slow and high internet connection can play an important role in how quickly you can attend the class and do not miss any live session. There is a possibility in downloading some information related to the subject, blurred videos etc., to overcome this situation just need to find a high-speed internet connection at your home. And know where you can get technical support for your connection and other technical issues related to software and tools for effective learnings students should be provided access to support devices that can help them solve technical problems via call, email or live chat. You should pay attention to your instructor during the process of solving the issues. Learning from home is an amazing experience you might expect things around you to be like a school campus.

But at home things are different for example, you might want a massive classroom, parks, playgrounds, canteens, friends, teachers around you to guide and learn. But with online learning have to manage everything in one room with parents around you, you can be easily distracted by small things at home. To overcome distraction at home, parents and friends should be informed about the time of online learning.

So that there will be no distractions from their side. Restrict the study area for others to come during live sessions and video calls. To get better learning outcomes, it is important to understand the learning styles. Learning through interaction, visual presentations, audio classes or written notes. That helps in enhancing the learning experience.

CONCLUSION

Online teaching tools like Zoom, Google meet, Byjus, Vedantu, Go to meeting, Next learning platform, Edmodo, Socrative, Project, Thing link, ck-12 etc. are the various online platforms available for teaching. Covid-19 coronavirus outbreak has put the world in lockdown. In such a situation, people have been asked to stay at home by the government to prevent the

spread of the virus. Strict limits have been enforced on gatherings, non-essential shops, businesses, parks, playgrounds, libraries, schools and colleges and other establishments have been closed in response to the accelerating pandemic. With schools forced to shut down, the gap in the delivery of knowledge is currently being bridged using “Online platforms and Tools”. The silver lining to this cloud is that teachers can use the craze for technology in Students opportunities even outside the conventional classroom setup. No wonder schools in different parts of the world are shifting to online learning with many teachers recording video classes for their students in the current world scenario. Alternatively, any teacher can also use personalized “Tools and Technologies”.

Students find it difficult to adopt to an online learning environment immediately after traditional classroom learning. Due to the sudden change, they are not able to adopt to the computer-based learning. Students who have been always studying in the traditional classroom mindset are not able to focus on online platforms. It is important for them to accept the new learning environment with an open mind. Now a days adaptive learning is using artificial intelligence to adjust the content according to individual needs. It helps in providing personalized course to identify their weaknesses and strengths for better learning outcomes.

REFERENCES

Journals

1. Amanda Athuraliya, ‘Online Teaching Tools for Educators and Students’, AASL, 4(1),2020, pp. 200-250.
2. Ashutosh Chauhan, ‘Digital Education Tools for Teachers and Learners’, EducationalTechnology, 187(4), 2018, 1172-1176.
3. Bell B.S, ‘E Learning in Postsecondary Education the Feature of Children’, InternationalJournal of Education, 23(1), 2013, 165-185.
4. Cole M.T, ‘Online Instruction, E Learning and Students Satisfaction’, EducationalInstitution, 3(2), 2014.
5. Craig J, ‘Effective Pedagogical Practices for Online Teaching’, Internet and HigherEducation, 25(1), 2009, 152-155.

Websites

www.link.springer.com
www.sciencedirect.com
www.telrp.springeropen.com
www.google scholar
www.researchgate.net

IOT BASED SOLUTION FOR MONITORING OF POLLUTION FRUIT PESTICIDES IN FRESH FRUITS AND VEGETABLES IN MARKET

Ms.T.Thenmozhi¹, Ms.S.Nishanthi²

*Department of Computer Applications, Assistant Professor
Marudhar Kesari Jain College For Women, Vaniyambadi, Tamilnadu, India.*

ABSTRACT:

This paper presents a system based on acetyl cholinesterase (AChE) biosensor and internet of things (IoT) for pesticide residues detection and agricultural products traceability. The system we presented based on AChE biosensor and IoT aims to extend the benefits of the pesticide residues detection date—remote control ability, data processing and sharing, agricultural products traceability and so on—to detection devices (purusmalus computer) in detection locations. These detection data got from detection devices were further aggregated, processed and analyzed by purusmalus computer in order to extract useful information (detection time, production places, detection samples, pesticide residues values, detection inspector) which were effective in protecting the quality and safety of agricultural products. In this study, the code of useful information was used in form of QR code, tracing and retracing the safety of agricultural products was achieved efficiently and reliably. In view of above, we design and implement a system based on AChE biosensor and IoT for pesticide residues detection and agricultural products traceability. It provides safe fruit and vegetable information for consumer, and lay foundation for the traceability of agricultural product.

Keywords: acetyl cholinesterase biosensor; pesticide residues rapid detection; QR barcode ; internet of things; agricultural products traceability.

1. INTRODUCTION:

Pesticides are chemical substances applied to crops at various stages of cultivation and during the post-harvest storage of crops. The use of pesticides is intended to prevent the destruction of food.

1.1 These authors contributed equally to this work.

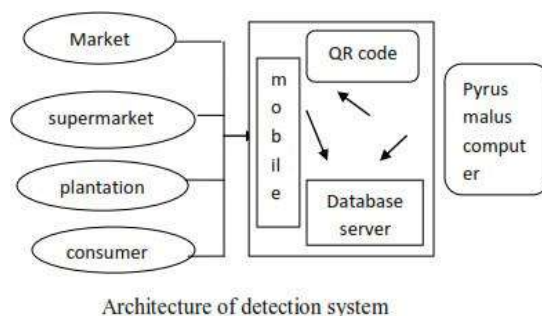
Crops by controlling agricultural pests or unwanted plants and to improve plant quality [1]. Pesticide use in commercial agriculture has led to an increase in farm productivity. Despite the wide ranging benefits of using pesticides in agriculture, several incorrect applications can result in high and undesirable levels of the compounds in the produce that reaches consumers.

These include inappropriate selection of pesticides used on foodstuffs, over use of pesticides and harvesting the crops before the residues have washed off after application [2,3]. Monitoring of pesticides in fruit and vegetable samples has increased in the last years since most countries have established maximum residue level (MRL) for pesticides in food products [4-6]. With the gradual advance of urbanization construction, the procurements of vegetables and fruits are most in markets and supermarkets. However, these procurement locations almost have no pesticide residues detection devices. Gas chromatography (GC), liquid chromatography (LC) or combinations (GC-MS or LC-MS/MS) are traditional analytical techniques for identification and quantity determination of pesticides residues [7-9]. Although these methods offer quantitative analysis with sensitivity and selectivity, they are slow, expensive, laborious and not convenient to popularize and promote. Moreover, they don't have the ability of information sharing and remote control. Therefore, they are not suitable for rapid detection and agricultural products traceability. Biosensors account for an easy method to determine pesticides [10] in environmental and food matrices [11]. The use of biosensors as screening devices is cost effective and decreases the number of samples to be analyzed by traditional analytical techniques mentioned above. With the explosive growth of smart phones, wireless technologies and sensor technologies have become a fundamental tool for everyday life around the world. The coming wave of interconnected devices, appliances, sensors, meters and countless other "things" represents the next generation of a hyper-connected world, the IoT [12]. Interconnected entities can open a communication channel with each other based on the IoT. Many technologies serve as the building blocks of this new paradigm, such as QR barcode, cloud services, machine-to-machine interfaces (M2M), and so on. Also, this paradigm has a multitude of application domains [13]. The IoT we used in this pesticide residues detection system on the one hand allows purusmalus computer to receive detection data from dispersed pesticide residues detection devices based on biosensor (purusmalus computer). The purusmalus computer turn database links into barcode-like images that can be scanned using a mobile phone or a QR code reader. On the other hand, people can obtain database link after having read the QR code and then access data in the database. With this we can establish an interconnection between those heterogeneous objects as long as they have access to the Internet [14-16]. The information sharing platform will take care of the centralization of the data of each purusmalus computer, allowing them to interact and communicate with purusmalus computer through the creation of a ubiquitous network by solving the interconnection problem [17].

without any programming knowledge. This system can be used in supermarkets, markets and plantations. Moreover, this system also can be used in the areas of purchasing, storage and transportation. The consumers and purchaser just have to scan the QR code affixed to the commodity to get the useful information at any time, any place, and in any way [18-20]. The remainder of this paper is structured as following: In section 2, there is a description to the whole system including structure and function. In section 3 we describe the structure, function and detection principle of the detection device (purusmalus computer). Furthermore, we present our research on the detection principle. Section 4 and section 5 show the study of purusmalus computer and information sharing platform, respectively. In section 6 we cover the evaluation and discussion of the data obtained from experimental and system testing. Finally, in Section 7 we give the conclusion of the whole paper and describe the performance, stability, application fields and promotion value of this system.

