PROCEEDINGS OF

INTERNATIONAL CONFERENCE ON ICT FOR DIGITAL, SMART, AND SUSTAINABLE DEVELOPMENT

ICIDSSD '19

PROCEEDINGS OF

INTERNATIONAL CONFERENCE ON ICT FOR DIGITAL, SMART, AND SUSTAINABLE DEVELOPMENT

March 15-16, 2019

Editors

Md. Afshar Alam
Ranjit Biswas
Jawed Ahmed
Syed Imtiyaz Hassan

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
JAMIA HAMDARD (*DEEMED TO BE UNIVERSITY*), NEW DELHI

Knowledge Partner: IJIT (BVICAM's International Journal of Information Technology)

Technical Partners:

AICTE (All India Council for Technical Education)

ISTE (Indian Society for Technical Education)

IETE (The Institution of Electronics and Telecommunication Engineers)

CSI (Computer Society of India)

Published by
Department of Computer Sciences & Engineering,
School of Engineering Sciences & Technology,
Jamia Hamdard (*Deemed to be University*),
New Delhi – 110062 (India)

ISBN 978-93-5351-608-6

Disclaimer
The opinion expressed and figures provided in this proceeding of ICIDSSD '19 are the sole responsibility of the authors. The organiser and the editor bear no responsibility in this regard. Any and all such liabilities are disclaimed.

Preface

The International Conference on ICT for Digital, Smart, and Sustainable Development (ICIDSSD '19) aims to provide an annual platform for the researchers, academicians, and professionals from across the world. ICIDSSD '19, held at Jamia Hamdard, New Delhi, is the first international conference of this series of conferences to be held annually. The conference majorly focuses on the recent developments in the areas relating to Information and Communication Technologies and contributing to Sustainable Development.

ICIDSSD '19 has attracted research papers pertaining to an array of exciting research areas. The selected papers cover a wide range of topics including but not limited to Sustainable Development, Green Computing, Smart City, Artificial Intelligence, Big Data, Machine Learning, Cloud Computing, IoT, ANN, Cyber Security, and Data Science. Papers have primarily been judged on originality, presentation, relevance, and quality of work. Papers that clearly demonstrate results have been preferred.

After the formal process of peer review, the editorial board has finally selected the most relevant papers to be included in this volume. We are sure that these research works will enrich our knowledge and motivate us towards exploring the latest avenues in research.

We would like to thank our Hon'ble Vice Chancellor, Prof. Seyed Ehtesham Hasnain, for his constant and commendable support extended to us towards the path of excellence. Alongside him, we would like to thank the Pro-Vice Chancellor, Prof. Ahmed Kamal, the Registrar, Mr. Syed Saud Akhtar and other officials of the University for supporting this conference. We thank our esteemed authors for having shown confidence in us and entrusting us with the publication of their research papers. The success of the conference would not have been possible without the submission of their quality research works.

We thank the members of the *International Advisory Committee*, *Technical Program Committee* and members of all the other committees for their advice, guidance, and efforts. Also, we are grateful to our technical partners, viz. ISTE, AICTE, IETE, and CSI for sponsorship and assistance. We also thank the Department of Higher Education, MHRD for the timely issuance of ISBN for the proceedings of the conference. Finally, we are thankful to all who have contributed to the success of this conference.

Editors

Md. Afshar Alam
Ranjit Biswas
Jawed Ahmed
Syed Imtiyaz Hassan

Editorial Board

Conference Organization

Chief Patron
Prof. (Dr.) Seyed Ehtesham Hasnain
(*Vice Chancellor, Jamia Hamdard*)

Patron
Prof. Ahmed Kamal
(*Pro-Vice Chancellor, Jamia Hamdard*)

Mr. Syed Saud Akhtar
(*Registrar, Jamia Hamdard*)

Prof. Ranjt Biswas
Dean, School of Engineering Sciences and Technology

Conference Chair
Prof. Md. Afshar Alam
Head, Department of Computer Science and Engineering

Conference Secretaries
Dr. Jawed Ahmed
Assistant Professor, CSE, Jamia Hamdard

Dr. Syed Imtiyaz Hassan
Assistant Professor, CSE, Jamia Hamdard

Chief Guests
Inaugural
Prof. K.K.Aggarwal
Chairman, National Board of Accreditation (NBA), New Delhi,
Founder Vice Chancellor, GGSIP University, New Delhi.

Valedictory
Freek Jan Frerichs
Head of Department, Holland Innovation Network, Netherlands Embassy, New Delhi

Guests of Honour
Prof. Mohammad Yamin
The Australian National University, Canberra, Australia

Dr. Yasser Ades
Greenwich University, London

Mr. Rajesh Pathak
GM-HR, Newgen Software Technologies Ltd.

International Advisory Committee

Dr. Patrick Appiah-kubi,
Chair UMUC, University of Maryland University College, Washington

Dr. Hamid Saremi,
Vice-Chancellor, Azad Islamic University of Iran-Quchan Branch, Iran

Dr. Eiad Yafi
Computer Engineering, MIIT, Universiti Kuala Lumpur, Malaysia

Dr. Pourya Shamsolmoali
Department of Automation, Shangai Jiao Tong University, China

Dr. Syed Khalid Mustafa
University of Tabuk, KSA

Dr. Deepak Kumar Jain
CRIPAC, IOA, Chinese Academy of Science, Beijing, China

Prof. Ewa Lechman
Gdansk University of Technology, Gdansk, Poland

Dr. Adam Marszk
Gdansk University of Technology, Gdansk, Poland

Dr. Zeba Khanam
CCI, Saudi Electronic University, KSA

Conference Coordinators

- Dr. Farheen Siddiqui, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Sameena Naaz, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Parul Agarwal, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Ihtiram Raza Khan, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Harleen Kaur, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Shabina Ghafir, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Bhavya Alankar, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Safdar Tanweer, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Mohd. Abdul Ahad, *Assistant Professor, CSE, Jamia Hamdard*
- Mrs. Suraiya Parveen, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Md. Tabrez Nafis, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Vinita Kumari, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Syed Shahabuddin Ashraf, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Syed Sibtain Khalid, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Siddhartha Sankar Biswas, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Aqeel Khalique, *Assistant Professor, CSE, Jamia Hamdard*

- Dr. Sapna Jain, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Sherin Zafar, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Nafisur Rahman, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Md. Omair Ahmad, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Shakeel Ahmad Siddiqui, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Refia Wiquar, *Assistant Professor, CSE, Jamia Hamdard*
- Ms Asfia Aziz, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Mirza Rahil Beg, *System Analyst, Computer Centre, Jamia Hamdard*

Technical Program Committee

- Dr. M N Hoda, *Director, BVICAM, New Delhi*.
- Dr. S. Arockiasamy, *University of Nizwa, Sultanate of Oman*
- Dr. Pinnamaneni Bhanu Prasad, *Kelenn Technology (France)*
- Dr. Ravins, *Centre for Interdisciplinary Research Basic Science, JMI, New Delhi*
- Dr. Omid Mahdi Ebadati, *Head, ICT, Centre Kharazmi University, Iran*
- Dr. Suraiya Jabin, *Department of Computer Science, JMI, New Delhi*
- Dr. Syed Munawar Fazal*, College of Commerce Arts & Science, Patna*
- Prof. K. Mustafa, *JMI, New Delhi*
- Prof. R.K. Singh, *IGDTUW, Delhi,*
- Gaurav Agarwal*, Inderprastha Engineering College, Ghaziabad, U.P*
- Prof. A.K Verma, *IIT Mumbai*
- Prof. Anupam Agrawal, *IIIT Allahabad*
- Dr. Mayank Dave, *NIT, Kurukshetra*
- Dr. Naveen Kumar, *Delhi University, Delhi*
- Dr. P. Chitra, *Prathyusha Institute of Technology and Management, Chennai*
- Dr. Aman Sharma, *Himanchal Pradesh University, Shimla*
- Dr. Amirtharanjan, *SASTRA University*
- Dr. Adwitya Sinha, *Jaypee Institute of Information Technology, Noida*
- Dr. Anurag Singh Bhagel*, Gautam Budh University, Greater Noida, U.P*
- Dr. Nagarthna, *NLU, Banglore*
- Dr. Pranab Das, *Assam Don Bosco University, Assam*
- Dr. Sandeep Sharma, *GBU, Greater Noida, U.P*
- Dr. Sudeep Tanwar, *Nirma University, Gujrat*
- Prof. S.M.K. Quadri, *KU, Kashmir*
- Dr. Anil Kumar Sagar, *Galgotias University, Greater Noida, U.P*
- Prof. Sumant Katiyal, *DAV, Indore*
- Dr. D. K. Sharma, *NSIT, New Delhi*
- Dr. Saurabh Kumar Pandey, *IIT Indore*
- Dr. Gagandeep Kaur, *Jaypee Institute of Information Technology*
- Dr. Pradeep Kumar, *Maulana Azad Urdu University*
- Dr. Anuradha Chug, *Guru Gobind Singh Indraprastha University, Delhi*
- Dr. Ashish Joshi, *USICT, Guru Gobind Singh Indraprastha University, Delhi*
- Dr. Khalid Raza, *Department of Computer Science, JMI, New Delhi*

- Dr. Mohd. Sadiq, *Faculty of Engineering, JMI, New Delhi*
- Dr. Yasha Hasija, *DTU, Delhi*
- Dr. Sumit Yadav, *Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi*
- Dr. Surjit Singh, *NIT Kurukshetra*
- Dr. B. B. Gupta, *NIT Kurukshetra*
- Dr. Sarika Jain, *NIT Kurukshetra*
- Dr. Ashish Saini, *Dayalbagh Educational Institute*
- Dr. Sudesh Kumar*, Indira Gandhi National Tribal University, Amarkantak*
- Dr. Abhilasha Vyas, *G D Goenka University*
- Dr. Anuj Sharma, *Panjab University, Chandigarh*
- Dr. Mahtab Alam, *Noida International University, U.P*
- Dr. Shikha Mehta, *Jaypee Institute of Information Technology*
- Dr. Shariq Murtuza, *Jaypee Institute of Information Technology*
- Dr. Gajendra Pratap Singh, *JNU, Delhi*
- Dr. Karan Singh, *JNU, Delhi*
- Dr. Mohd. Ashraf, *Maulana Azad Urdu University*

Organizing Committee

- Mr. Tabish Mufti, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Gautami Tripathi, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Anam Saiyeda, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Sapna Jain, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Richa Gupta, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Zeeshan Ali Haq, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Imran Hussain, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Faraz Doja, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Anil Kumar Mahto, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Shah Imran Alam, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Samar Wazir, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Maria Afzal, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Naved Alam, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Amit Bhooshan Singh, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Piyush Kumar Gupta, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Jyotsana Grover, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Nida Iftekhar, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Gaurav, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Syed Ali Mehdi, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Neha, *Assistant Professor, CSE, Jamia Hamdard*
- Mr. Tabrej Ahmad Khan, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Pooja Gupta, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Waseem Khan, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Md. Nasiruzzaman, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Sohrab Ali, *Assistant Professor, CSE, Jamia Hamdard*
- Ms..Ayesha Kamal, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Taqdees Fatima, *Assistant Professor, CSE, Jamia Hamdard*

- Mr. Mohd. Shadab Alam, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Shahab Saquib Sohail, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Neha, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Nahid Sami, *Assistant Professor, CSE, Jamia Hamdard*
- Ms. Samia Khan, *Assistant Professor, CSE, Jamia Hamdard*
- Dr. Naseem Rao *Assistant Professor, CSE, Jamia Hamdard*

Technical Support Team

- Dr. Qamar Parvez Rana, *Chairman of the Committee, Jamia Hamdard*
- Mr. Pravendra Singh, *Department of CSE, Jamia Hamdard*
- Mr. Iklakh Mohammad, *Department of CSE, Jamia Hamdard*
- Mr. Md. Onais Ahmad, *Department of CSE, Jamia Hamdard*
- Ms. Tabassum Najm, *Department of CSE, Jamia Hamdard*
- Mr. Md. Rehbare Islam, *Department of CSE, Jamia Hamdard*
- Mr. Ishrat Khan, *Department of CSE, Jamia Hamdard*

Student Team

- Abdul Majid Farooqi, *Research Scholar, Department of CSE, Jamia Hamdard*
- Lubna Ansari, *Research Scholar, Department of CSE, Jamia Hamdard*
- Ashish Kumar Mourya, *Research Scholar, Department of CSE, Jamia Hamdard*
- Shafqat-Ul-Ahsaan, *Research Scholar, Department of CSE, Jamia Hamdard*
- Sheikh Mohammad Idrees, *Research Scholar, Department of CSE, Jamia Hamdard*
- Insha Naz, *Research Scholar, Department of CSE, Jamia Hamdard*
- Roshan Jameel, *Research Scholar, Department of CSE, Jamia Hamdard*
- Sana Mairaj, *Research Scholar, Department of CSE, Jamia Hamdard*
- Awwab Mohammad, *Research Scholar, Department of CSE, Jamia Hamdard*
- Ibrahim Haleem Khan, *B.Tech (CSE) VIII Semester, Department of CSE, Jamia Hamdard*
- Shiv Kumar Jha, *B.Tech (CSE) IV Semester, Department of CSE, Jamia Hamdard*
- Syed Faisal Javed, *B.Tech (CSE) VI Semester, Department of CSE, Jamia Hamdard*
- Nikhil Tiwari, *B.Tech (CSE) VI Semester, Department of CSE, Jamia Hamdard*
- Adil Rehman, *B.Tech (ECE) IV Semester, Department of CSE, Jamia Hamdard*
- Fardeen Wassey, *B.Tech (ECE) IV Semester, Department of CSE, Jamia Hamdard*
- Mohd. Kashif Saleem, *B.Tech (ECE) IV Semester, Department of CSE, Jamia Hamdard*

Contents

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Implementation of ICT Technology for Sustainable Development Goals

Sapna Jain *
Jamia Hamdard University
New Delhi
drsapnajain@jamiahamdard.ac.in

M Afshar Alam
Jamia Hamdard University
New Delhi
aalam@jamiahamdard.ac.in

Nilofer A Kazmi
ExSecretary University Grants Commision
niloufer.adil27@gmail.com

Abstract - *During this modern time technology plays a vital role in serving to to cut back international carbon emissions. just in case of massive scale engineering technologies that facilitate to capture carbon and storage. we want will to derive answer to decrease the world carbon emissions by removing GHGs from the atmosphere in massive power generators. The increased renewable generation will facilitate to decrease grid intensity. ICT will facilitate to deliver substantial carbon reductions that helps to improve business aggressiveness with further economic edges and efficiencies, whereas investment existing digital competencies to drive aggressiveness and growth. The developing world is especially liable to dynamical weather conditions and isn't well served with web and voice communications. property development is needed to meet the wants of the current that they owe to the future generations the wants that are used particularly for the essential has to the native individuals which preponderating the priority and therefore the limitations which is employed by the social organisation on the environment's ability to satisfy future needs. The bridging of the digital divide helps to help the developing world to adopt new potentialities that job in tough conditions. we've got mentioned the techniques which may help to scale back carbon emissions with the assistance of ICTs tools in property surroundings.*

Keywords: ICT, property Development, smart tools.