2. SYSTEM DESCRIPTION:

Information sharing platform Pesticide residues detection system.



Architecture of detection system

System's architecture can be divided in three parts, as it can be seen in purusmaluscomputer. Based on AChE Biosensor, purusmalus computer based on LabVIEW platform and information sharing platform. AChE was immobilized on the working electrode and reacted with the substrate to produce the weak current signal. Purus maluscomputer detection device collected weak current signal generated from AChE biosensor and transformed weak current signal into 0-5V standard voltage signal as an output signal. We could get the pesticide residues concentration based on the changed voltage signal. The detection information was transmitted to the purusmalus computer through wireless transmission.

The QR code was printed by bar code printer at the same time and people can access the detection information in the database through scanning the QR code which encoded a date link. The prototype of this detection system was shown in Fig. 2.

2.1 Purusmalus computer detection devices

The purusmalus computer detection device of this detection system was made up of electrode module based on AChE biosensor, signal detecting and processing module, printing and storage module, power supply module and data transmission module. The purusmalus computer schematic based on the single-chip microcomputer was shown in Fig. 3. The purusmalus computer realized the real-time detection to the pesticide residues of real samples and sent the detection data to the purusmalus computer. AChE was immobilized on the working electrode and reacted with the substrate to produce the weak current signal. Three-electrode module collected weak current signal generated from AChE biosensor. The detection of the weak current signal was realized by using the signal detecting and processing module. The weak current signal generated by the AChE biosensor was transformed into 0-5V standard voltage signal as an output signal by this module. Microcontroller played a critical role during the signal detecting and processing process. The hardware circuit purusmalus computer was shown in Fig. 4 and the prototype of purusmalus computer was shown in Fig. 5.

2.2 Epigynous computer

A rapid real-time purusmalus computer system with the National language LabVIEW was designed to control and manage purusmalus computer detection devices. The system is centered round a computer running application developed for the purpose [21]. The previous detection devices can't realize multi-channel acquisition and detection, leading to low detection efficiency. The establishment of the purusmalus computer on the one hand can realize the multi-channel data acquisition and detection, on the other hand, can realize the monitoring and management of the purusmalus computer detection devices. With this system, it is possible for the computer to receive data automatically and import the detection data into the database then the QR code was printed by bar code printer which encoded a date link at the same time. The functional flow diagram of purusmalus computer system was shown in Fig. 6.

The purusmalus computer system provided a convenient human- computer interaction means between the purusmalus computer detection devices and operators. The front panel and block diagram of purusmalus computer system were depicted in Fig.7a and Fig.7b respectively.

2.3 Information sharing platform

Aiming at the necessity to construct the information sharing platform of the pesticide residues system, this paper analyzed the function request of platform, integrated the technology of Access Database and ASP, at the end formed the system structure and design project of the system platform. The ASP program is executed on a network server and the generated HTML files from the execution results are applicable to the different browsers, people can access to the Access Database with ASP Webpage [22]. This information sharing platform has on the one hand provided a good platform for the co- construction and sharing of information resources, on the other hand greatly reduced the cost for information retrieving and knowledge acquiring, making it easy to look for information. People can access the detection information in the database through the information sharing platform by scanning the QR code which encoded a data link at any time, any place. The model of the proposed information sharing platform was depicted in Fig.8. Using this model, after successful registration, the remote user can reach a specific sensor node directly through the Internet and does not need to first connect with the gateway node, thereby ensuring a more straightforward approach.

EXPERIMENTAL:

3.1 Apparatus

Electrochemical measurements were performed on a CHI660D electro-chemical workstation from Shanghai Chenhua Instrument Ltd. (Shanghai, China). Three electrodes were purchased from Aida technology Co. (Tianjin, China). The working electrode was gold electrode ($d=1\text{mm}$). A saturated calomel electrode (SCE) and platinum electrode were used as reference and auxiliary electrodes, respectively. Pesticides residues detection instrument was made in our laboratory.

3.2 Reagents and materials

Acetylcholinesterase (Type C3389, 500 U/mg from electric eel), acetylthiocholine chloride (ATCl) and chlorpyrifos were purchased from Sigma (USA). SnO_2 were obtained from Sinopharm Chemical Reagent Co., Ltd. Multiwall carbon nanotubes (MWNTs) (purity>95%) was purchased from Shenzhen Nanotech Port Company (China) and chitosan (CHIT) was from Shanghai Chemical Reagent Company (China). The 0.1 M pH 7.5 phosphate buffer solutions (PBS) were prepared by mixing the stock solutions of NaH_2PO_4 and Na_2HPO_4 .

3.3 Preparation of Nafion/AChE/MWNTs-SnO₂- CHIT/Au Biosensor

The Au electrode surface was freshly polished with 0.3 μm and 0.05 μm alumina powder, respectively, and then rinsed with ultrapure water after each polishing, finally cleaned ultrasonically with 95% ethanol and acetone for 3 min, respectively. The SnO₂ nanoparticles and MWNTs with a mass ratio of 1:3 were dispersed in 0.2% CHIT solution and stirred at room temperature for 3h. The obtained highly dispersed black suspension would be named as MWNTs-SnO₂-CHIT. A 2.5 μm of MWNTs- SnO₂-CHIT suspension was coated on the Au electrode surface and air dried naturally to obtain MWNTs-SnO₂-CHIT/Au. Similarly, SnO₂- CHIT/Au and MWNTs-CHIT/Au were prepared under the same procedure as illustrated in MWNTs- SnO₂-CHIT/Au preparation just without MWNTs or SnO₂ existing, respectively. The obtained electrode (MWNTs-SnO₂-CHIT/Au) was washed thoroughly with ultrapure water and then dried in air at room temperature. After the water was evaporated, MWNTs-SnO₂-CHIT/Au was coated with 2.5 μL AChE solution to obtain the AChE/MWNTs-SnO₂- CHIT/Au. Finally, the AChE/ MWNTs-SnO₂- CHIT/Au electrode was coated with an extra 2.5 μL 0.5% Nafion to maintain the stability of modified electrode.

RESULTS AND DISCUSSION:

4.1 Electrochemical detection of pesticides

For the measurement of pesticides, the obtained AChE/MWNTs-SnO₂-CHIT/Au was first immersed in pH 7.5 PBS containing different concentrations of standard pesticides solution for 10 min, and then transferred to the electrochemical cell of pH 7.5PBS containing 2 mMATCl to study the electrochemical response by cyclic voltammetry (CV) between 0.7 and 0.2V as shown in Fig. 9. The highest oxidation current of thiocholine could be obtained when the working potential was 500 mV between working electrode and reference electrode after the bare gold electrode was modified by the AChE/MWNTs-SnO₂-CHIT. The inhibition rate of pesticides was calculated as follows: $\Delta I (\%) = (I_0 - I_1) / I_0 \times 100\%$ or $\Delta U (\%) = (U_0 - U_1) / U_0 \times 100\%$ (3) $\Delta I (\%)$ is the inhibition rate, I_0 is the oxidation current obtained in the absence of pesticides and I_1 is oxidation current obtained after biosensor exposure to pesticides. The oxidation current signal generated by the AChE biosensor was transformed into 0-5V standard voltage signal as an output signal by current-to-voltage conversion circuit of purusmaluscomputer detection device based on the principle of $U=I \cdot R$. Microcontroller played a critical role during the signal detecting and processing process.

4.2 The establishment of inhibition ratio curve

This purusmalus computer detection device of the pesticide residues detection system has been used to test standard pesticides and real samples to compare with electrochemical analysis method. The linear regression equations for the modified electrode with using different detection methods were $\Delta I\% = 0.10667C (\mu\text{g/L}) + 45.05473$ (Fig. 11a) and $\Delta I\% = 0.10528C (\mu\text{g/L}) + 40.57149$ (Fig. 11b) in the range 2 to 500 $\mu\text{g/L}$, the limit of detection was 2 $\mu\text{g/L}$. Clearly, the correlation between electrochemical analysis method and the mentioned instrument was observed to be linear with a similar slope detect for the biosensors with Nafion/AChE/MWNTs- SnO₂-CHIT/Au composite film as shown in Fig. 9. This means the standard pesticides could be evaluated by the pesticides residues detection instrument and the results indicated that the purusmalus computer detection device was suitable for direct analysis of pesticides.

4.3 Detection of real samples

The precision of the detection device was evaluated by analyzing eight kinds of vegetables and fruits samples. The real samples test was studied by using fresh cucumber, long bean, spinach, cabbage, kidney bean, apple and tomato, lettuce. Table 1 was showed the actual detection results of different kinds of vegetables and fruits by using mentioned purusmalus computer detection device of the pesticide residues detection system, the pesticide residues concentration of real samples were all lower than the detection limit, so pesticide residue concentration displayed as a vacancy concentration. The detection results show that the pesticide residues did not exceed the standard. The detection results show that the pesticide residue did not exceed the standard. Gas chromatography (GC), liquid chromatography (LC) or combinations (GC- MS or LCMS/MS) are traditional analytical techniques for identification and quantity determination of pesticides residues [23-25]. Although these methods offer quantitative analysis with sensitivity and selectivity, they are slow, expensive, and laborious. Therefore, they are not suitable for rapid detection and field application.

Alain Hildebrandt et al. have designed and developed a portable biosensor for the analysis of organophosphorus (OP) and carbamate insecticides in water and food [26]. Gilmo Yang et al. have developed an opto-fluidic ring resonator biosensor for the detection of organophosphorus pesticides [27]. Vangelis G. Andreou et al. have developed a portable fiber-optic pesticide biosensor based on immobilized cholinesterase and sol-gel entrapped bromocresol purple for in-field use [28]. All these methods are highly competitive with traditional analytical techniques in terms of shorter time response and lower cost, but they are not enough instrumented, and on the other hand,

rather complex procedures make them unsuitable for industrial or commercial applications. The pesticide residues detection system mentioned in this paper integrated the function of pesticides residues rapid detection, detection devices supervising and detection information sharing based on the AChE biosensor and IoT. The detection performance of the purusmalus computer detection device showed the capability of the pesticides residues detection with good sensitivity and high practical value. The limit of detection was $2\mu\text{g/L}$ and the accuracy of measurement could meet rapid detection of pesticides residues requirements.

4.4 Performance evaluation

The working process of the mentioned pesticide residues detection system can be divided into 4 parts: purusmalus computer detection device detect the pesticide residues and get the concentration of the pesticide residues. purusmalus computer system received detection information from purusmalus computer based on the wireless data transceiver module and the detection information would be stored in the database at the same time. purusmalus computer integrated detection information, output QR code and supervise the purusmalus computer detection devices. Customers got the detection information by scanning the QR code which encoded a data link. In order to evaluate the stability and reliability of the proposed pesticide residues detection system during the working process, interference factors were added into the detection and transfer process to investigate their effect on the detection results. The delivery result of original pesticide residues detection data in each part was acquired from eight kinds of vegetables and fruits samples, as shown in Table 2. The equal pesticide residues concentrations show that the system was stable and reliable. The pesticide residues detection data which users obtained is accurate and original. The overall performance of the present system showed the capabilities of the pesticide residues rapid detection, detection information wireless transmission, detection information management, detection devices remote dynamic supervision and detection information sharing, which lay the foundation for the traceability of agricultural products.

CONCLUSION:

In this paper, a system consisting of purusmalus computer detection devices, purusmalus computer and information sharing platform was developed for pesticide residues detection and agricultural products traceability. The purusmalus computer in this system has rapid detection and wireless transmission function and can be used on-side. In order to achieve purusmalus computer management and multi-channel detection, purusmalus computer has been proposed and carefully designed for both the dynamic supervision and QR code printouts. The application of this

pesticides residues detection instrument has been performed on real samples. The system showed to be successful in pesticide residues detection and agricultural products traceability. For chlorpyrifos extracts, the detection system based on biosensor permitted to determine concentrations of 2µg/L, thus indicating the performance of this system can satisfy the pesticide residues detection and information sharing requirement of real vegetables and fruits samples. The detection system based on AChE biosensor and IoT for pesticide residues detection can be used in every link in the agricultural products traceability.

REFERENCES:

1. G. T. Bakirci, Y. Hisil, Food Chem, 135(2011) 1901
2. S. V. Boxstael , I. Habib , L. Jacxsens , M. D. Vocht , L. Baert , E. V. D. Perre , A. Rajkovic ,
- 3.F. Lopez-Galvez , I. Samplers and P. Spanoghe , *Food Control*, 2013, **32** , 190 -197
- D.A. SOUZA, Development of analytical methods for multi-residue determination of pesticides in drinking water of São Carlos – SP (Desenvolvimento de metodologia analítica para determinação de múltiplos resíduos de pesticidas em água de abastecimento de São Carlos – SP), Doctoral thesis, São Paulo University,
4. G. O. Guler, Y. S. Cakmak, Z. Dagli, Food Chem Toxicol, 48(2010)1218
- 5.A. R. Boobis, B. C. Osendorp, U. Banasiak, Toxicology letters, 180(2008)137
- 6.Agüera, M. Contreras, A. R. Fernandez-Alba, J. Chromatogr A, 655(1993)293
- 7.J. Ye, J. Wu, W. Liu, Trac-Trend Anal Chem, 28(2009) 1148
- 8.H. Guan, W. E. Brewer, S. T. Garriss, J. Chromatogr A, 1217(2010)1867
- 9.M. P. Marco, D. Barceló, Meas Sci Technol, 7(1996)1547
- 10.Nedelkoska, T. V.; Low, G. K. C.; Anal. Chim. Acta 2004, 511, 403
- 11.D. Shin, J. Telematics and Inform, 31(2014) 519
- 12.R. Roman, J. Zhou, J. Lopez, J. Comput Netw, 57(2013)2266
- 13.G. C. González, G. Pelayo, B. C. Bustelo, J. Comput Netw, 64(2014)143
- 14.L. Atzori, A. Iera, G. Morabito, J. Comput Netw, 54(2010)2787
- 15.L. Tarjan, I. Šenk, S. Tegeltija, J. Comput Electron Agr, 109(2014)1
- 16.L. Zhao, J. Qian, L. Chang, J. Data Knowl Eng, 69(2010)737
- 17.D. Lorenzi, J. Vaidya, S. Chun, J. Government Information Quarterly, 31(2014)6
- 18.L. Bai, F. Wang, M. Zhang, J. Procedia Soc Behav Sci, 138(2014) 350
- 19.Y. Qin, Q. Gong, J. Opt Commun, 310 (2014)69