I INTRODUCTION

ICT uses property in organisations that have necessary think about common which are crucial to the core operation of the organisation. once approached well, associate degree organisation property strategy should address all operations, products, services and worker activities and creating that strategy all encompassing. ICT is equally essential to the whole organisation. it's important to contemplate the capabilities of all departments, as well as ICT, once making a property strategy. property development involves over growth. It needs a amendment within the content of growth, to form it less Material- and energy-intensive and a lot of just in its impact. The changes that are required in countries to keep up the stock of ecological capital by providing the equal distribution of financial gain that helps to beat the matter of the unequal distribution of resources. [1]. the strategy for economic development should be a lot of soundly primarily based upon the realities of the stock of capital that sustains it. this is often seldom drained either developed or developing countries. for instance, financial gain from biology operations is conventionally measured in terms of the worth of timber and different product extracted, minus the prices of extraction. the prices of make the forest don't seem to be taken into consideration, unless

*Corresponding Autor

cash is truly spent on such work. therefore, calculation profits from work seldom takes full account of the losses in future revenue incurred through degradation of the forest. Similar incomplete accounting happens within the exploitation of different natural resources, particularly within the case of resources that don't seem to be capitalized in enterprise or national accounts: air, water, and soil. altogether countries, made or poor, economic development should take full account in its growth of the advance is essentially stressed .in most organisations. in step with National international Action arrange survey drained 2007, seventy four of the ICT employees that are deployed in ICT departments don't seem to be in step with CSR agenda in numerous corporations. The survey by MSN surroundings in 2008 showed that over eightieth of staff would like their workplaces to be property and inexperienced initiatives may be taken by employers to achieve sustainability [1]. It will facilitate the workers to retain the potential new employees to your organisation. ICT may deliver a fifteenth reduction in international carbon dioxide by 2020, delivering a value saving in fuel, energy and therefore the value of carbon of over a trillion dollars18. in step with the Climate cluster, the carbon footprint of the ICT sector is foreseen to increase by seventy five by 2020. However, good use of ICT may contribute a discount in carbon emissions in different sectors that's five times bigger than the carbon emitted by the ICT sector. our buildings and the way we tend to use them might be much more efficient. From occupancy primarily based lighting and heating solutions to automatic systems to capture daylight or offer shade from unwanted heat, ICT features a role to play. India has good grid power generation accounted for over 0.5 the country carbon footprint in 2007, but a 3rd of this was lost in transmission creating it an entire waste. Demand management and good meters can facilitate those supply energy run a lot of economical systems, while interactive time period energy displays prompt users to contemplate their energy consumption. The Dematerialisation is swapping high carbon activities with low carbon alternatives like e-billing rather than paper request, teleconference rather than travel to conferences and e-media instead of manufacturing CDs or newspapers. good motors introducing ICT to the producing sector to vary the quantity of energy employed by production lines, instead of motors either being on or off. Using good provision ICT will facilitate higher communication and coming up with, whether or not this is often for draw networks or client delivery rounds. several come back journeys of vehicles carry no product and with eightieth of Everglade Statets having but five vehicles, higher coordination, communication and cooperation is crucial to cutting carbon. In 2008 DEFRA (Department for surroundings, Food and Rural Affairs) taken off the coverage tips on Greenhouse Gas emissions [2]. The intense implications for renewable energy tariffs and peple organizations that claim Carbon Neutrality. Previously, if organisations purchased their electricity from a renewable energy or inexperienced tariff they might have zero carbon emissions from the employment of that electricity. Defra has declared that renewable energy tariffs ought to be wont to generate electricity electricity grid and carbon emissions should resolve imputable use. The reasoning for this is often two fold. Firstly, zero-rating electricity doesn't incentivise users to scale back electricity consumption. Secondly, renewable energy tariffs don't create further carbon reductions as Utilities have a legal demand, set by Government, to provide a particular quota of renewable energy. Those corporations that are shopping for renewable electricity and claiming zero carbon can currently need to report abundant higher carbon emissions from electricity consumption. for instance, a corporation that was disbursement 1,000,000 each year on electricity from a renewable energy tariff antecedently coverage zero carbon emissions can currently report half dozen,500 tonnes of carbon from this electricity. At a median of 20 per MT of carbon, the corporate can currently need to pay an extra 130,000 to offset that carbon if they need to assert carbon neutrality. This renewable energy stance, including the Carbon Reduction Commitment, offers organisations a lot of incentive to target cutting energy consumption and people departments that contribute most to the bills ICT being one in all those which contribute to environment.

II CARBON EMISSION EFFECTS

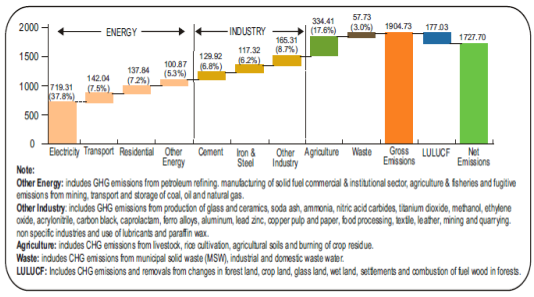
The known potential for ICT based. reductions depends on two factors i.e. the start line of every sector in terms of energy potency and therefore the role that ICT will play in facultative those reductions. It is thus not stunning that the 2 largest areas of chance known are transportation and residential services, specifically to a lot of economical use of energy in buildings. These are sectors that are presently not coated by the ETS directive, and wherever amendment has been less pressing than in, for instance, industry. Achieving reductions in these sectors would force observance and providing relevant incentives to a really fragmented set of call manufacturers, with all that this suggests in terms of knowledge gathering, process and knowledge transmission. In addition to those 2 sectors, a substantial chance was known in higher management of electricity generation, transmission and consumption. ICT-based real time military operation, transmission and process will enable higher coming up with and usage of electricity, facilitate bigger use of renewable sources in electricity generation, and turn out substantial savings across a broad vary of sectors. The impact of ICTs upon the surroundings may be distinguished in 3 dimensions [3], the primary one includes impacts connected on to the life cycle of ICT as well as production, use, recycling, and disposal. The second refers to the ICT applications which will amendment the processes of production, transport, and consumption, including, for instance, the replacement of a product with a service. The last dimension includes environmental impacts associated with behavioural changes and economic structures that are results of the ICT accessibility and usage. Therefore, ICT might be seen each as an answer and a controversy to environmental sustainability [3]. ICT systems and technologies contribute to environmental harm in numerous ways in which throughout their production and their disposal: energy use in production, consumption of serious amounts of electricity, gas emissions contribution, and high toxicity of the product, inappropriate waste management and reverse provision. it's price mentioning that with the quick growth of ICTs in business and everyday life, the overall electricity consumption of ICT is predicted to double by 2022 compared to the extent in 2010, whereas tripled by 2030 to one,700 terawatts. It has been calculable that the ICT sector (personal computers, peripherals, telecommunication networks and devices, and knowledge centers) generates two carbonic acid gas of the overall calculable emissions. Moreover, it's noticed that carbonic acid gas emissions, similarly because the power consumed in knowledge centers are bigger than those of the many countries and cities. The carbonic acid gas footprint of the ICT business is getting ready to exceed the one in all aviation industry [4]. In addition to the higher than, the poisonous substances, like lead and mercury, contained in ICT instrumentation increase the environmental burden at their disposal. The ICT instrumentation when the top of its helpful life is commonly disposed while not special treatment inflicting severe environmental issues. Environmentally friendly style, production and life cycle management of ICT may be created a lot of economical and environmentally property. The negative impact of ICT on the surroundings should be self-addressed by governments, international authorities, ICT industries, businesses and consumers. initiatives towards manufacturing and managing a lot of environmental friendly products. They will use ICTs in numerous ways in which, like [5]. The impact of ICTs upon the surroundings may be distinguished in 3 dimensions [6]. The last dimension includes environmental impacts associated with behavioural changes and economic structures that are results of the ICT accessibility and usage. Therefore, ICT might be seen each as an answer and a controversy to environmental property. ICT systems and technologies contribute to environmental harm in numerous ways in which throughout their production and their disposal: energy use in production, consumption of serious amounts of electricity, gas emissions contribution, and high toxicity of the product, inappropriate waste management and reverse provision. it's price mentioning that with the quick growth of ICTs in business and everyday life, the overall electricity consumption of ICT is predicted to double by 2022 compared to the extent in 2010, whereas tripled by 2030 to one,700 terawatts [7]. it's been calculable that the ICT sector which includes personal

computers, peripherals, telecommunication networks and devices, and knowledge centers generates two carbonic acid gas of the overall calculable emissions. Moreover, it's noticed that carbonic acid gas emissions, similarly because the power consumed in knowledge centers are bigger than those of the many countries and cities. The carbonic acid gas footprint of the ICT business is getting ready to exceed the one in all aviation industry [7]. In addition to the higher than, the poisonous substances, like lead and mercury, contained in ICT instrumentation increase the environmental burden at their disposal [8][9]. The ICT instrumentation when the top of its helpful life is commonly disposed while not special treatment inflicting severe environmental issues. Environmentally friendly style, production and life cycle management of ICT may be created a lot of economical and environmentally property. The negative impact of ICT on the surroundings should be self-addressed by governments, international authorities, ICT industries, businesses and consumers. Initiatives towards manufacturing and managing a lot of environmental friendly product. they will use ICTs in numerous ways [9].The amendment in ICTs that has daily average life which resulted within the increase of inexperienced house effectand increase in emission of carbon footprints .The Carbon footprint (CF) that is named Carbon profile is total quantity of carbonic acid gas (CO2) and different gas (GHG) emissions like alkane, inhalation anaesthetic .It helps in supply-chain end-life recovery which ends up in amendment in result [9]. It is thought of that the energy generated is pollution and shown in Figure 1 .It is considered that the energy generated is pollution and shown in In 2007, the Gartner studied which the total amount of CO2 emissions from the ICT industry which is around 2% of global carbon emissions. The ICT applications are acknowledged in the global environmental protection strategy help to huge potential to improve performance across the economy and society, as it concerns the remaining 97-98% [10].

	2009	2015	2020
Data Centers	121.30	229.87	369.48
PCs	126.69	222.41	516.55
Mobiles	1.54	3.74	6.58
Gaming Consoles	11.23	26.04	40.22
Carbon Conversion Number (CCN)	1.3	1.265	1.23
Total	260.77	482.06	932.84

Figure 1. ICT carbon footprint in megatones of CO2

In 2007, the Gartner studied that the overall quantity of carbon dioxide emissions from the ICT business which is around two of world carbon emissions. The ICT applications are acknowledged within the international environmental protection strategy facilitate to very large potential to enhance performance across the economy and society, because it issues the remaining 97-98% [8]. Figure 1. ICT carbon footprint in megatones of CO2 Th Ministry of surroundings and Forests (MoEF) calculated the online GHG emissions by Asian country, that contains of Land Use Land Use amendment and biology (LULUCF). The 1727.71 million tonnes of CO equivalent (MtCOeq) was emitted in 2007[11]. It helps to assess calculate GHG emission from energy as fifty eight per cent, from business twenty two per cent, from agriculture seventeen per cent, and waste product it's three per cent of the overall web CO emissions. The per capita CO emissions for Asian country that contains of LULUCF was calculated as one.5 tonnes per capita in 2007. The emissions were 1100.06 MtCO equivalent of that 719.31 MtCO equivalent were emitted from electricity generation and 142.04 in 2007.The MtCO emissions like 412.55 CO eq. The LULUCF sector was sequestering 177.03 MtCO equivalent as shown in. The LULUCF sector was sequestering 177.03 MtCO eq as shown in (Figure 2).



Source: India: Greenhouse Gas Emissions-2007 (May 2010), Indian Network for Climate Change Assessment, MoEF, Govt. of India

Figure 2 Sectoral GHG (MtCO2eq)in 2007

Figure 2 shows Sectoral GHG (MtCO2eq)in 2007 The GHG emissions calculated estimates done in 1994 and 2007 depicts that total GHG emissions while not LULUCF have full-grown from 1251.95 Mt in 1994 to 1904.73% MtCO equivalent in 2007 that have combined annual rate of growth (CAGR) of three.3 per cent It includes the LULUCF the CAGR that is 9 per cent. The reports from years 1994 and 2007 shows the expansion in GHG emissions in areas like cement production (six per cent), electricity generation (5.6 per cent) and transport (4.5 per cent) [11]. The small print of GHG emissions by varied sectors is conferred within the Table 1. The details of GHG emissions by various sectors is presented in the Table 1.

	1994		2007		Compound Annual Growth Rate (Per Cent)
	$MtCO_2e$	$MtCO_2e$	$MtCO_2e$	$MtCO_2e$	
Electricity	355.03	28.04	719.30	37.8	5.6
Transport	80.28	6.4	142.04	7.5	4.5
Residential	78.89	6.3	137.64	7.2	4.4
Other Energy	78.93	6.3	100.67	5.3	1.9
Cement	60.87	4.9	129.92	6.8	6.0
Iron & Steel	90.53	7.2	117.32	6.2	2.0
Other Industry	125.41	10.0	165.31	8.7	2.2
Agriculture	344.48	27.6	334.41	17.6	-0.2
Waste	23.23	1.9	57.73	3.0	7.3
Total without LULUCF	1251.95	-	1904.73	-	3.3
LULUCF	14.29	-	-177.03	-	-
Total with LULUCF	1228.54	-	1727.71	-	2.9

Table 1: Sectoralwise GHG emissions of 1994 and 2007

III ICT TOOLS

"The following methods have helped to put ICT emissions reductions into practice:"

Corporate attitudes company attitudes though the energy use of ICT is moving up the company agenda, there's abundant scope for improvement in sure areas. The Measurement shows forty two of executives responding to associate degree social scientist. A survey mentioned their organisation will not monitor its ICT-related energy disbursement. The Accountability is usually the ICT department of a corporation isn't to blame for the company is energy bills. The information centre simulation tool being developed by the BCS may be wont to break down the energy use within the data centre between the services being run there, allowing it to be beaked befittingly and procurement. Sixty three aforementioned responsibleness was a crucial think about ICT acquisition and thirty two said worth. solely twelve-tone system aforementioned that energy potency was crucial. We use advertising in some businesses are involved that greenness. It is getting used by makers as a sale static, while not them providing spare impartial and strong proof of the advantages of exploitation their equipment. This has become referred to as green-washing.

Legislation the Energy-using Products Directive Legislation the Energy-using product Directive was adopted by the Parliament in 2005. It aims to encourage makers to style product, as well as ICT instrumentation, with whole life environmental impacts in mind. for every product cluster coated, minimum standards for energy potency are set, and

solely merchandise meeting them are eligible available within the EU. General measures proscribing the overall carbon emissions of corporations might have a lot of impact than ICT-specific legislation. for instance, the 2007 Energy study introduced the Carbon Reduction Commitment (CRC), a compulsory emissions mercantilism theme for big organisations. The BCS has expressed concern that the CRC because it presently stands excludes outsourced knowledge centres from a company. It reports emissions that it's to blame for. This creates associate degree incentive for corporations to source knowledge centres, doing nothing to scale back emissions and doubtless moving skilful jobs overseas. Laws concerning knowledge retention might have a harmful result on ICT potency. Requiring that a lot of knowledge should be keep in associate degree instantly recoverable type will increase the requirement for storage in data centres.

GreenICT: an initiative by Indian Government. The strategies to scale back the carbon foot print by ICT business in India. The Indian telecommunication sector, carbon emissions has categorised emissions in 2 ways in which In the Direct emissions methodology the captive power generation of diesel generator and within the Indirect emissions method the acquisition of electricity from the Centralized power system and acquisition and therefore the use of IT and installation is completed .The Diesel generators may be wont to operate the Mobile Network towers that add the absence of electricity which is totally utilized to work and run mobile network towers utilized in areas of nonconnectivity. The inexperienced telecommunication standards offer incentive to the adoption of comparable client electronic product like Refrigerators and Air-conditioners, tringent energy potency standards. together with USA independent agency energy star that has created the point of framework exploitation carbon emission for telecommunication sector and carbon emission underneath the inexperienced House Gases (GHG) protocol of the globe Resource Institutes (WRI) .This initiative is wide utilized in Industrial sector for carbon emission accounting and use of the carbon emission of telecommunication corporations has to be coated in all dimensions within the Direct Emissions ,Stationary that is that the production of electricity Diesel generator is utilized in Mobile consumption and e-waste disposal. Fugitive emission are principally from refrigeration cooling whereas Indirect Emissions are from Stationary combustion i.e. consumption of purchased electricity, heat or steam). In the Indirect Emissions and Stationary combustion the assembly of purchased materials with method emissions and Mobile combustion. It includes transportation of raw materials with equipments and waste worker business travel and employee travel.

Labelling**:** Labelling a way to market ICT emissions reductions is by labelling instrumentation to indicate its energy potency. This informs client choices, and therefore the needs of explicit labels is also nominative for acquisition contracts. Labels may be voluntary or obligatory, and may address simply in-use energy or attempt to study the full life impact. the foremost wide used label is ENERGY STAR, which focuses principally on the energy potency in-use in on, standby and off modes. In 2008, the EU ENERGY STAR regulation was amended to want that central government organisations specify energy potency needs for ICT instrumentation purchases no less exacting than ENERGY STAR.