CRYPTOGRAPHY FROM OLDER TO MODERN ERA

*K. Durgadevi¹, V. Harisha², K.R. Soundarya³
Marudhar Kesari jain college for women, vaniyambadi, TN, India*

ABSTRACT:

Our world occupied with full of data's whether it can be single person or family or village or town or city or state or country database which is exchanged between others. In older days, each and every single data present in written form only and that data is important but not secured, so the data exchange is done only between trusted persons which is hard and time consuming. But, in this modern era which is internet era, each and every individual person details are in soft document form which could be saved anywhere like mobile, laptop, cloud and PC etc., through internet. Even though data making and data exchange between other person is easy and time saving, there is a big risk, that his data will expose to anyone who uses internet.

INTRODUCTION:

Cryptography was the secret writing and it was the oldest art, and it was a first used and documented around 1900 B.C. Egyptian transcriber used only non-standard symbols in writing. There is some expert's claim that cryptography appeared spontaneously sometime after writing was created, with applications ranging from diplomatic letters to war-time battle plans. There will not be any surprise, then, that new systems of cryptography came into effect after the extensive development of computer communications.

Cryptography is a technique of protecting data information and data communications through use of codes, so that only the person for whom the information is proposed can read and process the data. The preface "crypt" means to "hidden" and the postfix "graphy" stances "writing" refers to hidden writing.

The five main functions of cryptography are as follows:

Privacy/confidentiality:

To Ensure that no one can read the sent message except the proposed receiver.

Authentication:

The method used to verify single's identity.

1.Integrity:

Assurance was given to the receiver that he received the message without any altered form the original.

2.Non-repudiation:

Evidence that the sender really sent this message.

3.Key exchange:

Technique by which the crypto keys are shared between the sender and the receiver.

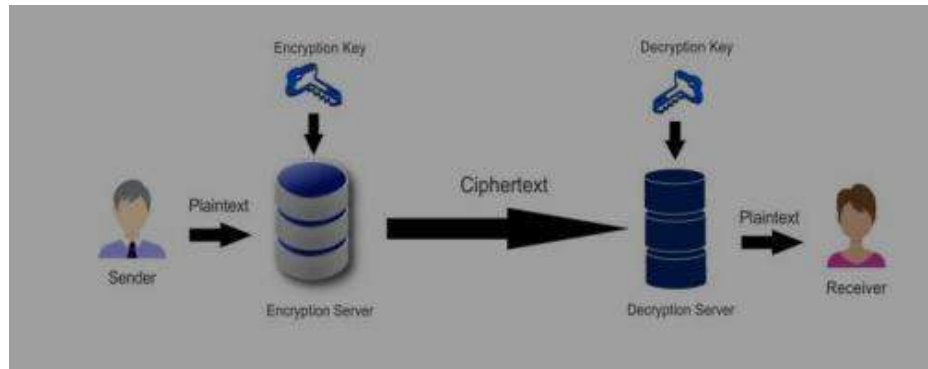


Figure: 1 – Cryptography Working

Cryptography starts with unencrypted data, stated to plaintext. Plaintext is encrypted into ciphertext, which will in usually decrypted into practical plaintext. The encryption and decryption are based upon the type of cryptography scheme being active and some method of key. We can write the same in formulas, this process is written as example:

$$C = E_k(P)$$

$$P = D_k(C)$$

Where, **P** = Plaintext, **C** = Cipher Text, **E** = Encryption method, **D** = Decryption method, and **k** = Key.

OLDEN DAYS CRYPTOGRAPHY:

Ciphers (secret writing), codes and other encryption methods have been used throughout the history by most civilized peoples in some form to prevent non-approved people from understanding communications sent.

CAESER SHIFT CIPHER – ROMAN ARMY:

One of the oldest, earliest and simplest cryptography method created by Julius Caesar and it was used by Rome Army.

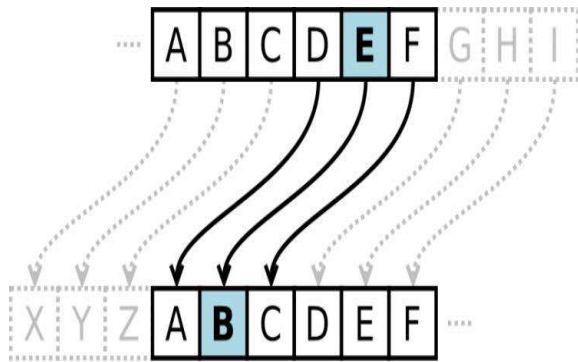


Figure: 2 – Caesar Shift Cipher

SCYTALE CIPHER – SPARTANS:

The Scytale cipher is a transposition cipher used by the Spartans and the ancient Greeks to send battle attack messages. Scytale Cipher encryption method uses rods of wood with equal radius, once the message created, the rod winds a strip of leather around it. Message to be written across the rod, so that anonymous person unwinds the leather stripe the letters show jumbled in a meaningless fashion.



To decrypt, all one must do is **wrap the leather strip around the rod and read across**. The ciphertext is: "Iryatbhmvaehdlurip". Every fifth letter will appear on the same line, so the plaintext (after re-insertion of spaces) becomes: "I am hurt very badly help".

Figure: 3 –Scytale Cipher

STEGANOGRAPHY – ANCIENT GREECE:

Steganography is also an ancient method of hiding messages sent from battle field and also from kingdom 'in plain sight' by masking it as something else. This method is done by shaving the head from kingdom's most trusted servant, "marking" the message onto his scalp, and then allowing the servant's hair to regrow again until the messages hides. Then the servant was sent to the receiver and ordered him to shave his head once again to read the hidden message.

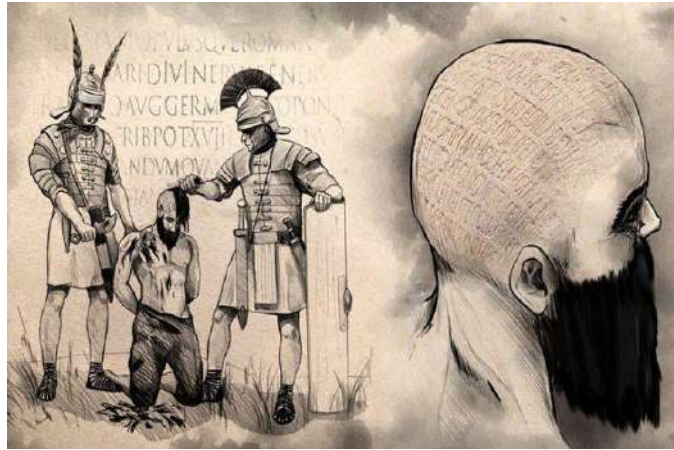


Figure: 4 – Steganography

PIGPEN CIPHER – FREEMASONS:

Pigpen Cipher was a simple geometric swap cipher. This method of cipher uses symbols to encode the letters within a message which was sent to the concern person. It is encrypted and decrypted by creating a grid or set of grids to produce the following symbols.

a	b	c	d	e	f	g	h	i	j
└	┐	┌	┘	□	├	└	┐	┌	┘
k	l	m	n	o	p	q	r	s	t
┐	┌	┘	└	├	└	┐	┌	┘	└
u	v	w	x	y	z				
<	^	v	>	<	^				

Figure: 5 - Pigpen Cipher

PLAYFAIR CIPHER – UNITED KINGDOM

Playfair Cipher method encrypts pairs of letters, instead of single letters as in the unassumingswitch cipher. Playfair Cipher is significantly much harder to break since the frequency analysis used for simple substitution ciphers does not work with it.