Mandatory labelling: The EU introduced obligatory energy labelling for white merchandise in 2006, ranking energy potency from A to G. an identical system has been advised for ICT instrumentation, but makers argue that over-simplistic as ICT instrumentation can be operated in many alternative modes. They would favour to label merely with the facility utilized in outlined by typical modes, however customers might not perceive wattages, associate degreed changing them into an energy value depends on the electricity worth that individual users pay. Intellect is piloting a green sheet.to accompany all instrumentation, which might collect knowledge on energy consumption and knowledge to assist shoppers optimize the potency of their devices.

Waste Management with smart ICT: Waste management with good ICT Rapid advances are created within the ICT business in step with Moore's Law. The variety of transistors which will be placed inexpensively on a microcircuit has doubled more or less each 2 years that ends up in premature devolution and generation of waste. Tremendous

energy savings is employed be created by employment ICT hardware, avoiding the requirement to extract raw materials, particularly extremely energy intensive materials like rare earths. bigger use of employment and safe disposal of ICT waste will therefore assist in reducing temperature change and the unharness of GHGs and introduce property of offer to the ICT business.BT has rumoured productivity enhancements of twentieth wherever it's introduced smart operating practices .Bradford Council has been concerned in an exceedingly pilot theme to see the advantages of tele-working. The teleworkers achieved associate degree improvement in productivity of between eleven and thirty eight which is measured as claims processed per hour, with accuracy levels remaining similar. There was additionally associate degree eightieth reduction in absence rates, and employees rumoured improved job satisfaction and quality of life. Sun Microsystems has audited its openwork programme this year on the average every worker edges from a 2-hour reduction in commute per week. Over one year, this reduction ends up in two tonnes of carbon savings (including rebound effects) and a private money saving of 1,000 in fuel and non-depreciation conjugation prices. Swedish company Telia has reduced its carbon dioxide emissions by fiftieth in 3 years by reducing workplace house, working a lot of flexibly and eliminating most essential business travel.40 United Kingdom of Great Britain and Northern Ireland employees waste up to 689,000 hours a month waiting in airports. The ensuing economic prices are calculable to exceed 450 million p.a..41 Universal wireless access across all conveyance would facilitate to extend productivity. Codes of observe for home operating would facilitate business and people to maximise the advantages of tele-work. nowadays Denmark and eire are the sole countries in Europe with such codes [12].

Datacenters**:** Datacenters are the quickest growing emissions sub-sector. Growth in capability, processing power and variety of servers are driving up emissions. This accelerated growth has return to the eye of all stakeholders, from Governments, corporations and makers, to the final public. However, research and description of best-practices has shown that there's substantial area for reduction in datacenters emissions. Currently data centers consume energy each in server power consumption and in cooling and power systems. Existing knowledge shows that non-operation connected consumption (e.g. cooling) represents up to fiftieth of datacenters total consumption. Best-practice benchmarking has shown that up to fortieth energy savings are getable through intelligent system and housing style, most through reduction of cooling and powering needs not associated with actual server usage. in addition, developments in multi-processing programs and systems are permitting a rise in server virtualisation, reducing the quantity of servers required by pooling resources and utilizing unused server power. Overall, a fortieth reduction in server energy wants, and twenty eight reduction in variety of servers required is also achieved through the introduction of each best-practices and new virtualisation technologies. The result's a potential forty four reduction in datacenters subsector emissions.

Green ICT: Inexperienced ICT inexperienced ICTs support police work systems for natural surroundings and resources, as a method to guard and restore natural ecosystems potential (forests, rangelands, lakes, rivers, water reservoirs, wetlands etc.) and initiate hindrance activities, like remote observation systems tele-detection and alarm systems in reference to forest fires, floods, erosions and usually because it issues temperature change indexes, environmental observance, GIS technology etc. device tele-detection exploitation wireless network technology will facilitate to gather from remote ecosystems time-series of environmental knowledge, that are sent wirelessly and registered into native databases in institutes, labs etc. Further, environmental knowledge is analysed getting to study sensitive environmental indexes whereas visual graphs are ceaselessly created for comparison and checking of development trends through simulation models that imitate effectively the environmental reality. in addition, future forecasts are provided, whereas call-making stakeholders will value various future situations with decision support systems, and given sure parameters we will offer property environmental management, wise use of natural resources, and hindrance of natural disasters [13].

Green Informatics Technique: Inexperienced science Technique -It will contribute through the strategies as well as Inventing innovative energy saver systems, technologies and smart devices, exploitation ,smart energy management, Applications for energy saver policies exploitation renewable sources, solar power and electrical phenomenon, wind energy, bio-fuel, bio-climatic technology, anti-pollutants technology. The employment and reducing of e-waste like previous IT systems, chips, PC, hardware, printers, mobile phones, etc. The fortieth of the overall energy consumption is because of households, thus innovative smart houses, inexperienced constructions of bio-climate material and green design creating use of innovative energy device IT systems are able to do to live, manage and cut back electricity consumption and air-conditioning needs. throughout the last decade, technical and industrial product makers were basically obligated to vary direct, regarding their energy consumption, as a results of the financial condition additionally to the increased environmental awareness of the general public. the priority taken by the producer towards energy reduction covering each data processor, from the laptops and mobiles to the information centres, has been and can be presumptively productive. Actually, shoppers show their preference for good devices, new, perceptibly less energy-consuming technologies, renewable energy sources and updated, a lot of economical cooling systems with improved energy management software package . Qualifying product are rewarded with official certification for meeting or extraordinary potency tips. The transition to a low-carbon society by 2050 includes the vision to measure and add low-energy and low-emission buildings, with intelligent heating and cooling systems. Cars and trains ought to be electrical and hybrid and thus our living-environment are less impure and cleaner. it's additionally necessary throughout international financial condition that the world of inexperienced construction and house energy improvement will result in new job places. In Figure three, it's conferred the the typical per capita final energy consumption of households in 2005 and 2010, divided into electricity. In Figure 3, it is presented the the average per capita final energy consumption of households in 2005 and 2010, divided into electricity “.

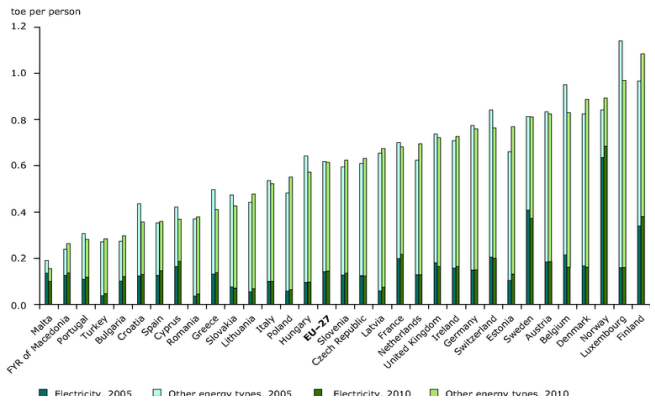


Figure 3. Average per capita final energy consumption of households in 2005 and 2010, divided into electricity consumption and other energy types (Source: EEA)

IV CONCLUSION

A property development approach, as delineated higher than, is essentially holistic. Analyzing the impact of economic, social and environmental trends, and therefore the interactions between them, is barely doable from a basis of thorough, researched understanding each of gift circumstances and of probably changes in those circumstances. info and Communication Technologies will play a key role within the environmental protection, the environmental property, the environmental education and therefore the rural property development. inexperienced IT enhances the property of computing through producing lower impact materials and product, reduced energy consumption of information centers and computers, and higher employment and finish of life management [13].

exploitation inexperienced science tools, services and technologies will contribute to the Environmental and Rural property. Yet, inexperienced science cannot substitute individuals and their behaviour that also constitutes the foremost crucial issue among environmental protection and property.

REFERENCES

[1] Avital, M., Lyytinen, K., King, J. L., Gordon, M. D., Granger- Happ, E., Mason, R. O., and Watson, R. T. 2007. Leveraging Information technology to Support Agents of world benefit. Communications of he AIS (19), 567-588.Broadband Commission.2012

[2] Ernst and Young. 2011. The Role of Green ICT in Enabling Smart Growth in Estonia. Available at: http://www.pamlin.net/new/wpcontent/uploads/EY_MKM_Green_ICT_study_2011_FINAL-REPORT2.pdf

[3] Towards Green ICT Strategies: Assessing Policies and Programmes on ICT and the Environment. OECD Conference on "ICTs, the environment and climate change", Helsingør, Denmark, 27-28 May 2009 www.oecd.org/sti/ict/green-ict.

[4] Zacharoula S. Andreopoulou, Green Informatics: ICT for Green and Sustainability. Journal of Agricultural Informatics. 2012 Vol. 3, No. 2

[5] http://www.gsmworld.com/documents/mobiles_green_manifesto_11_09.pdf

[6] Center for Sustainable Systems, 2011,http://css.umich.edu

[7] Lorenz M. Hilty and Wolfgang Lohmann,an annotated bibliography of conceptual frameworks in Ict for sustainability

[8] Andreopoulou, Z. (2012). Green Informatics: ICT for green and Sustainability. Journal of Agricultural Informatics. Vol. 3, No 2.

[9] Andreopoulou, Z.S., B. Manos, N. Polman and D. Viaggi (2011). Agricultural and Environmental Informatics, Governance, and Management: Emerging Research Applications.IGI Global. USA.

[10] The Stern Review on the Economics of Climate Change, available from www.webcitation.org/5nCeyEYJr.

[11] SMART Portugal 2020: Reducing Emissions and Increasing Energy Efficiency through ICT

[12] The Broadband Bridge: Linking ICT with Climate Action for a Low-Carbon Economy. Available at: http://www.broadbandcommission.org/Documents/Climate/BD-bbcomm-climate.pdf

[13] Accenture (2008). Data Centre Energy Forecast Report. Final Report, Silicon Valley Leadership Group.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Heterogeneous Smart Devices and Cloud Computing Based Architecture for Handling Big Data in Smart Farming: A Sustainable Approach

Harleen Kaur *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
harleen.unu@gmail.com

Shafqat-Ul-Ahsaan
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Ashish Kumar Mourya
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract - *Sensing devices are currently used everywhere for capturing and gathering information for personal and commercial use. The information generated by these sensing devices has heterogeneity and huge in volume like big data. There are so many areas such as farming, remote sensing and weather forecasting where smart devices and cloud computing has been applied. Farming using these kinds of devices has gained more attention for boosting productivity and revenue as well. Smart Farming is an innovative agricultural management concept using advanced techniques and sensing devices to increase the variety and quantity of crops and their products. Crop productivity may be increase by understanding and forecasting yield performance in different ecological conditions. In this study, we proposed an architecture based on smart devices for handing big farming data in Cloud computing. The proposed architecture is based on internet of things to sense and analyze the condition and store information in cloud for processing and extracting meaningful information.*

Keywords - *Sensor Networks, Internet of things, Smart Farming, Wireless network, Big data, Heterogeneity*

I. INTRODUCTION

Smart devices technology like Internet of things and cloud computing are two different technologies that has changed the digital era and became important part of life. Internet of Things (IoT) is a collection of objects that connected with each other in such a fashion that they form a network. These networks of objects are capable of communicating with each other by means of internet, mobile phones, sensors that have implanted at specific areas, Radio Frequency Identification (RFID), General Phone Radio Service (GPRS) and computers [1]. These objects communicate with the help of unique addressees that allow them to be verified and addressed over the network. The objects in the network are assigned with a specific task and send reports to the users [2]. The power of Internet of Things based networks can be used in multiple fields to make living easy and smart. The implementation of smart devices in various fields especially where the supervision is needed is gaining more focus and popularity. When we implement the IoT based networks, it lessen the human effort of reporting certain things on regular basis in a cold storage system, the temperature is set at a specific range and it requires a regular supervision in order to check whether the temperature is adequate or it has increased or decreased. Therefore, if we

*Corresponding Autor

implement a sensor in such a storage system, the temperature is set at a certain point and sensor senses the temperature at regular intervals [3]. The sensor tunes the temperature in a way feasible for the safe storage of fruits or vegetables that have stored in a cold storage system. In the same way, IoT based networks can be utilized in agriculture in a variety of circumstances. The main challenge that has been spotted in the domain of agriculture is to offer the necessary information and timely help to the farmers. It seems very hard to trace out the knowledge that will help in sustainable agriculture [4].

There are a range of web applications that come up with tricky solutions to increase exchange and acquirement of information regarding sustainable agricultural techniques. These applications consist of Expert Knowledge System, Agricultural Information Retrieval System, Knowledge Management System and Agro Advisory System having the capability of exchanging and reusing agriculture-based knowledge.

II. BACKGROUND

Susan F. Ellakwa [1], proposed a system for multi-matching and integration method. The supported programming languages used for this system are ASP.Net and C# to compare crop entities. An ontology system is built using XML. The information-based retrieval system allows farmers to get answers to their queries as per their requirement. Xiangyu HU et al. [2], presented an IoT based application where Radio Frequency Identification (RFID), smart sensors and GPS are implanted to exchange information from the agricultural fields. RFID were used to record information and the fields are monitored by means of sensors implanted at specific locations. Richard K. et al. [5], presented a consistent distributed mobile based structural design to offer farmers timely information. A spray quality application was developed by creating a three-layered design with cloud hosted middle-ware. The proposed system allows farmers to gain access over a variety of information through wi-fi or 3G. Within [6], the authors' have proposed an advice-giving system for cotton crop. The developed system is an integration of web services, cotton ontology and Development Advisory System. With the help of this system, farmers are allowed to submit their queries and are answered accordingly. The structural design for this system comprised mainly of five components- the SQL database, the Resource Description Framework (RDF) knowledge includes data regarding cotton, data concerned to a specific geographical area mapped through Google earth. The systems that make use of IoT and cloud highlights the architectures that makes timely information available to the farmers at their disposal. In [7], presented a system provides offline communications over Bluetooth among farmers. The system is improved with a significant edge by using Computer Graphics Metafile (CGM) models over IoT. IoT based networks using sensors and cloud is very efficient and reliable but nobody can deny the truth that it is difficult to install. Depending on wireless technology like GPS, GPRS, Wi-Fi etc. can be devastating in case of unaccepted crash.

Xiaojing et al. [8], focus on the use of cloud-based systems to confirm the link between cloud and IoT as for as the agricultural prospects are concerned. The authors' declared that smart agriculture is the application of Internet of Things, which in fact has a wide purpose and dazzling future. ZigBee is a low-cost, less power consumption and wireless networking standard [9]. The ZigBee knowledge system provides the platform to identify pests in crops, prediction of drought and moisture level in soil. If the farmer has all this information available at a real-time interval, the irrigation to the fields is controlled; pesticides are sprayed in advance to control the diseases that may occur. Joe et al. [10], studied that smart agriculture has happened to be an important issue. In agricultural applications, sensor networks and IoT play a key role while monitoring environment related parameters to enhance the agricultural production as these technologies offer a high resolution over sensing data captured from real world analog/physical signals. Maumbe et al. [11], proposed a framework regarding applications of information and communication technology (ICT) in agriculture. The author declared that ICT is a powerful tool for improving rural development and agriculture standards.

III. IoT Systems in Smart Farming

Internet of Things (IoT) provides a platform that enables the farmers to lessen poverty and support the standard of living in rural areas. The farmers living in rural areas can grow a variety of crops harvesting those using organic greenhouses; those crops are consumed locally in addition they can also export to rest of the world. This type of agricultural practices allows farmers to produce extra income and improve their standard of living and also take part in gross domestic product (GDP). With the help of IoT, agricultural surveillance programs can be held for public awareness about different types of diseases, how pests spread, how to prevent the growth of plant eating pets, provide knowledge to farmers about pesticides spray schedule, give knowledge to farmers about the symptoms of most affecting crop diseases. This proves to be fruitful for farmers to take preventive measures prior to the condition gets out of control [12].

IoT technologies make it easy to track farm products from source to their destination while transporting. This is beneficial for farm products that need additional processing, so the buyers get the information in advance about the arrival of farm products to their destinations and plan their duties accordingly. As we know, that rural areas are thinly populated, so transportation of farm products is a problem. With the help of IoT, transporters can get the information about who is in need of transport. As a result, transporters have no need to wait u until they have full truck load, they can leave anytime as they have information in advance that farmers are waiting for transport [7]. If livestock or crop smart health cards are came in practice which save information about livestock or crops that are affected proves to be beneficial for both the agriculture or veterinary officer and the farmer. The smart card contains all the essential information related to the affected livestock or crop, so it becomes very easy and is the efficient way to make diagnosis regarding any disease and to prescribe medicines.

IV. Sensing devices in Smart Farming

Working of Internet of Things (IoT) based smart farming for sustainable development in agriculture. The Level of methane in field can also be monitored the help of gas sensor for management of cattle feed to achieve minimum methane emission which is a greenhouse gas. In addition to environmental methane data, data of soil moisture, temperature and humidity level could also be acquired from soil moisture sensor and digital humidity and temperature sensor [13]. The captured data could be stored in real time on cloud based IoT platform like Ubidots from Arduino Uno. The IoT platform Ubidots could provides real time dashboard which can be accessed from portable devices like smart phone, computer, laptop or tablets from anywhere. The captured data on IoT platform could be used for further analysis and control [14, 15].

If it is possible for the ruler areas to install satellite transmission systems, it offers plenty of jobs for locals who could provide low-cost solutions, cheaper network services to the communities. The satellite transmission allows farmers living in rural areas to get information about various products and their market prices, services offered by government. The facility of mobile internet in combination with inexpensive sensors allows farmers to communicate directly with consumers and this way bypassing the middlemen who make use of farmers and derive benefit from them [16].

With the help of IoT, animal theft can be prevented to a large extent by fitting radio frequency identifiers (RFIDs) that make it easy to trace animals. In case if animal theft occurs, the location of the animal can be seen over a map in a control centre by means of data that is sent over a wireless medium [17]. In rural areas, there is a daily activity of animal grazing, there is a chance of animals get lost while grazing. Livestock are fixed with radio frequency identifier chips and RFID readers are installed at multiple locations to exchange information with the agricultural services centre to find the location of the animal [18].

V. Proposed Architecture of Smart devices and Cloud computing in Smart Farming

The Proposed modelist rely on the Internet of Things, that intended with the sensing belief and supports almost any networks provided by different services providers. The first part of proposed architecture provides observations to sensors and sensor-related data, using the underlying ontology and inters connected Data principles.

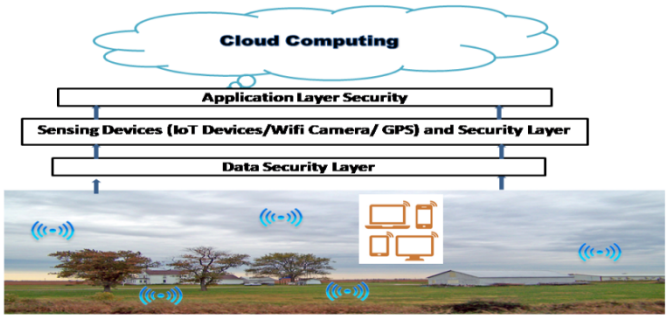


Fig. 1 Proposed Smart device Architecture using cloud computing as backbone in smart farming

The sensing devices like cameras and IoT embed system send climate related information to different locations of farms and poultry grounds. The proposed model has also designed to conserve electricity and communication network. Different environmentalist and geologist used this information for decision-making and implementing policies. The cloud computing technology provides interface to authenticate user and access information for gathering. Mobile systems may be also considered for sending commands to smart devices to make them energy saving. The cloud-computing paradigm has revolutionalized the wireless networking and data mining. The users from different locations can access and send any message to smart devices.

Cloud computing service providers have ensured the application layer security of farmer's data. In the proposed model, this layer is also responsible if any security breach affects its user's service communications. A cloud space provider offers various authentication checks to validate the identity of the users.

Data security layer has assured safe and sound peer-to-peer information transportation from sensing devices to upper layer. Our proposed security mechanism is more secured in comparison to the existing sensing device models. This layer is also responsible for fault tolerance and network breach related problems.

VI. Conclusion

Smart Farming is an innovative research area where different environmentalist and scientist doing research for making farming easiest to generate more revenue. Smart devices technologies such as RFID, IoT, and Sensor make it easy to track farm products from source to their destination while transporting. This is beneficial for farm products that need additional processing, so the buyers get the information in advance about the arrival of farm products to their destinations and plan their duties accordingly. This paper has proposed an architecture that use smart devices and cloud computing for handling sensors data for useful information extraction. The acquired and stored data in cloud could be processed and analyzed to achieve the threshold values which cloud be set for controlling the sensing and connecting devices in future.

REFERENCES

[1]Ellakwa, S.F., El-Azhary, E.S. and El-Kafrawy, P., 2012. Integrated ontology for agricultural domain. *International Journal of Computer Applications*, *54*(2).

[2] Hu, X. and Qian, S., 2011, November. IoT application system with crop growth models in facility agriculture. In *2011 6th International Conference on Computer Sciences and Convergence Information Technology (ICCIT)* (pp. 129-133). IEEE.

[3] Lomotey, R.K. and Deters, R., 2014, June. Management of mobile data in a crop field. In *2014 IEEE International Conference on Mobile Services* (pp. 100-107). IEEE.

[4] Chaudhary, S., Bhise, M., Banerjee, A., Goyal, A. and Moradiya, C., 2015, January. Agro advisory system for cotton crop. In *2015 7th International Conference on Communication Systems and Networks (COMSNETS)* (pp. 1-6). IEEE

[5] Awuor, F., Kimeli, K., Rabah, K. and Rambim, D., 2013, May. ICT solution architecture for agriculture. In *2013 IST-Africa Conference & Exhibition* (pp. 1-7). IEEE.

[6] Xiaojing, Z. and Yuanguai, L., 2012, September. Zigbee implementation in intelligent agriculture based on internet of things. In *2nd International Conference on Electronic & Mechanical Engineering and Information Technology*. Atlantis Press.

[7] Dlodlo, N. and Kalezhi, J., 2015, May. The internet of things in agriculture for sustainable rural development. In *2015 international conference on emerging trends in networks and computer communications (ETNCC)* (pp. 13-18). IEEE.

[8] Jiang, J.A., 2014, August. Becoming technological advanced-IOT applications in smart agriculture. In *APAN 38th meeting*(pp. 11-15).

[9] Maumbe, B.M. and Okello, J.J., 2013. Uses of Information and Communication Technology (ICT) in agriculture and rural development in sub-Saharan Africa: Experiences from South Africa and Kenya. In *Technology, Sustainability, and Rural Development in Africa* (pp. 113-134). IGI Global.

[10] Georgakopoulos, D.; Jayaraman, P.P. Internet of things: From internet scale sensing to smart services. Computing 2016, 98, 10. [Google Scholar] [CrossRef]

[11] Dawes, N.; Kumar, K.A.; Michel, S.; Aberer, K.; Lehning, M. Sensor Metadata Management and Its Application in Collaborative Environmental Research. In Proceedings of the IEEE Fourth International Conference on eScience, eScience '08, Bloomington, IN, USA, 7 December 2008; pp. 143–150.

[12] Jeung, H.; Sarni, S.; Paparrizos, I.; Sathe, S.; Aberer, K.; Dawes, N.; Papaioannou, T.G.; Lehning, M. Effective Metadata Management in Federated Sensor Networks. In Proceedings of the 2010 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC), Newport Beach, CA, USA, 7–9 June 2010; pp. 107–114.

[13] Michel, S.; Salehi, A.; Luo, L.; Dawes, N.; Aberer, K.; Barrenetxea, G.; Bavay, M.; Kansal, A.; Kumar, K.A.; Nath, S.; et al. Environmental Monitoring 2.0 (Demonstration). In Proceedings of the 25th International Conference on Data Engineering (ICDE), Shanghai, China, 29 March–2 April 2009.

[14]A.B. Zaslavsky, C. Perera, D. Georgakopoulos, Sensing as a service and big data, in: International Conference on Advances in Cloud Computing, ACC-2012, Bangalore, India, July 2012, p. 219.

[15] V. Chang, Y.-H. Kuo, M. Ramachandran, Cloud computing adoption framework: A security framework for business clouds, Future Gener. Comput. Syst. 57 (2016) 24–41.

[16] L. Lamport, Password authentication with insecure communication, Commun. ACM 24 (11) (1981) 770–772.

[17] J. Katz, P. MacKenzie, G. Taban, V. Gligor, Two-server password-only authenticated key exchange, J. Comput. System Sci. 78 (2) (2012) 651–669.

[18] T. Xiang, K. Wong, X. Liao, Cryptanalysis of a password authentication scheme over insecure networks, J. Comput. System Sci. 74 (5) (2008) 657–66

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Optimized Aggregation Approach for Sustainable MANET

Sherin Zafar *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
zafarsherin@gmail.com

Deepa Mehta
FET, MRIU, Haryana
deepa.mehta12@gmail.com

Samia Khan
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Gautami Tripathi
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract - *Recent Years have accentuated a remarkable rise in the popularity in the field of mobile computing owing to the size reduction of the computing devices and is also contributed by the exceptional increase in the processing power availability of mobile laptops rendering various computer applications to reach the ever-rising population segment. This increase in the number of users poses more challenges in the field of networking. Aggregation has so far has been able to curb the problems caused in order to improve the scalability of the network. This Paper studies and analyses various aggregation Techniques applied in MANET points out the improvements in various QOS parameters resulted due to the application of aggregation on routes, data and address in the field of MANETs. It also demonstrates the security enhancements possible in the field of MANET contributed by aggregation.*

Keywords *- Address Aggregation. Route Aggregation Data Aggregation, MANET, Security.*

I. INTRODUCTION

An Ad-hoc network [5] is a collaborative participation of a set of self-administrating mobile nodes not requiring deployment of any infrastructure. Each node acts as a specialized router forwarding data from source to the destination in accordance with a set of rules which are specifically designed for ad-hoc networks capable of working taking into account its unique features. Over the years various Routing Protocols [2] [4] have been proposed and implemented, however, due to the limitations of the MANETs the Routing Protocols design gets restricted and researches are on trying to refine the protocols in order to achieve better performance due to specific architecture of MANET as depicted in Figure.1.

The various challenges faced by MANET are [1]:

- **Limited radio range**: MANET works in a wireless range hence is characterized by a limited bandwidth. Bandwidth thus needs to be used

*Corresponding Autor

ideally; one such way to achieve this is to design a routing Protocol with minimum overhead. The limited range of transmission also enforces restriction on routing strategies in order to sustain the information pertaining to topology of the network. Due to high mobility nodes in MANET, the network topology keeps changing and thus retaining the topological information for each and every node involves high overhead resulting in added use of bandwidth

- **Characteristics of wireless medium**: The wireless channel suffers from a lot of disorders pertaining to its nature such as fading, obstruction, etc. These features pose a limitation on the data rate and make the cordless transmissions inconsistent.
- **Mobility**: MANET consists of highly mobile nodes with freedom to get involved in the network whenever required and exit the network at will, resulting in a constant change in the topology of the network leading to numerous link breakages. Such situations often require alterations in the routes. This imposes additional constraint in the development of efficient routing protocol.
- **Energy constraints**: The mobile devices comprising the MANET are restricted in power resources. In the absence of any significant research for improving the battery life of the devices, the entire consequence is borne by the routing protocol which needs to be energy efficient.

Fig.1 A basic-architecture of mobile ad-hoc network-MANET

- **Routing**: The mobile nodes with distinct speeds prompt a change in the network topology with their movement. Routing thus becomes a major challenge as it causes degradation in performance keeping in view the unicasting and multicasting demands posed by the nodes.
- **Security issues**: The wireless environment poses additional challenge on the ad-hoc networks. Since all the devices share a common channel. Data transmission through a node is acknowledged by entire set of participating nodes, making it possible for any invader to sneak in and access the information communicated within the network.

In order to accommodate the limitations of ad-hoc networks and derive an efficient routing protocol, several researches have undertaken the concept of aggregation applied to addresses, routes and data

to save scarce resources in MANET. This paper studies the various aggregation schemes and analyses the impact of each kind of aggregation on the performance of ad-hoc networks and also on the security of the network.
The next sections of this paper briefly explain concepts, advantages and applications of Address, Route and Data Aggregation followed by Conclusion and References.

II. ADDRESS AGGREGATION

Shiflet, C. et. al. [8] have appraised the potential of aggregation of addresses in MANETs to comprehend the effect of aggregation on the routing overhead. The paper takes into account the attribute of grouping flat address spaces in MANETs if restrained physically. The paper tries to improve the scalability of the network by using address aggregation to abate the impact of routing overhead. Aggregation is applied on AODV [7] as it includes all the constructs required which support aggregation which includes the existence of network prefix field in the Route Reply as well as in the routing tables. Various modifications formulated while implementing aggregation in AODV is mainly addition of support for Longest Prefix match algorithm, enabling AODV to acquire capability of supporting the much needed aggregation, where multiple nodes can be addressed with a single route and thus a single reply can serve multiple requests as depicted in Figure.2.

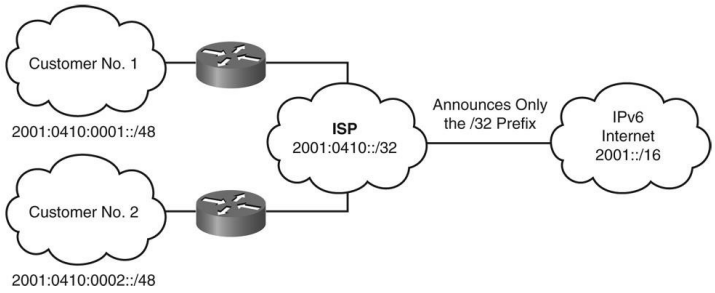


Fig.2 Example of Address Aggregation

Intrinsic nature of aggregation demands an assigned node in the subnet to be held accountable for routing on the behalf of other remaining nodes of the subnet and is designated by subnet router. Forming a subnet is difficult in MANETs due to the highly mobile nodes, an effective grouping of nodes is a challenging task. It also requires a proper configuring of the addresses as the determination of aggregated nodes will be dependent on the relationship between their addresses. Various applications can be used for configuring the membership in the subnet and maintain the addressing. The paper employs static methods for the determination of leader node. Route Request in an address aggregated AODV is broadcasted and the Reply is handled by the Subnet Router. It takes a decision after finding the destination of route request and its own address and mask to be the same. The results of the research paper record a reduction in routing overhead from 42% to 82%. Delay was decreased from 30% to 61%.

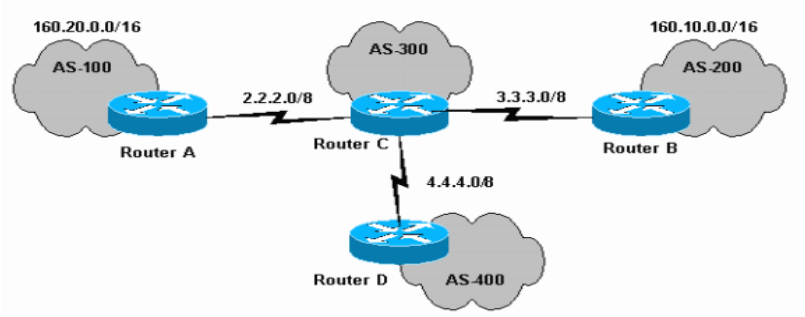
III. ROUTE AGGREGATION

Naito, K. et. al. [6] have exemplified a multicast routing protocol designed to diminish the packet collisions with the help of aggregation. The primary cause behind the collision of packets is the data forwarding by the nodes which are at the same hop distance from the source node and the retrieval and further forwarding of the data packet happens at the same time instance resulting in multi forwarding of data packets and thus causing collisions. The proposed protocol employs a mechanism to aggregate

multicast routes at the forwarder nodes, to avoid the collision of packets due to hidden node problems. As a result, the forwarding of redundant packets is reduced as depicted in Figure.3.

The protocol also employs the mechanism to control the transmission power of the data packets in accordance with the downstream device information provided which is the signal to noise ratio. The scheme uses ODMRP (On Demand Multicast Routing Protocol) [9] as the base protocol. The changes in the base protocol include the editing of joint query and addition of the number of downstream nodes and the route transmission power fields in the extension field. Joint Reply is edited by adding the hop count and transmission power required.