• For example,
Keyword: PLAYFAIR
Plaintext: OC

P	L	A	Y	F
I	R	B	C	D
E	G	H	K	M
N	O	Q	S	T
U	V	W	X	Z

7:54Z

Figure: 6 -Playfair Cipher

TYPES OF CRYPTOGRAPHY ALGORITHM:

The types of cryptography algorithm were categorized based on the number of keys that involved in encrypting the data and decrypting the data and it is divided into three types.

Symmetric Encryption Method.

Asymmetric Encryption Method.

Hash Functions.

SYMMETRIC ENCRYPTION METHOD:

Symmetric Encryption algorithm in cryptography that uses same key to encrypt plain text and decrypt cipher text. The key is shared between sender and receiver to encrypt and decrypt the message. Using same key for encryption and decryption was the main drawback in this method.

Symmetric Encryption

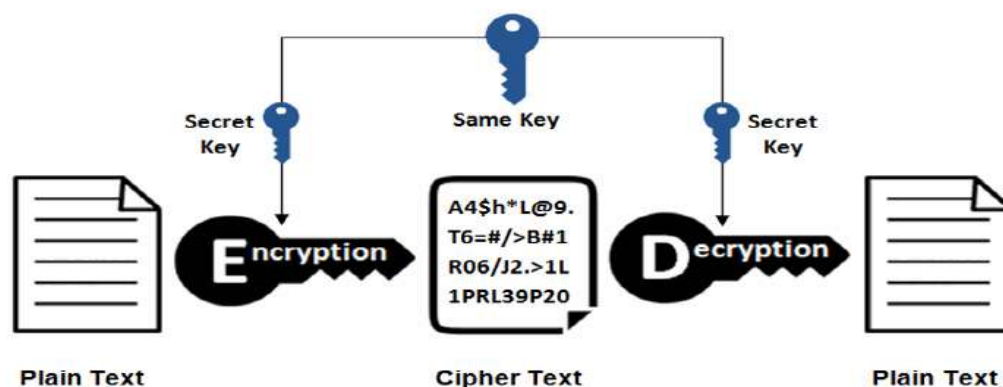


Figure: 7 - Symmetric Encryption Method

ASYMMETRIC ENCRYPTION METHOD:

The asymmetric encryption method uses multiple keys to encrypt and decrypt data. The asymmetric encryption method contains two encryption keys that are scientifically related to each other, those keys are known as the public and private key. Without private key, a person cannot decrypt the data sent from source, if sender shares his public key to anyone.

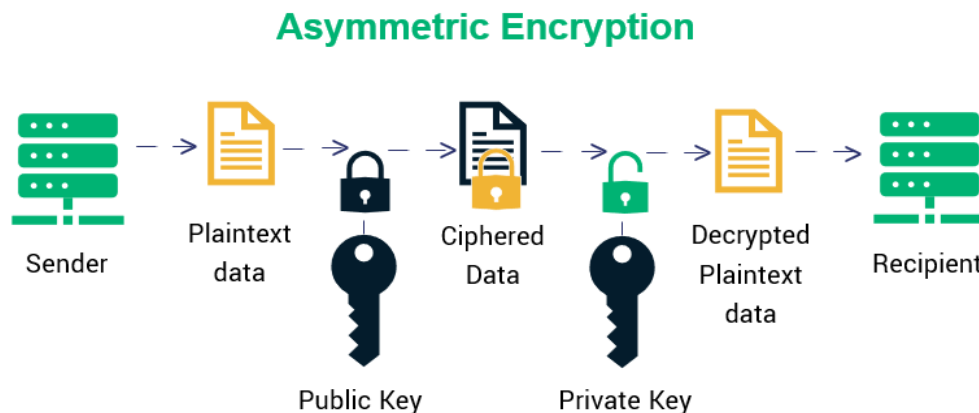


Figure: 8 - Asymmetric Encryption Method

HASH FUNCTIONS:

Hash function is a mathematical function which converts any numerical input value into another compressed numerical value. The input given to the hash function is of arbitrary length but output is always of fixed length. This method is mostly used in sending software updates over the cloud.



Figure: 9– Hash Function

CONCLUSION:

Cryptography was the main role in securing data from hackers, this cryptography concept continuous to modern era from older era. There are many cryptography algorithms, which has easy algorithm to decode and some are very difficult algorithm to decode. We can realize that the data security is still in a big question mark, because someone develops some new algorithm or modify old algorithm, always there is a group to decode the algorithm to show this algorithm is failure and it continuous from olden day era to modern day era. We can see history that each and every algorithm was cracked by some people at some point of time. We can also see that data security is improving day by day, because someone creates new or modified algorithm before someone crack the older algorithm. Each and every cryptography algorithm has some advantages and some disadvantages based on the purpose of security involves.

REFERENCES:

1. <https://www.garykessler.net/library/crypto.html#fig01>
2. <https://interestingengineering.com/11-cryptographic-methods-that-marked-history-from-the-caesar-cipher-to-enigma-code-and-beyond>
3. <https://en.wikibooks.org/wiki/Cryptography/Scytale>
4. <https://medium.com/@.Qubit/symmetric-key-algorithm-in-cryptography-3d839bba8613>
5. <https://sectigostore.com/blog/types-of-encryption-what-to-know-about-symmetric-vs-asymmetric-encryption/>

Brain Intelligence model: Fusing Artificial Intelligence and Artificial Life

Dr.Vignesh Ramamoorthy H¹

*Assistant Professor, Department of Information Technology and Cognitive Systems
Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India*

Ms.Priyadharshini N²

*Assistant Professor, Department of Computer Science
Sri GVG Visalakshi College for Women, Udumalpet, Tamil Nadu, India*

Dr.LovelineZeema J³

*Assistant Professor, Department of Computer Science
Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamil Nadu, India*

Abstract

Artificial intelligence (AI) is an important technology that supports daily social life and economic activities. It contributes greatly to the sustainable growth of Japan's economy and solves various social problems. In recent years, AI has attracted attention as a key for growth in developed countries such as Europe and the United States and developing countries such as China and India. The attention has been focused mainly on developing new artificial intelligence information communication technology (ICT) and robot technology (RT). Although recently developed AI technology certainly excels in extracting certain patterns, there are many limitations. Most ICT models are overly dependent on big data, lack a self-idea function, and are complicated. In this paper, rather than merely developing next-generation artificial intelligence technology, we aim to develop a new concept of general-purpose intelligence cognition technology called Beyond AI. Specifically, we plan to develop an intelligent learning model called Brain Intelligence (BI)[^] that generates new ideas about events without having experienced them by using artificial life with an imagine function. We will also conduct demonstrations of the developed BI intelligence learning model on automatic driving, precision medical care, and industrial robots.