Fig.3 Example of Route Aggregati



The join query messages are broadcasted by the source node into the entire network. The messages include the transmission power required for the route and also the number of nodes downstream. The downstream nodes receiving the joint query messages start estimating the SNR of the received messages along with the calculation of the transmission power required per route. The next node with minimum transmission power required is set as the next hop address in the routing table. The new route transmission power is also calculated. The node with minimum power required is selected as the upstream node and multicast routes can be aggregated, thus avoiding collisions. Advantages of Route Aggregation [3] include:

- Costing: Aggregation helps in reducing the size of routing tables thus occupying less memory making the routing process less cumbersome for the computing resources.
- Performance Enhancement: The reduction in the size of routing tables leads to a quicker lookup as a result the processing and queuing time is reduced also the route advertisements are lessened.
- Improvement in Network Stability: Aggregation requires a head of the subnet to route the data for the entire subnet. In case of failure of head, the instability is not propagated throughout the network.

IV. DATA AGGREGATION

Infrastructure less, power constrained MANETs face challenges employing techniques to monitor and detect cyber-attacks. Data aggregation techniques have been proved efficient to achieve security in MANET. **Zhang, D. et. al.** [14] have developed certain aggregation techniques supplementing the reduction in energy cost during the transition of information and minimizing the consumption of energy. The paper details two kinds of aggregation techniques:

a. Lossless Aggregation: This aggregation amalgamates the detection information and ensures no loss of data when decompressed.
b. Lossy Aggregation: The amalgamation in this aggregation ensures the data to contain almost all the vital information but with some loss of precision.

The data aggregation techniques have been validated to achieve security in MANET [10] [11] [12]
[13] along with managing energy efficient transmission of information.

Table.1 Aggregation as a Source of Improvement in MANET

Technique	Delay	Packet Delivery Ratio	Security	Efficiency
Address Aggregation	Reduced	Improved	No Effect	Improved
Route Aggregation	Reduced	Improved	No Effect	Improved
Data Aggregation	Reduced	Improved	Improved	Improved

V. CONCLUSION

This paper discusses about various routing protocols have been proposed and implemented, however, due to the limitations of the MANETs the Routing Protocols design gets restricted and researches are on trying to refine the protocols in order to achieve better performance through aggregation methodology as

presented in this paper. Aggregation can be classified as Route, Data and Address Aggregation whose advantages on MANET are depicted in Table.1. Intrinsic nature of aggregation demands an assigned node in the subnet to be held accountable for routing on the behalf of other remaining nodes of the subnet and is designated by subnet router. Forming a subnet is difficult in MANETs due to the highly mobile nodes, an effective grouping of nodes is a challenging task. It also requires a proper configuring of the addresses as the determination of aggregated nodes will be dependent on the relationship between their addresses. Various applications can be used for configuring the membership in the subnet and maintain the addressing. Data aggregation provides its advantages over cost, scalability and performance enhancement whereas data aggregation is a biggest boost over maintaining security in MANET.

REFERENCES

[1] Aarti and Tyagi, S. S. 2013. Study of MANET: Characteristics, Challenges, Application and Security Attacks. International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3.

[2] Corson S, Macker J. 1999. Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations. Network Working Group, RFC2501.

[3] Loo, J., Mauri, J. L., and Ortiz, J. H. 2012. Mobile Ad Hoc Networks: Current Status and Future Trends. CRC Press.

[4] Mehta, D., Kashyap, I., Zafar, S. 2017 Routing Optimization in Cloud Networks. International Journal of Advanced Research in Computer Science, vol 8, no. 2, ISSN 0976-5697.

[5] Mobile Ad-Hoc Networks (MANET) Charter IETF, http://www.ietf.org/html.charters/manet- charter.html

[6] Naito, K., Fujii, M., Mori, K. and Kobayashi, H. Multicast routing protocol for ad-hoc networks with route aggregation and transmission power control. Japan.

[7] Perkins, C. E., Belding-Royer E. M. and Das, S. R. 2002. Ad Hoc On Demand Distance Vector (AODV) Routing. Internet Draft, Internet Engineering Task Force (IETF), Cornell University, NY.

[8] Shiflet, C., Belding-Royer, E. M. and Perkins, C. M., 2004. Address Aggregation in Mobile Ad hoc Networks. In Proceedings of the IEEE International Conference on Communications (ICC), Paris, France.

[9] Yi, Y., et. al. 2002. On-Demand Multicast Routing Protocol (ODMRP) for Ad Hoc Networks. IETF MANET Working Group.

[10] Zafar, S., Soni, M. K. 2014. Biometric Stationed Authentication Protocol (BSAP) Inculcating Meta-Heuristic Genetic Algorithm. I.J. Modern Education and Computer Science, 28-35.

[11] Zafar, S., Soni, M. K. 2014. A Novel Crypt-Biometric Perception Algorithm to Protract Security in MANET. Genetic Algorithm . I.J. Computer Network and Information Security, 64-71.

[12] Zafar, S., Soni, M. K. 2014. Secure Routing in MANET through Crypt-Biometric Technique. Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA),713-720.

[13] Zafar, S., Soni, M. K., and Beg, M. M. S. 2015. QOS Optimization in Networks through Meta- heuristic Quartered Genetic Approach. ICSCTI, IEEE.

[14] Zhang, D., Ge, L., Hardy, R., et.al., 2013. On effective data aggregation techniques in Host-based Intrusion Detection in MANET. In IEEE 10th Consumer Communications and Networking Conference (CCNC), Las Vegas, NY, pp. 85-90.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Prediction of PM10 and PM2.5 Pollutants using Machine Learning

Mukta Satsangi
Department of Botany
School of Chemical and Life Sciences,
Jamia Hamdard, New Delhi

Vinita Kumari *
Department of Computer Science and Engineering,
School of Engineering Sciences and Technology,
Jamia Hamdard

Abstract- *Air pollution has become a global environmental concern. The air pollutants like nitrogen dioxide (NO_2), sulphur dioxide (SO_2), particulate matter (PM10 and PM2.5), ozone (O_3) and carbon monoxide (CO) in air shows harmful effects on human beings by affecting their health when increased above permissible limits. Most harmful air pollutants are PM10 and PM2.5 which due their small size moves easily from lungs to blood circulation. In this work a predictive model for PM10 and PM2.5 concentration has been developed using Machine Learning, a computational method. The algorithm used for this purpose is multiple linear regression (MLR), a supervised machine learning algorithm. This algorithm is applied on air pollution dataset collected from three polluted areas of Delhi: Anand Vihar, Punjabi Bagh and R.K. Puram. This model could be helpful in forecasting the concentration of PM10 and PM2.5 so that levels of these pollutants could be kept in check.*

Keywords- Air pollution, Machine learning, PM10, PM2.5, Multiple linear regression (MLR)

I. INTRODUCTION

Air pollution is a worldwide problem affecting both environment and health of human beings. Mixture of solid and liquid particles suspended in air is called particulate matter (PM). PM10 and PM2.5 are dangerous for human health because of their small particle size [1]. PM2.5 can also move into bloodstream resulting into cardiovascular problems like stroke and heart diseases and thus PM2.5 is called as "The Invisible Killer"[2-3].Factors responsible for production of particulate matter are emissions from vehicles, combustion, smoke produces by forest fires, burning of agricultural crops and various industrial processes[4]. Various air pollution monitoring stations have been established to record concentration of pollutants and also give details regarding ambient air quality.

Computational method like machine learning can be applied on large datasets to analyse data, extract hidden information, discover new patterns and develop models to predict future outcome. Machine learning is a subdivision of artificial intelligence (AI) which utilises mathematical and statistical models for forecasting predictions. Machine learning techniques can be supervised learning, unsupervised learning and reinforcement learning. In supervised learning, dataset contains both input and output values. Thus, it is trained and then tested for the data, whereas in unsupervised learning only inputs are available in the dataset and the machine has to explore the data on its own and find patterns and predictions. Using various algorithms, computer is trained and made to understand with inputs of datasets and then can give predictions or results based on its understanding

*Corresponding Author

[5-6]. Depending upon whether classification, correlation, prediction or pattern recognition is required, a machine learning model is developed using algorithms on training set. Then the developed model is applied on testing set and accuracy and performance is analysed. Machine learning shows application in healthcare sector through prediction of diseases like cancer, diabetes, liver disorders [7-8]. Machine learning can also be applied to forecast pollutant concentration, prediction in variations of PM2.5 concentration based on hourly data of PM2.5, air quality assessment and prediction [9-12].

To predict concentration of air pollutant, meteorological data is also required. By predicting concentration of air pollutants, information can be used to develop strategies to reduce emission of pollutants. We have developed a model for prediction of PM10 and PM2.5 using machine learning for air pollution dataset of three stations of Delhi: Anand Vihar, Punjabi Bagh and R.K. Puram using multiple linear regression (MLR).

II. MATERIALS

A. DATA SOURCE

Air pollution dataset for four years (January 2015-December 2018) has been obtained from Central Pollution Control Board (CPCB) for three air pollution monitoring stations in Delhi namely Anand Vihar, R.K. Puram and Punjabi Bagh. The site can be accessed at http://www.cpcb.nic.

B. DATASET DESCRIPTION

Using CPCB Real time Air Quality data monitoring application, data for three air pollution monitoring stations has been collected for four years. Anand Vihar had 1079 observations, Punjabi Bagh had 1124 observations and R.K. Puram had 1294 observations. The three air pollution monitoring stations showed two major air pollutants like particulate matter PM2.5 and PM10 and six meteorological parameters such as relative humidity (RH), solar radiation (SR), ambient temperature (AT), bar pressure (BP), wind speed (WS) and wind direction (WD).

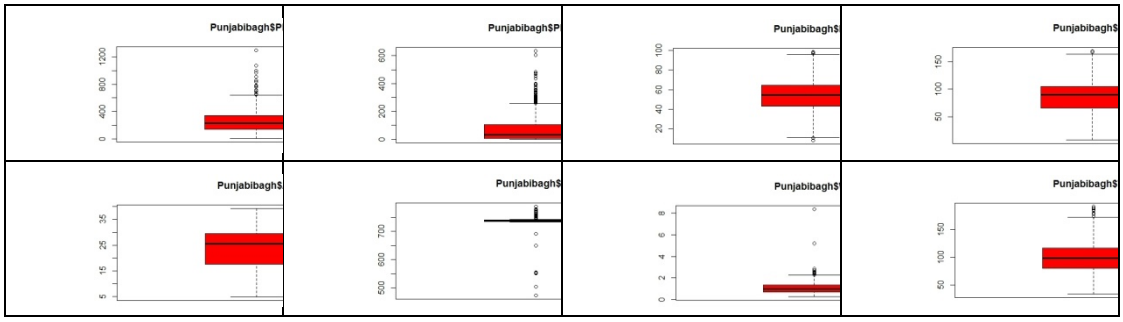
C. STUDY AREAS

Anand Vihar locality is a densely populated area, located in trans-Yamuna region of New Delhi. It is associated with providing transportation services due to presence of bus terminal (Anand Vihar ISBT) and railway station (Anand Vihar Terminal). It is always occupied by automobiles, causing congestion on roads and contributing to air pollution in large amount by releasing harmful gases from automobiles. R.K. Puram is a residential colony in South West Delhi for Central Government Employees, but as it is sandwiched between Ring Road and Outer Ring Road, it is always busy whether day or night. In night, the roads get occupied by heavy vehicles like trucks which are transporting goods from one place to other and commercial vehicles like taxis, thus releasing harmful gases in environment. Punjabi Bagh is a residential locality located in west side of Delhi. In past years lot of commercial growth has taken place due to which population has increased in this locality. As population increases, number of vehicles also increases thus contributing to air pollution.

III. METHODOLOGY

The air pollution dataset for all the three monitoring stations were imported to RStudio (version 3.5.2) data manipulation tool. Next, the data was pre-processed, in this pre-processing of air pollution dataset involved feature selection, imputation of missing values and filtering of data to remove outliers. Missing values (NAs) in dataset makes model less effective. These missing values were imputed using *k*-NN (*k*-Nearest Neighbour) imputation. Outliers were studied through boxplots (Figure1.) carefully and removed by selecting appropriate range for each variable. Multiple linear regression (MLR) is used to develop a two sets of models each for PM10 and PM2.5 to predict

concentration of air pollutants PM10 and PM2.5. This algorithm is used as concentration of air pollutants are not only affected by emission sources but also by various meteorological parameters which disperses these pollutants. It tells which parameters significantly affect concentration of PM10 and PM2.5. For this purpose dataset of three different stations of Delhi-Anand Vihar, Punjabi Bagh and R.K. Puram are used.



Fig,1. *Boxplots of air pollutants and meteorological data showing outliers for area Punjabi Bagh*

A. MULTIPLE LINEAR REGRESSION

Multiple Linear Regression is a statistical technique that is used to study the influence of independent variables (predictor variables) on a dependent variable (response variable). For a given dataset with response variable as Y and predictor variables as X_1, X_2,…, X_k the equation of multiple linear regression that fits the dataset is

$$\mathbf{Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots\ldots\ldots\ldots\ldots\ldots + b_k X_k}$$

where b_0 is the intercept and b_1, b_2, b_k values are called as regression coefficients which tells about association between each predictor variable and the response variable. Multiple linear regression can be used to determine the impact of various independent variables on a dependent variable, by what amount each independent variable can bring variation on the response variable and values of dependent variable can be predicted for future. To analyse the results, we check coefficient of determination (R^2). R^2 is the proportion of the variance explained by the model. It takes into account that all the independent variables are going to affect the dependent variable even if they are not showing significance impact. R^2 value can range from zero to one. So it's better to check adjusted R^2 value as it explains variations only based on the independent variables which actually show significant impact on dependent variable. Thus, adjusted R^2 can be used to judge goodness of model. Multiple linear regression has been used to study quality of water [13-14], malaria mapping through GIS and remote sensing [15].

B. MODEL DEVELOPMENT

To develop a predictive model (Figure2.) for concentration of air pollutants PM10 and PM2.5 using MLR , pre-processing of raw data was done for the three areas of Delhi: Anand Vihar, Punjabi Bagh and R.K. Puram, which involved imputation of missing values by *k*-NN imputation for meteorological parameters (SR, AT, RH, BP, WS and WD) and removal of outliers for air pollutants (PM10 and PM2.5) and meteorological parameters (SR, AT, RH, BP, WS and WD). Two models each for PM10 and PM2.5 were developed. For PM10, the first model developed using multiple linear regression contained predictor variables as RH, SR, AT, BP, WS and WD. The second model contained predictor variables as RH, SR, AT, BP, WS, WD and previous day's PM10 concentration. For PM2.5, the first model developed using multiple linear regression contained predictor variables as RH, SR, AT, BP, WS and WD. The second model contained predictor variables as RH, SR, AT, BP,

WS, WD and previous day's PM2.5 concentration. The model was trained on 70% training data and tested on 30% remaining data.

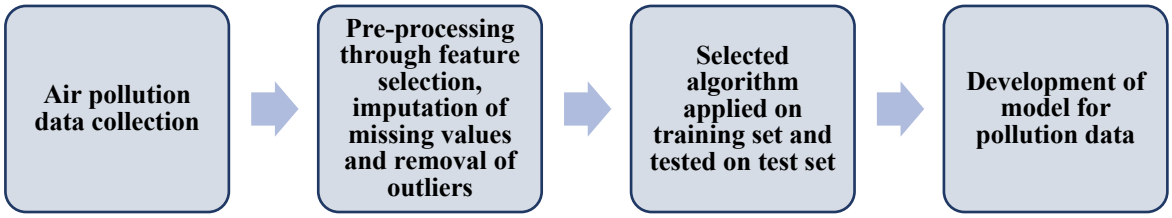


***Fig.2**. Steps of model development*

IV. RESULTS AND DISCUSSION

Predictive models were developed for PM10 and PM2.5 using multiple linear regression algorithm in "RStudio". In case of Anand Vihar, for PM10 when only meteorological parameters (AT, SR, BP, RH, WS and WD) were taken, it gave adjusted R^2 value 0.35. The adjusted R^2 value increased to 0.63 when previous day's PM10 concentration was added with meteorological parameters. For PM2.5 when only meteorological parameters (AT, SR, BP, RH, WS and WD) were taken, it gave adjusted R^2 value 0.45. The adjusted R^2 value increased to 0.71when previous day's PM2.5 concentration was added with meteorological parameters. In case of Punjabi Bagh, for PM10 when only meteorological parameters (AT, SR, BP, RH, WS and WD) were taken, it gave adjusted R^2 value 0.40. The adjusted R^2 value increased to 0.71 when previous day's PM10 concentration was added with meteorological parameters. For PM2.5 when only meteorological parameters (AT, SR, BP, RH, WS and WD) were taken, it gave adjusted R^2 value 0.34. The adjusted R^2 value increased to 0.73 when previous day's PM2.5 concentration was added with meteorological parameters. In case of R.K. Puram, for PM10 when only meteorological parameters (AT, SR, BP, RH, WS and WD) were taken, it gave adjusted R^2 value 0.50. The adjusted R^2 value increased to 0.74 when previous day's PM10 concentration was added with meteorological parameters. For PM2.5 when only meteorological parameters (AT, SR, BP, RH, WS and WD) were taken, it gave adjusted R^2 value 0.37. The adjusted R^2 value increased to 0.82 when previous day's PM2.5 concentration was added with meteorological parameters.

In case of Anand Vihar, multiple linear regression explained 63% variations in PM10 and 71% variations in PM2.5. In case of Punjabi Bagh, multiple linear regression explained 71% variations in PM10 and 73% variations in PM2.5. In case of R.K. Puram, multiple linear regression explained 74% variations in PM10 and 82% variations in PM2.5. Thus, for all the three areas of Delhi, PM10 and PM2.5 concentration was not explained adequately by meteorological parameters like relative humidity (RH), solar radiation (SR), bar pressure (BP), ambient temperature (AT), wind speed (WS) and wind direction (WD). But when previous day's concentration of PM10 and PM2.5 were added with meteorological parameters, R^2 value increased indicating that they show significant impact on determining PM10 and PM2.5 concentration.

The performance of the models developed for PM10 and PM2.5 for areas Anand Vihar, Punjabi Bagh and R.K. Puram are represented through model fit graphs (figure 3-5).They show the best models developed for PM10 and PM2.5 when meteorological parameters and previous day's pollutant concentration are taken as predictor variables. They show good relation between real values and predicted values.

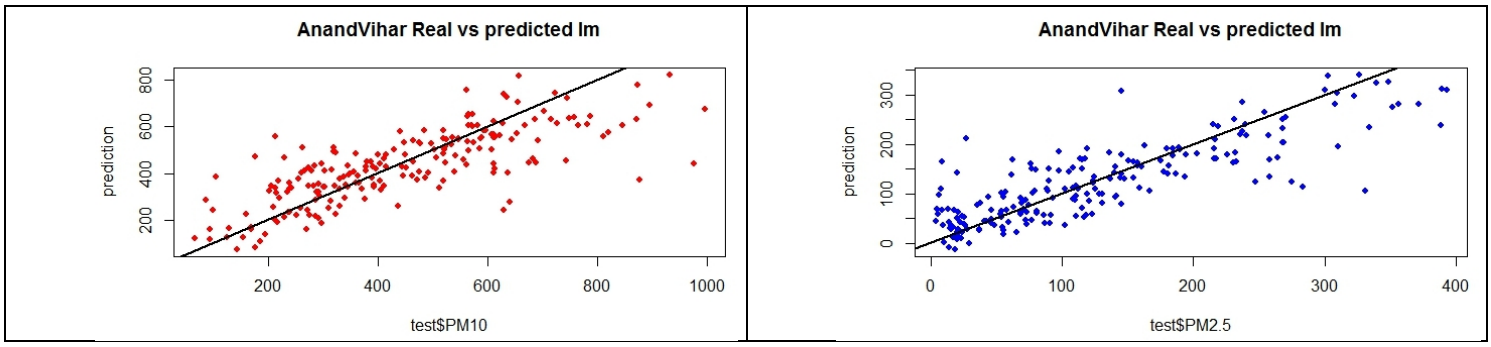


Fig.3. *Concentration of PM10 and PM2.5 actual versus prediction for Anand Vihar using multiple linear regression when all meteorological parameters and previous day's pollutant concentration are taken as predictor variables*

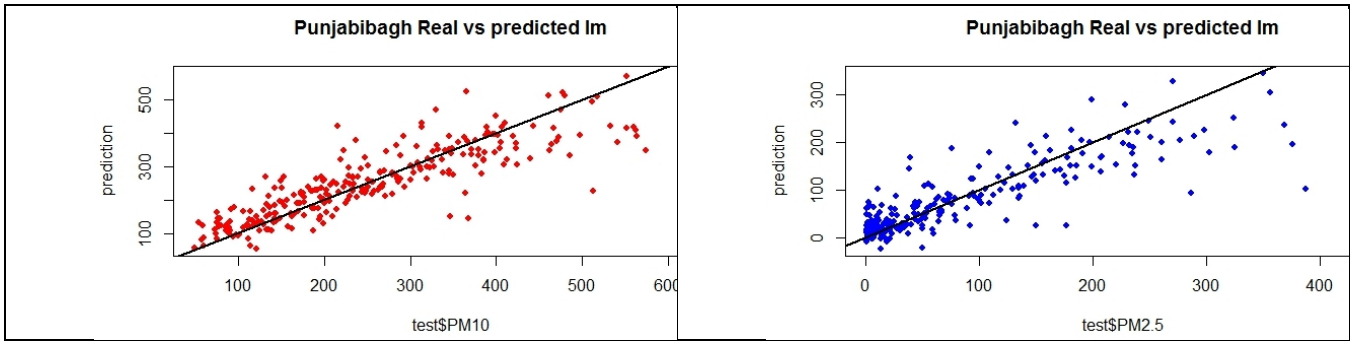


Fig.4. Concentration of PM10 and PM2.5 actual versus predicted for Punjabi Bagh using multiple linear regression when all meteorological parameters and previous day's pollutant concentration are taken as predictor variables

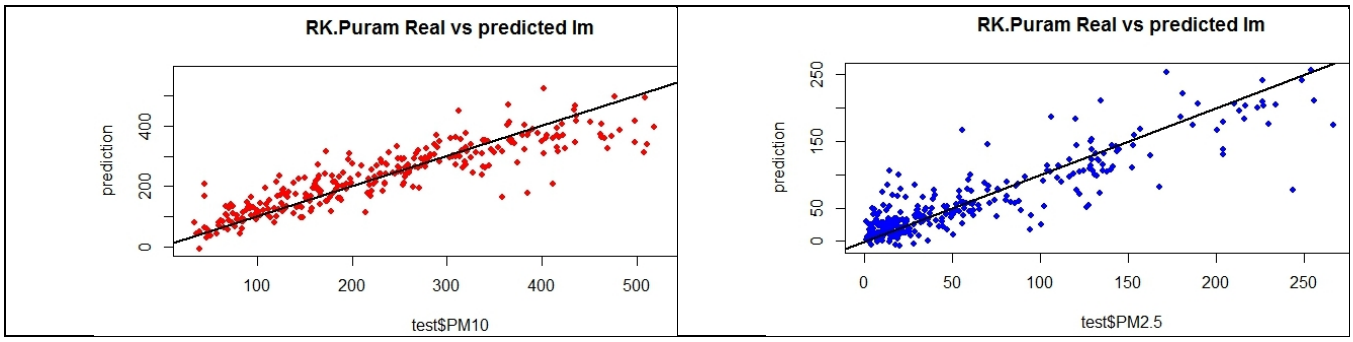


Fig.5 Concentration of PM10 and PM2.5 actual versus predicted for R.K. Puram using multiple linear regression when all meteorological parameters and previous day's pollutant concentration are taken as predictor variables

V. CONCLUSION

Air pollution dataset for three areas- Anand Vihar, Punjabi Bagh and R.K. Puram was collected and model to predict air pollutants PM10 and PM2.5 concentration was developed using multiple linear regression. The meteorological parameters like relative humidity, solar radiation, bar pressure, ambient temperature, wind speed and wind direction were used as predictors to develop model. These models showed less variance based on these parameters but when previous day's pollutant concentration was added to the model, R^2 increased significantly indicating that previous day's pollutant concentration plays important role in forecasting pollutant concentration. Thus, 63%, 71% and 74%

variance in PM10 concentration for areas Anand Vihar , Punjabi Bagh and R.K. Puram was developed using multiple linear regression. 71%, 73% and 82% variance in PM2.5 concentration for areas Anand Vihar , Punjabi Bagh and R.K. Puram was developed using multiple linear regression.

REFERENCES

[1] Dimitriou, K., Kassomenos, P.A., Paschalidou, A.K.: Assessing air quality with regards to its effect on human health in the European Union through air quality indices. Ecol. Indic 27, 108–115 (2013).

[2] Wolf K, Stafoggia M, Cesaroni G, et al.: Long-term Exposure to Particulate Matter Constituents and the Incidence of Coronary Events: In 11 European Cohorts. Epidemiology (26), 565-74 (2015).

[3] Habre R., Coull B., Moshier E., Godbold J., Grunin A., Nath A., et al: Sources of indoor air pollution in New York city residences of asthmatic children. J. Expo Sci Environ Epidemiology (24), 269-78 (2014).

[4] Morawska L., Thomas S., Bofinger N. D., Wainwright D., Neale D. Comprehensive characterisation of aerosols in a subtropical urban atmosphere: particle size distribution and correlation with gaseous pollutants. Atmos Environ 32 (14-15), 2461–2478 (1998).

[5] Shiliang S.: A survey of multi-view machine learning, Neural Comput. Applic (23), 7–8 (2013).

[6] Jordan M.I., Mitchell M.: Machine learning: trends, perspectives, and prospects. Science 349 (6245), 255–260 (2015).

[7] Hwang, Kyu-Baek, et al.: Applying machine learning techniques to analysis of gene expression data: cancer diagnosis. Methods of Microarray Data Analysis. Springer US, 167-182(2002).

[8] Hashem E.M. and Mabrouk M.S.: A study of support vector machine algorithm for liver disease diagnosi. Am. J. Intell. Sys.4 (1), 9–14 (2014).

[9] Daniel P. C., Jeffrey A. W., Stephen E. W., Robert M.S., Richard A. B.: The Steubenville Comprehensive Air Monitoring Program(SCAMP): Analysis of Short-Term and Episodic Variations inPM2.5 Concentrations Using Hourly Air Monitoring Data. J. the Air & Waste Manag. Asso. (55), 559–573(2011).

[10] Sarah N. K., Alaa M. El-Halees.: Implementation of Data Mining Techniques for Meteorological Data Analysis. International Journal of Information and Communication Technology Research 1(3), 96-100, (2011).

[11] Pandey G., Bin Z. and Le J.: Predicting sub-micron air pollution indicators: a machine learning approach.; Environmental Science Processes Impacts; 15(5) 996-1005 (2013).

[12] Caselli M., Trizio L., de Gennaro G., Ielpo P.: A Simple Feedforward Neural Network for the PM10 Forecasting: Comparison with a Radial Basis Function Network and a Multivariate Linear Regression Model. Water Air Soil Pollution (201), 365–377(2009).

[13] Adamu M. and Ado A.: Application of Principal Component Analysis and Multiple Regression Models in Surface Water Quality Assessment. J. Env. and Earth Sci. 2(2) 16-23 (2012).

[14] Aziz R. and Muhammad O.C.: Regional interpretation of river Indus water quality data using regression model. Afri. J. Envi. Sci. and Tech. 8(1) 86-90 (2014).

[15] Ra P. K., Nathawat M. S., Onagh M.: Application of Multiple Linear Regression Model through GIS and Remote Sensing for Malaria Mapping in Varanasi District. Health Sci. J. 6(4)731-749(2012).

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

A Dwarf Survey on ASR (Automatic Speech Recognition) in Noisy Environment

Syed Sibtain Khalid *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
sibtain1977@gmail.com

Safdar Tanweer
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
safdartanweer@yahoo.com

Naseem Rao
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Bhavya Alankar
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract - *These Speech is very special and comfortable means of communication for humans from the very ancient time to till date. It is tried to implement the same way of communication with machines. But there are various issues in this communication, e.g. ascent, language, race, sex, age, health and many more. It is not the only problem that exist in the ASR but one of the major problems is to recognize the speech in presence of environmental noise. There is drastic improvement in ASR due to rigorous effort by researchers in this field. But still ASR in noisy environment is a big challenge to researchers. In this paper we have tried to present the development of the ASR in noisy environment in recent past.*

Keywords - *ASR, Environmental Noise, Feature Extraction Techniques, Noise Classifier*

I. INTRODUCTION

The presence of environmental generated noise in background of speech signal severely change the recovery process and make them difficult to complex. To retrieve the originality of information to be recognized, most of the environmental born noise isolation are done with the digitally empowered software and hardware. During this dwarf survey the noise considered for studies are traffic noise, industry generated noise, rail and market noise etc. [1]. The instruments so far available to measure the intensity of such class of noise not much capable to filter them up to mark. Hence various classification schemes are lodged to measure their percentage proportion in the environment. The popular classifiers whose performance for classification of such class of noise are LDA, QDA, SVM, KNN, GMM, HMM, ANN etc.

*Corresponding Author

In the speech feature analysis to know the important feature of speech through my wide review it is summarized that LPC. MFCC, RASTA filtering, PLP, MP are few whose performance are moderate to excellent depending upon the percentage of background noise and the kind of speech spectrum. We have found that ANN and KNN performance seems to be better for ASR in noisy environment [2]. On the other hand, MFCC results are better for feature selection in ASR.

II. NOISE CLASSIFIERS

The purpose of the noise classifier is to estimate and analyze the background noise available in speech signals. It reveals the contamination caused due to environmental noise in the speech. Some of the researcher works on the noise samples data. These results shows a significant accuracy by the selected feature are classified on the basis of their application and use to find the percentage of background noise few of them are listed below [3,4,5]:

A. LDA (Linear Discriminant Analysis) and QDA (Quadratic Discriminat Analysis)

It is one of the statistical ways to classify speech in exhaustive and exclusive group proposed by fisher's. it is widely accepted in machine and pattern recognition field. It is capable to differentiate various features on linear scale. However, in case of QDA[1,5], the boundaries to differentiate the features are determined by application of quadratic function.

B. SVM (Support Vector Machine)

This method is beauty to separate input speech vector into two separate location using decision linear boundary. The element location in the separated boundary greatly varied according to the data set they are been classified [6].

C. ANN (Artificial Neural Network)

It is comprised of cluster of neurons. The cluster of these neurons are used for sharing information by forming their network. All such network have divided into input, hidden, output layers[7]. The information data is asserted to the input which is processed in hidden layer and the result are noticed at output layer. Such network have very important adaptive learning ability as per they are been trained.

D. GMM (Gaussian Mixture Model) & HMM (Hidden Markov Model)

GMM is the widely accepted methods for data classification using PDF (probability distribution function) of weighted Gaussians. HMM is popularly known as predictive model, that analyses the previous history of situation happens in the past on the basis of that, it is modeled time series data and establishes transition probability from current state using previous state [8].

III. FEATURE EXTRACTION TECHNIQUES

There are various methods for extracting features. Few of the efficient feature extraction techniques in the field of ASR are STSA, FBA, STFA, LPC, PLP, MFCC, MP, Combined LPC and MFCC etc. The dominant features are extracted in the form of parameterized data by using different feature extraction techniques [9]. The universal feature of speech can be identified as energy, phone, pitch, spectral identity etc. Sometimes the speech has some kind of environmental noise in the background which is spread over entire spectrum of speech. Therefore, the preservation of speech feature which is parametrized using feature extraction techniques is difficult and would not get the recognition of speech properly. So, ASR in noisy environment using parametrization needs to be some more robust techniques for ASR process. Two or more techniques for feature extraction techniques are used to get better. The important feature extraction which are also widely accepted are as follows [10]:

- Short Time Speech Analysis (STSA)
- Filter Bank Analysis (FBA)

• Short Time Fourier Analysis (STFA)

• Linear Predictive Coding (LPC)

• Perceptual Linear Predictor (PLP)

• RASTA PLP

• Mel Frequency Cepstral Coefficient (MFCC)

IV. CONCLUSION

This paper is about the brief review of ASR in the presence of environmental noise. We have discussed various environmental noise present in the background which contaminated the original speech and causes to take certain measure to estimate the effect of noise. We have seen and compare the different techniques to classify the environmental noise. It is evident from many researchers that ANN, KNN and HMM are the most favorable to carry out the above said task. The noise classification is the key thrust in estimating the environmental noise.