Keywords:

Brain intelligence, Artificial intelligence, Artificial life

1. Introduction

From SIRI [1] to AlphaGo [2], artificial intelligence (AI) is developing rapidly. While science fiction often portrays AI as robots with human-like characteristics, AI can encompass anything from e-Commerce prediction algorithms to IBM's Watson machines [3]. However, artificial intelligence today is properly known as weak AI, which is designed to perform a special task (e.g., only facial recognition or only internet searches or only driving a car). While weak AI may outperform humans at a specific task, such as playing chess or solving equations, general AI would outperform humans at nearly every cognitive task.

In recent years, the US government has supported basic research on AI, which is centered on robots and pattern recognition (voice, images, etc.). Microsoft has announced real-time translation robots and innovative image recognition technologies [4]. Amazon uses artificial intelligence for autonomous robots in delivery systems [5]. Facebook has also developed facial recognition technology based on artificial intelligence called DeepFace[^] [6]. Robots and artificial intelligence are being actively studied in university institutions in the United States. Innovative technologies, such as corporate cooperation and deep learning, are emerging. The robot car developed by the Artificial Intelligence Laboratory at Stanford University has set a faster time than an active racer [7]. The Computer Science and Artificial Intelligence Laboratory at Massachusetts Institute of Technology has developed a cleaning robot and a four-foot walking robot [8].

a world-class Btechnological superpower[^], the Japanese government has formulated projects such as the BScience and Technology Basic Plan [9][^] and robots to solve various problems. However, through some research, we found that recent ar- tificial intelligence technologies have many limitations. In the following, we list some representative limitations and analyze the reasons why recent AI cannot break through these inherent disadvantages.

1.1 Limitations of artificial intelligence

In recent years, artificial intelligence technologies have devel- oped dramatically due to improvement in the processing ca- pacity of computers and the accumulation of big data. However, the results of current artificial intelligence technol- ogies remain limited to specific intellectual areas, such as im- age recognition, speech recognition, and dialogue response. That is, current AI is a specialized type of artificial intelligence acting intellectually in a so-called individual area (see Fig. 1). Examples include techniques such as Convolutional Neural Networks (CNN) or Deep Residual Learning (ResNet) for visual recognition, Recurrent Neural Networks (RNN) or Deep Neural Networks (DNN) for speech recognition, and Represent Learning (RL) for dialogue understanding. All of these are a part of the intellectual work carried out by each area of the human brain; they are only a substitute and do not perform all of the functions of the human brain. In other words, AI has not been able to cooperate with whole-brain functions such as self-understanding, self-control, self- consciousness and self-motivation. Specifically, we conclude that the limitations of the recent artificial intelligence technol- ogies are the following:

(1) Frame Problem

Considering all the events that can occur in the real world, since it takes a large amount of time due to big data training, AI is typically limited to a single frame or type of problem. For example, if you restrict the algorithm to apply only to chess, shogi, image recognition, or speech recognition, only certain results can be expected. However, when trying to cope with every phenomenon in the real world, there is an infinite number of possibilities that we have to anticipate, so the extraction time becomes infinite due to overloading of the database.

(2) Association Function Problem

Machine learning and artificial intelligence are excellent at extracting a particular pattern. However, the results of ma- chine learning are easy to misuse. Current artificial intelli- gence technology depends on large-scale data and can obtain results using only numerical values, but it does not have the association function like the human brain. That is, a single part of the brain cannot be as intelligent as the whole brain.

(3) Symbol Grounding Problem

It is necessary to link symbols with their meanings, but this task is often not resolved in current artificial intelligence. For example, if you know the individual meaning of the word Bhorse[^] and the meaning of the word Bstripes[^], then when you are taught that Bzebra = horse + stripes[^], you can under- stand that Ba zebra is a horse with stripes[^]. However, the computer cannot make the same connections between ideas.

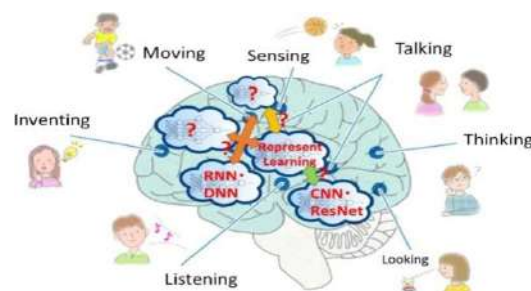


Fig. 1 Shortages of current artificial intelligence

(4)Mental and Physical Problem

What is the relationship between the mind and body? That is, if the mind is generally thought of as non-material, how can the physical body be affected by it? Whether or not this is possible has not been elucidated. In conclusion, we can see that there are many problems unsolved in current artificial intelligence. In this paper, we first review the most recent algorithms for weak AI. Then, we introduce the next-generation intelligence architecture, Brain Intelligence, which is an advanced artificial intelligence for solving the disadvantages of weak AI algorithms.

2Artificial Intelligence

The market and business for AI technologies is changing rapidly. In addition to speculation and increased media attention, many start-up companies and Internet giants are racing to acquire AI technologies in business investment. Narrative Science Survey found that 38% of enterprises have been using AI in 2016, and the number will increase to 62% in 2018. Forrester Research expects AI investment in 2017 to grow by more than 300% compared with 2016. IDC estimates that the AI market will grow from \$ 8 billion in 2016 to \$ 47 billion in 2020 [11]. Current artificial intelligence includes a variety of technologies and tools, some time-tested and others that are relatively new. To help understand what is hot and what is not, Forrester has just released a TechRadar report on artificial intelligence (application developers), detailing the 9 technologies for which companies should consider using artificial intelligence to support decisions.

- Natural language generation
- Speech recognition
- Virtual/augmented reality
- AI-optimized hardware
- Decision management
- Deep learning platforms
- Robotic process automation
- Text analytics and NLP
- Visual recognition

3Brain intelligence (BI)

There are many approaches [58–61] proposed to solve the limitations of recent AI. However, these models are simply extended from the current AI models. This paper introduces the following items for explaining the concept of BI, which is different from artificial intelligence, but extends upon current artificial intelligence. The BI intelligent learning model fuses the benefits of artificial life (AL) and AI. Currently, the mainstream research on deep learning is a method of learning expressions extracted from essential information of observational data by a deep neural network with a large number of layers. However, research on multitask learning that learns multiple tasks at the same time and transition studies that divert learning results for a certain task to other tasks is still insufficient. For this reason, AI models based on unsupervised learning and shallow neural networks will become trends in future. In this paper, we will combine various regional AI methods using a particular rule, especially unsupervised learning methods. It is essential to develop a new intelligent learning model with a small database and the ability to understand concepts. Therefore, we propose a Brain Intelligence model with memory and idea function in Fig. 3. The BI model network combines artificial life technology and artificial intelligence technology with memory function.

Research on current AI mainly focuses on individual areas such as dialogue comprehension, visual recognition, and auditory discrimination and so on. Research on whole-brain functions is still insufficient. For example, there are few studies on perceptual understanding models and self-thinking models.

Therefore, in this research, we will clarify the function and mechanisms of the whole brain and make efforts to realize it as artificial intelligence. BI network is consisted by many simple sub-networks. The parameters of each sub-networks is updated by S-system [62], which can modify the sub-networks by reproduction, selection, and mutation.