REFERENCES

[1] Safdar Tanweer , Abdul Mobin and Afshar Alam, "Environmental Noise Classification using LDA, QDA and ANN Methods", Indian Journal of Science and Technology, Vol 9(33), DOI: 10.17485/ijst/2016/v9i33/95628, September 2016, ISSN (Print) : 0974-6846, ISSN (Online) : 0974-5645

[2] Chu S, Narayanan S, Kuo CJ. "Environmental sound rec¬ognition with time frequency audio features", IEEE Trans¬action on Audio, Speech and Language Processing. 2009; 17(6):1142–58.

[3].J.-S. Lienard "Speech analysis and reconstruction using short-time, elementary waveforms", Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '87.

[4].P.Suresh Babu, Dr.D.Srinivasulu Reddy, Dr.P.V.N.Reddy, "Speech Signal Analysis Using Windowing Techniques", International Journal of Emerging Trends in Engineering Research (IJETER), Vol. 3 No.6, Pages : 257- 263 (2015) Special Issue of NCTET 2K15, ISSN 2347 – 3983.

[5].R. W. Schafer , L. R. Rabiner, "Design of digital filter banks for speech analysis", The Bell System Technical Journal (Volume: 50, Issue: 10, Dec. 1971), Print ISSN: 0005-8580, Page(s): 3097 – 3115,

[6].M. Hosseinpour ; M.N. Ranjbar ; M. Mousavinejad, "Speech Unit Selection Based on Matching Pursuit", Computer Systems and Applications, 2007. AICCSA '07. IEEE/ACS International Conference on SBN Information: Print ISBN: 1-4244-1030-4, CD-ROM ISBN: 1-4244-1031-2, ISSN Information: Print ISSN: 2161-5322, Electronic ISSN: 2161-5330

[7]. Syed Sibtain Khalid, Safdar Tanweer, Dr. Abdul Mobin, Dr. Afshar Alam, "A comparative Performance Analysis of LPC and MFCC for Noise Estimation in Speech Recognition Task ", International Journal of Electronics Engineering Research. ISSN 0975-6450 Volume 9, Number 3 (2017) pp. 377-390 © Research India Publications http://www.ripublication.com,

[8]. M. Portnoff, "Short-time Fourier analysis of sampled speech", IEEE Transactions on Acoustics, Speech, and Signal Processing (Volume: 29, Issue: 3, Jun 1981).

[9].Safdar Tanweer, Abdul Mobin, Afshar Alam, "Analysis of Combined Use of NN and MFCC for Speech Recognition", World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering Vol:8, No:9, 2014.

[10].Hynek Hermansky, "Perceptual linear predictive (PLP) analysis of speech", The Journal of the Acoustical Society of America 87, 1738 (1990); https://doi.org/10.1121/1.399423, Volume 87, Issue 4.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Towards Determining the Legal Status and Sustainability of Intelligent Agents

Nafisur Rahman *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
nafiis@gmail.com

Md. Onais Ahmad
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
oahmad@jamiahamdard.ac.in

Syed Imtiyaz Hassan
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
s.imtiyaz@gmail.com

Abstract – *In this paper we have presented the fundamental questions revolving around the legal status of Intelligent Agents. After a brief introduction, we have summarized the definition of Intelligent Agents followed by an account of the ideas of legal recognition of Artificial Agents and the questions surrounding them. We have highlighted the importance of figuring out if it is actually necessary to decide the legal status of robots and if it is viable to do so. After that, we have argued that until the legal status of Intelligent Agents is decided and settled, their sustainability will always be questioned. We have concluded that a collaboration of researchers from various fields viz. Computer Science, Engineering, Law, etc. is needed to find the answers surrounding this debate.*

Keywords *– Intelligent Agents, Autonomous Machines, Robots, Legal Status, Sustainability.*

I. INTRODUCTION

Artificial Intelligence [1] is the intelligence demonstrated by machines, commonly known as Intelligent Agents, by thinking and acting rationally as humans tend to do. For this, they make use of various techniques to reason like humans and take decisions accordingly. But the more they resemble humans, the more do they pose a problem of having their legal status determined. As humans are conferred with certain rights and they have certain obligations, should the Intelligent Agents be conferred with similar rights and should similar duties be attached to them? The answer to this question needs serious academic debates and discussions. The opinion of the research community is divided and the theoretical foundations for the outcomes of one or the other extreme are still developing.

*Corresponding Author

II. Intelligent Agents

Intelligent Agents [1] are the entities that sense an environment, take decisions, and act upon them affecting the environment in order to achieve some goals. The terms viz. Autonomous Machines [2], Smart Machines, and Robots may also be used to refer to the Intelligent Agents. The examples may include self-driving vehicles, smart weapons, and sophisticated humanoid robots. Although created by humans, when left to deal with an environment, the intelligent agents are not controlled by them. They make their own decisions based on the inputs they receive through the sensors and accordingly, they take actions that they find appropriate in the light of the knowledge base that they have. Their knowledge base may evolve over time and may depend on the outcome of their actions. This makes them learn and influences their future decisions and actions. In many ways, this is similar to how humans sense, act, and learn.

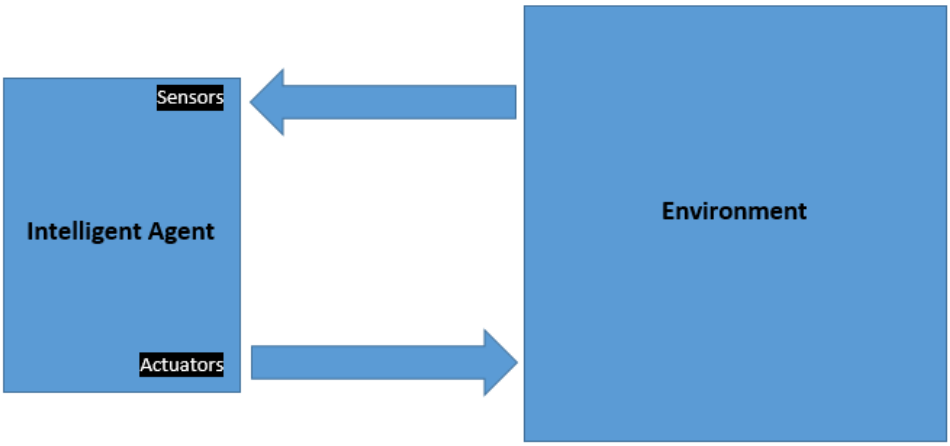


Fig. 1 Intelligent Agents

Fig. 1 describes how an Intelligent Agent takes inputs from the environment through sensors and gives responses through actuators.

III. Legal Status of Intelligent Agents

By the Legal Status of an entity, we refer to how it is regarded by law. In the context of Intelligent Agents, the discussion primarily revolves around whether to grant legal personhood [3] to these electronic persons or not and the consequences of doing so. The ideas of granting legal personhood and not doing so have been doing the rounds for quite some time but any discussion on the topic is neither complete nor conclusive. For example, on one hand, Fischer and Ravizza [4] have argued against granting legal personhood to Intelligent Agents citing absence of intention, will, and consciousness in them. Hage [5], on the other hand, has strongly refuted the established notions of free will and intentionality and has argued that Intelligent Agents can be considered as legal persons just like humans. Most of the discussions on the subject deal with the possibility [6, 7], or the lack of it, of granting legal personhood to Intelligent Agents. But we view it slightly differently. Any attempt to grant legal personhood to Intelligent Agents or Robots should answer two fundamental questions:

1. Is it necessary to do so?
2. Is it practically feasible?

Academics and researchers have varying viewpoints [5, 8] regarding these and until this debate settles down, any decision, by the legal authorities, for or against conferring legal personhood to Intelligent Agents will invite a lot of criticism.

IV. Sustainability of Intelligent Agents

Determination of legal status of Intelligent Agents or Robots, i.e. whether they can be considered as legal persons [3, 9] or not, will have far reaching effects. If it is decided that they can be considered as legal persons, then their rights and obligations will dictate future course of action relating to their involvement in smart and sustainable development. If their rights and obligations are settled, scientists and manufacturers will have a greater clarity regarding the capabilities they intend to provide in the robots and the application areas or environments where they should operate autonomously. On the contrary if it is decided that Intelligent Agents cannot be legal persons, the very idea of their manufacturing and usage will have to be questioned. Because of many grey areas, despite the availability of tools and technologies, it may be argued that it is wiser not to employ them in making Intelligent Agents and this would not be unprecedented, the past example being the ban on human cloning. Thus, sustainability of Intelligent Agents will greatly be affected by their legal status.

V. Conclusion

The above discussion on the legal status and sustainability of Intelligent Agents highlights the need for a thorough academic debate on the subject and calls for meaningful research collaborations among computer scientists, engineers, legal experts, and people's representatives. We intend to take the discourse forward by working out some of these collaborations in our future work.

References

[1]. Russell, S. and Norvig, P. Artificial intelligence: a modern approach. Englewood Cliffs: PrenticeHall, 1995.

[2]. Franklin S, Graesser A (1997) Is it an agent, or just a program?: a taxonomy for autonomous agents. In: Mu¨ller JP, Wooldridge MJ, Jennings NR (eds) Intelligent agents III agent theories, architectures, and languages. ATAL 1996. Lecture Notes in Computer Science (Lecture Notes in Artificial Intelligence), vol 1193. Springer, Berlin.

[3]. Smith B (1928) Legal personality. Yale Law J 37(3):283–299.

[4]. Fischer JM, Ravizza M (2000) Responsibility and control: a theory of moral responsibility. Cambridge University Press, Cambridge.

[5]. Hage, J. Artif Intell Law (2017) 25: 255. https://doi.org/10.1007/s10506-017-9208-7.

[6]. Brożek, B. & Jakubiec, M. Artif Intell Law (2017) 25: 293. https://doi.org/10.1007/s10506-017-9207-8.

[7]. Bryson, J.J., Diamantis, M.E. & Grant, T.D. Artif Intell Law (2017) 25: 273. https://doi.org/10.1007/s10506-017-9214-9.

[8]. Chen, J. & Burgess, P. Artif Intell Law (2019) 27: 73. https://doi.org/10.1007/s10506-018-9229-x.

[9]. Solaiman SM (2017) Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy. Artif Intell Law 25(2):155–179.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Optimizing Learning using Student Activity Tracking Data with Machine Learning Techniques

Anam Saiyeda *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
anamsaiyeda@jamiahamdard.ac.in

Sheikh Mohammad Idrees
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

M. Afshar Alam
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract - *With technology being omnipresent and the increased availability of the internet, activity tracking task produces huge amount of a variety of data. Ranging from wearable technology, to GPS, sensor data or apps activity, this data can be applicable for an assortment of functions from improving health to recommendation systems. One application of this is in the sphere of learning analytics as well as in educational data miming where this data can be utilized for optimizing learning and teaching techniques. This study collects and summarizes the various ways in which student activities can be tracked and the data used for applying machine learning techniques and give useful insights to the educators. We try to identify how learning analytics has been employed in the area of education. Further we study its expected benefits for not just students and teachers but also for education institutions. Previous research and studies were collected and classified. The details of the studies were tabulated on the basis of their objectives, techniques and approaches. Use of technology and the fields of machine learning, artificial intelligence etc in education and learning will lead to improved education and ranking of institute.*

Keywords - e-learning, m-learning, learning analytics, educational data mining, activity tracking

---

I. INTRODUCTION

Technology has invaded every aspect of life. With the growing popularity of smart phones and smart wear activity tracking has become very common. Activity tracking is not limited to just wearable devices. Activity tracking can include tracking everyday activities through sensors or wearable devices, tracking activities of individuals through their mobile phones activities, tracking activities on their systems, laptops, collecting logs of certain applications and also tracking people on social media. With advances in hardware nowadays large storage is available and data processing is easier at lesser cost. Data has

*Corresponding Author

become the most valued entity. Big data, data analytics, machine learning and AI are important as they help us make sense of the large amount of data generated and utilize it for the betterment if the society. Similarly in the field of education, this data collection can be used for improving learning of students, improve teaching standards, and make the environment more interactive which further improves the functioning and ranking of institutes and also improves the education quality if the country. New models of learning like b-learning and e-learning are also available now for the students and teachers. Online and offline learning environments, social media, tracking lead to the production of large amounts of statistics related to education, learning and teaching processes. Valuable information, patterns can be extracted and this information can be utilized to improve students' performance. For this two fields are emerging in the education sector i.e. educational data mining and the other is the area of learning analytics. This study will be propitious for educational organizations and academics. It will capacitate them to be up to the minute with technology in these spheres. These emerging fields have a lot of potential for further exploration of this area and educators need to be kept apprised of the developments in these fields.

A. LEARNING ANALYTICS

Learning analytics aims to improve and optimize learning by analyzing data, collected from learners, educators, learning experiences, and programs. LA aims at understanding the whole system and focuses on leveraging human judgment. It informs and empowers students and instructors. It focuses on analysis, learner success prediction, sentiment analysis, concept analysis, discourse analysis, influence analysis and sense-making models [1].

Learning analytics can be generally categorized as learner analytics, learning experience analytics, and learning program analytics. Learning experience analytics focuses on a specific learning activity and understanding about it in detail. Usage patterns for a specific activity are observed like the duration of use, time of access, how much or the frequency of use, resources used the most, topics searched frequently. Learner analytics intends to comprehend further about a specific person or group of people occupied in activities. In this case learning is one of the outputs. It incorporates queries correlated to usage patterns, performance for explicit learners, high-potential employees, skills and aptitude of person/groups, topics of interest of the learners. Learning program analytics focuses on performance of an overall learning program. A learning program may encompass several learners, and learning experiences. It answers questions like behaviour differences of learner before and after training, improvement of organizational performance because of learning and the most effective learning methodology. [17]

Learning analytics helps in evaluation of the effectiveness of teaching for improvement. It helps monitoring of students' learning, find patterns, predict students' performance, and detect undesirable learning behaviour. This can help us identify students at risk, predict dropout rates, take prompt follow-up action and provide such students help and timely assistance. The learning experiences can be optimized, using this and judicious and astute observations from data can be made about the teaching effectiveness. Learning characteristics and patterns can be utilised to make it more interactive, distinctive, personal and agreeable. This will lead to overall improvement of not just students but also instructors and the universities on whole.

B. EDUCATIONAL DATA MINING

Educational Data Mining (EDM) is an upcoming interdisciplinary research area that incorporates data mining in the field of education. It engages with the utilisation of methods to traverses data stemming from an educational context. It studies educational data

generated by students and instructors using data-mining, machine-learning and statistical methods. It helps in optimization of learning processes. EDM uses various computational techniques and methods for analysis of educational data in order to study educational questions. Baker et al. [27] created 4 groups of EDM applications. The first one i.e. Student modelling designs customized learning processes using student data and EDM techniques. The second model does modelling of the knowledge organization of the realm. It uses techniques which coalesce psychometric modelling frameworks with space-searching algorithms. The fourth one is scientific research where applications aid in development and testing of educational scientific theories and in formulation of new hypotheses.

A wide variety of EDM techniques are present. Some are analogous to those used in other domains by means of data mining, while some are exclusive to educational data mining. Some of the major categories of methods frequently used in EDM are Prediction, Structure Discovery, Relationship Mining, and Discovery with Models. This is not an exhaustive collection of EDM methods.

II.RESEARCH METHOD

A. NEED FOR REVIEW

Kitchenham[28] proposes that a comprehensive review aids in finding the gaps in existing methodologies and to place the hypothesis in proper perspective. This section encapsulates the existing evidences concerning the machine learning techniques used for improving learning by using data generated from tracking applications. The impetus of the review is to classify the studies on the title. The work considers many primary and secondary studies on the topic.

B. SOURCES OF INFORMATION

In order to carry out the review quality databases were searched. The filtering was done as per the guidelines decide in advance regarding the purpose of this paper. The papers were obtained from the databases of ACM, IEEE, Springer and Science Direct.

C. LITERATURE SEARCH STRATEGY

The search convention included the papers associated with student activity tracking, learning analytic and educational data mining. The keyword <Student activity tracking> was used. The search was further refined with the keywords <Computer Science> and <learning analytics> and < educational data mining >. Irrelevant papers related to other disciplines were filtered out and the results are summarized below. Only papers from the last 5 years were considered i.e. from the year 2014 to 2018.