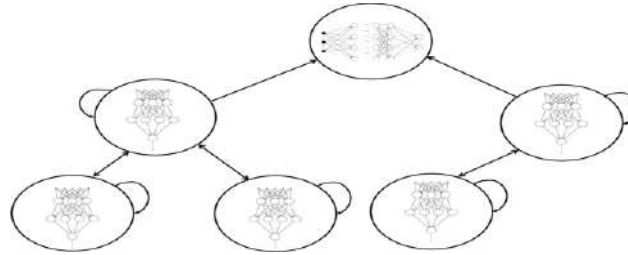


Fig.2 The concept of the BI model network. Different neural networks are connected by artificial life-based network, which can share the parameters, trained results, and structures to parents and sons

Different from NeuroEvolution of Augmenting Topologies (NEAT) [63, 64], the proposed BI model network does not just use the neural network structure and parameter optimization mechanism, it improves the structure of current AI models using S-system. hyperNEAT [65], a type of A-life based NN, which uses the Compositional Pattern Producing Network (CPPN) for pattern generation and uses NEAT for parameters optimization. hyperNEAT cannot overcome the drawbacks of the NEAT network. Other gene-based models, such as Gene Regulatory Network (GRN) [66] and Evolving Reaction Network (ERN) [67], are also studied by some researchers. These methods are inspired by biological characteristics, which do not take into account the usage of all the brain's function. Cognitive Computing (CC) [68] is proposed a new model from the view of human cognitive functions. The BI model network is investigated from an engineering point of view, in the future, we will develop a super-intelligent brain function model that intends to discover problems itself and autonomously enhance its abilities.

4. Conclusion

In this paper, we have presented state-of-the-art artificial intelligence tools for individual application areas, such as natural language processing and visual recognition. The main contributions of this work are as follows. First, this is an overview of current deep learning methods. We have summarized the nine potential applications in detail. Second, this paper puts together all the problems of recent AI models, which will direct future work for researchers. Third, in this paper, we first proposed the brain intelligence model, which is a model fusing artificial intelligence and artificial life. AL models, such as the S-system, have the benefits of an association function, which is different from generative adversarial networks (GAN), for building big data within a life evolution process. It is foreseeable that the BI model can solve the issues of the frame problem, the association function problem, the symbol grounding problem, and the mental/physical problem.

References

1. Siri, <https://en.wikipedia.org/wiki/Siri>. Accessed 20 April 2017
2. AlphaGo, <https://deepmind.com/research/alphago/>. Accessed 20 April 2017
3. IBM Watson, <https://www.ibm.com/watson/>. Accessed 20 April 2017
4. Microsoft Translator Speech API, <https://www.microsoft.com/en-us/translator/speech.aspx>. Accessed 20 April 2017
5. Amazon Prime Air, <https://www.amazon.com/Amazon-Prime-Air/b?node=8037720011>. Accessed 20 April 2017
6. Taigman Y, Yang M, Ranzato M, Wolf L (2014) BDeepFace: Closing the Gap to Human-Level Performance in Face Verification, *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR2014)*, pp.1–8

7. StanfordArtificialIntelligenceLaboratory,<http://ai.stanford.edu/>. Accessed 20 April2017
8. MIT BigDog, <https://slice.mit.edu/big-dog/>. Accessed 20 April 2017
9. The4thScienceandTechnologyBasicPlanofJapan,<http://www8.cao.go.jp/cstp/english/basic/>. Accessed 20 April2017
10. AIEXPO,<http://www.ai-expo.jp/en/>.Accessed20April2017
11. 2017 Will Be the Year of AI, <http://fortune.com/2016/12/30/the-year-of-artificial-intelligence/>. Accessed 20 April 2017
12. LeCunY, BengioY, Hinton G (2015) Deep learning. Nature 521(7553):436–444
13. BengioY,DucharmeR,VincentP,JanvinC(2003)Aneuralprobabilisticlanguage model.JMachLearnRes3:1137–1155
14. MnihA,HintonG(2007)BThreenewgraphicalmodelsforstatistical language modelling,^InProcofICML07,pp.641–648
15. T. Mikolov, M. Karafiat, L. Burget, J. Cernocky, S. Khudanpur, (2010) BRecurrent neural network based language model,^ In Proc of Interspeech 10, pp.1045–1048
16. SutskeverI,VinyalsO,LeQ(2014)BSequencetosequencelearning with neural networks,^ In Advances in Neural Information Processing Systems, pp.3104–3112
17. MeiH,BansalM,WalterM(2016)BWhattotalkaboutandhow? SelectivegenerationusingLSTMswithcoarse-to-finealignment,^ In NAACL-HLT, pp.1–11
18. LuongM,LeQ,SutskeverI,VinyalsO,KaiserL(2016)BMultitasksequencetosequencelearning,^InProcICLR,pp.1–10
19. BourlardH,MorganM(1994)Connnectionistspeechrecognition: ahybridapproach.KluwerAcademicPublishers,Netherlands
20. GravesA,MohamedA,HintonG(2013)BSpeechrecognitionwith deeprecurrentneuralnetworks,^InICASSP2013,pp.1–5

The Influence of Technology on English Language and Literature

*Dr.S.Varalakshmi, Vice-Principal,
Global College of Arts and Science for Women, Veppur,TN,India*

Abstract:

In today's world the importance of Technology is growing tremendously in every aspect, each and every country is harnessing the utilities of Technology in order to improve the status of their own. India particularly focuses on 'Digital India' in every aspect and especially in the fields of Education, Technology, Science etc., We could see the development and utilization to improve the standards and methodology of teaching, learning pedagogy, especially in the teaching of different subjects. The influence of Information and Communication Technology made teaching process flexible and innovative for the teachers as well as the staff members. English literature which talks about the life in various genres like drama, poetry, prose, etc., could be thought innovatively and creatively using different techniques of technology. The emotions, actions, dialogues connected with these genres as well could be expressed or understood explicitly with the aid of technology. My paper will discuss in detail about how technology fruitfully could be incorporated into the learning of literature and language and how new genres could be made because of it. The paper will further discuss about incorporating these elements into literature with the help of technology into the different genres like drama, prose and poetry....It will also focus on how ICT could make the teaching learning process of English Literature more interesting and more lively, especially to the teachers and students who learn English Language and Literature as a foreign language especially in countries like India.

Keywords: Language and Literature teaching and learning by incorporating Information and Communication Technology...

Introduction

Information and Communication Technology is decisive to appendage the conventional mode of education. This does not mean that technology could replace technology and no technology could address the emotional aspect of human beings. In the era of technological bang, it gets bit boring of teaching literature in conventional way, where the teacher alone speaks or teaches looking at the text for many classes. After certain tenure there may be a chance of forgetting what happened in the classroom, instead if the classrooms are aided with technology students could get rid of the monotonous way of listening and noting down notes from the teachers. For example: When I studied Nagamandala and Hayavadana in PG my teacher used to explain it with notes and wonderful description of events jotted down by Girish Karnad, but if it would have been played using technology I would not have lost the musical note behind the drama.....

Hence the technology incorporated Learning and Teaching may make the classroom livelier since it could make the classroom more interesting towards what was taught in the classroom. We will see how technology could be incorporated in the innovative teaching in the following lines....

Information and Communication Technology

ICT is an acronym of Information and Communication Technologies. The set of technologies developed to manage information and send it from one place to another. They cover a wide range of solutions. They include technologies to store information and retrieve it later.-Google.

Information is the raw material for any session and that such information is wealth we could get in fraction of seconds from the internet with the tremendous development of Communication and Technology.

The English word "Information" apparently derives from the Latin stem (information-) of the nominative (information): this noun derives from the

verb *īnfōrmāre* (to inform) in the sense of "to give form to the mind", "to discipline", "instruct", "teach" –Google & Oxford dictionaries.

Communication

Communication (from Latin *communicare*, meaning "to share" or "to be in relation with") is "an apparent answer to the painful divisions between self and other, private and public, and inner thought and outer word"

Communication is "the imparting or exchanging of information by speaking, writing, or using some other medium." It's also called "the successful conveying or sharing of ideas and feelings."-Oxford Dictionary.

Technology

Technology is the set of knowledge, skills, experience and techniques through which humans change, transform and use our environment in order to create tools, machines, products and services that meet our needs and desires. The word Technology has been derived from the Greek Word "Technologia" which means 'systematic treatment'.

Nowadays we have got many apps for teaching learning process as Byju's, Whitehat junior, Akash etc., all of these technology incorporated learning has made learning process more interesting and effective. In this context we could speak the connectivity of Television, video, Computers, tabs and satellite systems among others-a range of technologies.

English Literature

English Literature is which has its own life span of more than hundreds of years continues to be one of the most popular courses of study in colleges and Universities all over the world. It deals with the study of Literature in English Language especially in India it is called Indian Literature and the group of Asian Continent literature we call Commonwealth Literature. Writers from all over the different parts of the world cover literature with various styles

and genres. Especially the genres of English Literature such as Prose, Poetry, Fiction, drama and other forms have various writers with variety of styles and forms written all over the world.

Teaching of English Literature with Information and Communication Technology: An innovative Approach

English Literature and its nuances of elements like words, expressions; emotions, timing, moods and dialogues etc., have to be well expressed in order to make the students understand explicitly. Teachers of English Literature will have to explain all these traces without losing their originality and that could be possible if the classroom is aided with technological aspects like VCD, CD, PPT, Mike, Speaker, Projector, musical instruments and many more... These issues could well be resolved with the help of technology. The difficulty in handling these issues could well be handled with the help of technology aided learning and teaching process by incorporating perfectly and effectively.

Different genres of literature like prose, poetry, drama, fiction and novel could phenomenally be delivered by infusing literature with technology so that technology could judiciously be used for the innovative approach using ICT to enable the teaching learning process more interesting and effective.

Poetry:

Poetry as an art form predates written text.^[1] The earliest poetry is believed to have been recited or sung, employed as a way of remembering oral history, genealogy, and law. Poetry is often closely related to musical traditions

^[2] and the earliest poetry exists in the form of hymns (such as the work of Sumerian priestess Enheduanna), and other types of song such as chants. As such poetry is a verbal art. Many of the poems surviving from the ancient world are recorded prayers, or stories about religious subject matter, but they also include historical accounts, instructions for everyday activities, love songs

[3] and fiction. But now the poetry is not recited it is only read but it could be changed with the help of technology aided teaching.

For example in my classroom when I take poetry I used recite the poem with musical instruments where it is necessary and I used to show them videos of the poem and while explaining Power point presentation with suitable background as like if forest I used to display them forest as a background.....

Especially when I taught them ‘A Solitary Reaper’ written by William Wordsworth I took them to the our college valley and it was taught with musical instruments and one of my students enacted like a girl in the poem and all my students enjoyed it and I hope they will remember it forever.

Drama

Drama is the specific mode of fiction represented in performance: a play, opera, mime, ballet, etc., performed in a theatre, or on radio or television.^[4]

In English (as was the analogous case in many other European languages), the word *play* or *game* was the standard term for dramas until William Shakespeare's time—just as its creator was a *play-maker* rather than a *dramatist* and the building was a *play-house* rather than a *theatre*.^[5]

For example when the class of mine dealt with teaching Faustus by Christopher Marlowe I made the students to create a skit of Banquet Scene using technical aids as Video camera, musical instruments, dialogue delivery with mike and speaker and the scene could be ever green in the minds of my students forever.

Novel

A novel is a long, fictional narrative which describes intimate human experiences. The novel in the modern era usually makes use of a literary prose style.

We have classic example like Jane Eyre written by Charlotte Bronte came as a film in Tamil Movie as Shanthi Nilayam and it was a massive hit. Secondly Five Point Someone by Chetan Bhagat in Tamil and Novel like Pride and Prejudice by Jane Austen and Treasure Island by Robert Louis Stevenson were massive hits and its reach was remarkable. This has happened because of the incorporation of technology into English Literature and Literature of other languages in General. We could pile up many examples for an evidence and most prominently from Indian Literature Swami and his Friends by R.K.Narayan had reached every house as it got telecasted in Television. Hence technology combined Learning and Teaching of Literature could undoubtedly reach everyone without losing its texture.

When I taught Middlemarch by George Eliot I took the entire play explaining the scenes with PPT, Video and Audio aids and my students could ever remember.

When I teach them Language and Linguistics, their performance in Group Discussion used to be recorded in video cam and same I used to telecast it to the students to check their verbal and nonverbal communication.

The Influence of Technology on Learning:

The impact of technology on English language has two sides as the coin has two sides. Due to its impact social communication through mass media has revolutionized and it offers an unrestricted access to people all around the world.

Due to its influence a new form called text-speak is prominently used nowadays by everyone and this is also used like coding words between friends, students who don't want anyone to know about what their conversation is... At the same time there is an expected problem we have in par with Standard English, as the text-speak gradually deviates from the Standard English usage.

Standard English**Text-speak**

Oh my God

OMG

Good morning

Gn

Forgive

4give

Your

ur

These discussions show that text-speak could be the new form of language for future generations, either we accept it or not. Therefore the next/net generation could not be blamed for their usage which they could use for processing/analyzing and they use it to share their feelings, opinions, ideas and emotions. Of course with technology being heading in every field Dictionary of English Language also adding up new vocabulary every day, which no include cyberpunk fiction, steampunk, the New Weird, the post-ironic novel, long arc serial television, art games, electronic literature, interactive fiction, autobiographical comics, mini-fiction(shorter than short stories)etc., have come into existence....

It is not surprising that these short short stories will lack in vivid descriptions, articulate expressions of different characters but we have to accept the changes happening in the world of English Literature as a result of the impact of Technology. Opponents of short-short stories call the post- modern digital stories as plot less and lightweight accessories. Others argue that social media communications contributed to the linguistic ruin of generation”(Axtman 2002).

Technology and Literature

As it was discussed earlier technology is advancing in a high pace and literature is changing with it. Consequently when we think of changing teaching and learning pedagogy incorporating Technology, it could also give rise to new genres of literature as short short stories with different pseudonyms

Conclusion:

ICT has connected the world irrespective of geographical distances and it is even influencing the way we think, write, express our emotions etc., hence no doubt we could not avoid using technology in teaching English Language and Literature without its help...

My research focuses that English Literature and Language could be thought easily and can be made most joyful with the help of technology even though it has its another aspect. This paper is just a beginning of my new pedagogy I made in teaching English Literature and Language to the students to make them understand their lessons easily....Undoubtedly we could make teaching learning process the most effective, most interesting and livelier classrooms with the help of technology....

References:

1. "Poet goes from newsprint to verse". 29 January 2014.
2. Francis, Norbert (2017). Bilingual and mutlicultural perspectives on poetry,music and narrative: The science of art. Lanham MD: Rowman and Littlefield.
3. Arsu, Sebnem (14 February 2006). "The Oldest Line in the World". The New York Times.
4. Elam (1980, 98).
5. Wickham (1959, 32—41; 1969, 133; 1981, 68—69). The sense of the creator of plays as a "maker" rather than a "writer" is preserved in the word playwright. The Theatre, one of the first purpose-built playhouses in London, was an intentional reference to the Latin term for that particular playhouse, rather than a term for the buildings in general (1967, 133).
6. Axtman,K.(December 2002). r u online?: The evolving lexicon of weird teens. Christian Science Monitor. Retrieved from <https://www.csmonitor.com/>.
7. Patea,V.(Ed.).(2012). Short story theories: A twenty-first-century perspective.New York, NY: Rodopi <https://doi.org/10.1163/9789401208390>