Search with the keyword <Student activity tracking> yielded 24,780 results in the Springer database. The papers were ordered on the basis of relevance. After refining the search for the specified duration 13,578 results were obtained. Refining the search and including only the papers from the field of computer science 1646 results were obtained. These contained 763 Chapters, 2416 articles and 623 conference papers. In the years 2014-18 within CS results were obtained

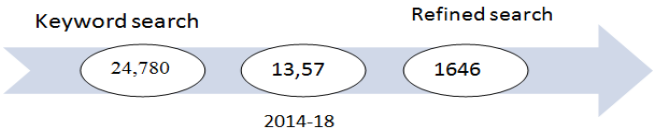


Fig.1: Selection of papers from Springer Database

In the ACM digital library the keyword <Student activity tracking> gave 75,200 results. The search was refined for the desired duration and 27,429 results were obtained. On the basis of abstracts the irrelevant results were filtered out.

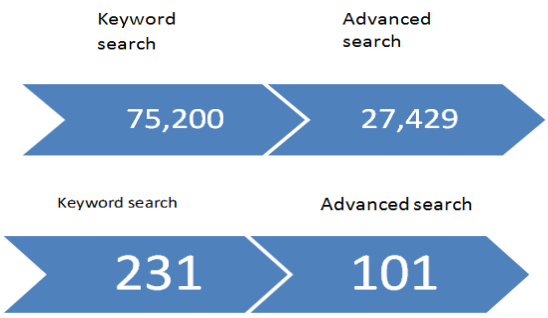


Fig 2: Selection of papers from the ACM database & IEEE database

In the IEEE database the keyword <Student activity tracking> gave 231results. The search was refined for the desired duration and 101results were obtained. On the basis of abstracts the irrelevant results were filtered out.

In Science Direct database search with the keyword <Student activity tracking> yielded 85,909 results. After refining the search for the specified duration 30,210 results were obtained. Refining the search and including only the papers from the field of computer science 1,263 results were obtained. These contained 261 articles, 1 review article, 5 editorials and 11 others.

Fig 5: Selection of papers from the Science Direct database

D. STUDY SELECTION

Initial filtering of papers was done on the basis of titles. Then on the basis of keywords further selection was done. The abstract of the remaining papers were analyzed. After reading the complete text the final set of papers was finalized.

E. DATA EXTRACTION

Data collection was done according to the Kitchenham guidelines. The quality of the papers was the prime benchmark for paper selection. For the study 11 papers were finalized and are explained in Table 1.

F. RESEARCH QUESTIONS

The purpose of this review is to enumerate the various techniques available for analysis of data collected from student tracking and the answer the research questions

mentioned. The assortment of papers was done in conformity with the following research questions. The questions are as follows.

RQ1. What are the various ways available in which student activities are tracked?

RQ2. How is the data which is collected analyzed, using what machine learning techniques?

RQ3. In what ways can this analysis done be used to optimize the learning experience?

RQ4. What are the areas in which this field is lacking?

III. LITERATURE REVIEW

Table 1: Literature review table

Title	Summary	Techniques
Hussain, Mohammed et al [2]	A framework is proposed for the data generated by universities. Analysis of the data is designed for the evaluation of the institution against standards of accreditation bodies. This can be used for the academic accreditation of higher education programs. This framework minimizes the human involvement in the process of accreditation. The paper gives a detailed design of the process of aligning assessment with student learning outcomes. The logs of the online learning module such as BlackBoard are used. The rate of student involvement in the learning module, the time between faculty posting files and student viewing, student responses for the learning module are analyzed.	Association, clustering and prediction
Rianne et al [3]	This paper analyzed 17 blended courses with 4,989 students in an institution using Moodle LMS. The logs were used to predict student performance from LMS predictor variables and from in-between assessment grades, using both multi-level and standard regressions.	Multiple Linear regressions
Wang et al [4]	This paper presents a student achievement predicting framework. It includes data processing and student achievement prediction. A layer-supervised multi-layer perceptron(MLP)-based method has been proposed using the data in data warehouse, to predict the achievement of students. To each corresponding hidden layer of MLP supervisions are fed to improve the performance of student achievement prediction.	multi-layer perceptron(MLP)
Daud Ali et al [5]	The ability to predict a student's performance can be help improve the modern educational systems. Existing techniques utilize features which are related to academic performance of students, parent's/family background and assets, and income. Features which include personal information of students like family expenditures are usually ignored. Learning analytics, discriminative, generative classification models are put in application to predict whether a student will be able to complete his degree or not.	Support Vector Machine (SVM), C4.5 Classification and Regression Tree, Bayes Network (BN), Naive Bayes (NB)
Leo et al [6]	This paper uses the students' usage of online learning material as a predictor of academic success. The amount of time that each element such as a text paragraph or an image was visible on the students' screen was recorded. Machine learning methods were applied to study to what extent material usage predicts course outcomes. This paper concludes that the time spent with each paragraph of the online learning material is a moderate predictor of student success.	Support Vector Classifiers (SVC) and ϵ-insensitive Support Vector Regressors (ϵ-SVR) for the purposes of predicting academic success

	This is even when corrected for student time-on-task. This information helps to identify at-risk students. The predictive performance of the model depends on the quantity of data. It is also observed that academic success can be predicted by course material usage. Also this data can be collected with minimal interference to the students' learning process	
Jayagopi et al [7]	The engagement or attention level of the students is analyzed from their facial expressions, head pose and eye gaze using computer vision techniques and a decision is taken using machine learning algorithms. Classification is done using SVM(Linear), SVM(RBF),Logistic Regression	SVM(Linear), SVM(RBF), Logistic Regression
Singh et al [8]	Analysis of data and finding hidden information manually is a very difficult task. To improvise educational data mining, i.e. performance and unambiguousness of obtained models clustering is used here. Data of 84 under-graduate students was used and students grouped according to their final marks they achieved in the course.	EM clustering algorithm
Marwala et al [9]	"This paper provides an efficient educational data mining approach to support e-learning. For this it collects and stores documents (text documents, images and the videos files).unique words are selected from it and K-means clustering is applied. The frequency of the common unique words becomes the input to the K-means clustering.	K-means clustering
Adam et al [10]	Courses are being taught using technology-mediated instruction and e-learning environments. This paper investigates how student motivated engagement profiles developed in educational research can be used as to predict student behavior. It uses tracked behaviors of 249 students from 7 CS1 courses over the span of 3 semesters. It shows that students with different engagement profiles behave differently in an online, wiki-based CSCL system while performing collaborative creative thinking exercises. Thus, such profiles could be useful as student models for providing customized support in e-learning environments	Kruskal-Wallis test
Falakmasir et al [11]	A novel data analysis pipeline is presented. Student Proficiency Inferred from Game data (SPRING) that allows modeling game playing behavior in educational games. A simple interpretable model is obtained that fits the data and predicts learning outcomes. Data is obtained from students playing 11 educational mini-games. Using test data logs from students playing educational games it predicts math assessments accurately. It uses Hidden Markov Models to model data sequences and then extracts features from the sequences. It predicts the score of the traditional assessment by using a regression model	Hidden Markov Model and regression
Marcal et al [12]	Mobile Learning (m-learning) in various education disciplines. This paper examines a post-graduate class consisting of 474 learners in which Short Message Service (SMS) was used. The main goal was to find out if his technology would increase student participation in the course. The paper concludes that, the students who received mobile messaging performed significantly better than the students from the control group. Data was obtained from records of Moodle	Statistical techniques (Kolmogorov–Smirnov test, z-test, f-test, t test)
Santillan et al [13]	This paper predicts the final student´s performance by analyzing their behavior from the Moodle LMS. Three kinds of classifying algorithms are used i.e. JRIP (a RIPPER implementation) which gives classification rules, , J48 (an open source Java implementation of the	JRIP, J48, Bayesian Network with neural network classification model

	C4.5) which has a classification model of the type decision trees and Bayesian Network with neural network classification model	
Elbadrawy wt al [14]	This study presents a class of collaborative multi-regression models. These are personalized to each student. They take into account features related to student's past performance, engagement and course characteristics. They estimate a small number of regression models shared by all students along with student-specific combination weights using historical information. This leads to information sharing and generation of personalized predictions. They are capable of improving the performance prediction accuracy by over 20%.	Regression
Arrigo et al [23]	Learning Analytics in Mobile Learning is a challenging topic.m-learning is characterized by the learners' mobility, localized data and information. Large amount of data can be collected by many users, the affordances provided by the technologies, social dynamics characterizing the context in which learning occurs. This paper provides a task-interaction framework which helps supporting educational decision-making in mobile learning. The framework is based on the factors which involve relationships between the different types of interactions which happen in a mobile learning activity and the various tasks that are pedagogically relevant for the learning activity.	Social Network Analysis (SNA) techniques) semantic web
Manhaes et al [15]	Architecture is proposed that uses EDM techniques to predict and to identify students at dropout risk. This can help the academic managers to monitor the progress of the students and identifying the ones in danger. student identification (id), course id, year and semester of admission, course status, semester id, status of the current semester, CGPA, GPA, number of credits, numeric score and alphanumeric grade(course final situation: approved, failed, absence) were the attributes considered. The data of 1 semester was used to obtain the prediction for the following semester. The classifiers considered were Naïve Bayes (NB), Multilayer Perceptron (MLP),Support Vector Machine with polynomial kernel (SVM1) and RBF kernel (SVM2) and Decision Table (DT).	Naïve Bayes (NB), Multilayer Perceptron (MLP), Support Vector Machine with polynomial kernel (SVM1) and RBF kernel (SVM2) and Decision Table (DT).
De Boer et al [16]	In Massive open online courses(MOOCs) every interaction is recorded which can be used to model performance over the course of the class. The paper is built on studies applying random-effect panel regression to individual student. This is used to explore whether time with resources and number of problem attempts predict achievement. Achievement is defined as longitudinal homework performance	Regression

IV. RESULTS

In this literature review the papers included and their summary and techniques used are summarized. The study indicates that most frequently used method for tracking student activity is through learning management systems. The logs of LMSs particularly Moodle are widely used to track the activities and perform classification, clustering, and prediction on the data generated. BlackBoard is another LMS used. Some techniques use mobile phones to try and optimize the learning experience. Game Activity Logs are also used to predict student performance and make predictions. Regression, multi-layer perceptron (MLP),Support Vector Machine (SVM) [discriminative], Naive Bayes (NB), K-means clustering, Bayesian Network, neural network classification model are some of the

techniques employed. The techniques are broadly classified into the ones using softwares like LMS to track activities, tracking activities online which also includes social media, use of mobiles and apps for such activities, and other methods like game logs, questionnaires etc.

Table 2: Machine learning techniques used to analyze student activity tracking data

Technique	Description
Classification	The task of putting data into several classes based on the training data or labeled set provided. The various techniques include Naïve Bayes Classifier, SVM, Neural Net
Clustering	It involves creating groups or cluster based on some similarity measure. Techniques are K means clustering, EM clustering
Association	Association involves association rule mining a procedure which observes frequently occurring patterns, correlations, or associations from datasets and generate rules based on them
Prediction	This techniques aims to make predictions based on some existing variables. Regression is the technique used for this.
Social Network Analysis	It uses networks and graph theory to investigate social structures

4.1 Tracking through learning softwares: e-Learning involves various forms of online learning mechanisms involving the instructor and learner interaction via the utilization of Information and Communication Technologies(ICT). From educational television shows like "The Magic school bus" and "Bill Nye the Science Guy" of the 90s to the learning apps on cell phones today and YouTube, e-learning has always been around. It provides the advantage of flexible learning from anywhere, anytime, at user's convenience and pace. One of the techniques of e Learning involves the use of Learning management systems (LMS), which are software applications. They are used for the purpose of administrative tasks, documentation, reporting, tracking and delivery of educational and guidance courses as well as programs, resources or development modules [18] Several Learning management systems(LMS) like Moodle [19], Blackboard [20] and Shakai [21] are widely used. Several research techniques study the data generated from these LMS to track student activity and use it to optimize learning experience.

a. **Tracking students' Online Learning Activities:** The internet has become commonly available to al today. Social media, YouTube, and other online activities provide a medium to study behavior of students. Instructors can get insights about how students behave online, access educational course materials, talk about educational topics on social media. Applications can be developed which, with the help of activity information available provide recommendations, automatic email reminders to students encouraging or instructing them to view educational content. Online learning has great potential in improving learning and teaching methods. It provides easier access to educational resources by bringing down the cost as well as time of commuting. It also allows students to study on a schedule, which is suited according to their needs. The growth of internet, enthusiasm around innovative, technology-based education initiatives, advent of social media, and acceleration in online course enrollments [22] makes educators ponder of the continuing expansion of online learning could be leveraged to improve the education sector. This can be

beneficial as it can increase the academic access, progression, identification of students at risk, success of low-income and underprepared college students.

b. **Tracking through m-learning:** M-learning, is an educational system which stands for mobile learning. With the advent of smart phones and internet everyone has access to android and i-OS based phones. Using these mobile devices, it's very easy to have a continuous access to the learning resources in the form of apps, e-books, YouTube and other such sources. But studies show that tracking activities through smart phones is limited. Earlier work has been done using SMS to improve learning [12]. But the research in these is still emerging. This field has a lot of potential. Tracking the usage of learning apps can help educators improve teaching and learning and make the process more interactive [29]. IT can aid in decision making as presented in the paper by Arrigo et al [23] .Chen et al[24] used Facebook to explore differences in student learning outcomes and satisfaction depending on various learning styles. This is an emerging field with many ways available which can be tapped to optimize learning.

c. **Other methods:** Other methods include a variety of techniques like tracking game logs of students and using it to predict their performance. Preparing questionnaires for students, tracking them by regular feedback, and using other techniques like tracking physical fitness to study the relation between physical health and student performance [25]. Another can be finding the links between stress, mental health [26] and education and so on.

V. CONCLUSION

To provide an insight into the current trends of this fields the review based on some longer review, summarizes the literature on activity tracking of students and provides recommendations for educators. The study shows that most of the work has been done in this field using learning management systems. More work can be done in the field of mobile learning as it is an emerging field. Mobiles are the most used technological devices; hence they are a mine of huge datasets. They can supply sensor data, which is in the form of GPS, accelerometer, activity data of app usage and so on. All this can aid in improving the education system. Student's interaction in the classroom can be improved by using the insights gained from the analysis of this data.

REFERENCES

[1]. Liñán, Laura Calvet, and Ángel Alejandro Juan Pérez. "Educational Data Mining and Learning Analytics: differences, similarities, and time evolution." *International Journal of Educational Technology in Higher Education* 12.3 (2015): 98-112.

[2]. Hussain, Mohammed, et al. "Mining educational data for academic accreditation: aligning assessment with outcomes." Global Journal of Flexible Systems Management 18.1 (2017): 51-60

[3]. Conijn, Rianne, et al. "Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle." (2017).

[4]. Qu, Shaojie, Kan Li, Shuhui Zhang, and Yongchao Wang. "Predicting Achievement of Students in Smart Campus." IEEE Access 6 (2018): 60264-60273.

[5]. Daud, Ali, et al. "Predicting student performance using advanced learning analytics." *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 2017.

[6]. Leppänen, Leo, et al. "Predicting Academic Success Based on Learning Material Usage." *Proceedings of the 18th Annual Conference on Information Technology Education*. ACM, 2017

[7]. Thomas, Chinchu, and Dinesh Babu Jayagopi. "Predicting student engagement in classrooms using facial behavioral cues." *Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education*. ACM, 2017

[8]. Mishra, Akansha, Rashi Bansal, and Shailendra Narayan Singh. "Educational data mining and learning analysis." *Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on*. IEEE, 2017.

[9]. Appalla, Padmaja, Venu Madhav Kuthadi, and Tshilidzi Marwala. "An efficient educational data mining approach to support e-learning." *Wireless Networks* 23.4 (2017): 1011-1024.

[10]. Eck, Adam, Leen-Kiat Soh, and Duane F. Shell. "Investigating differences in wiki-based collaborative activities between student engagement profiles in CS1." *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 2016

[11]. Falakmasir, Mohammad H., et al. "A data-driven approach for inferring student proficiency from game activity logs." *Proceedings of the Third (2016) ACM Conference on Learning@ Scale*. ACM, 2016

[12]. Marçal, Edgar, et al. "Mobile phone text messaging to increase student participation: an experience in a blended course." *Proceedings of the 31st Annual ACM Symposium on Applied Computing*. ACM, 2016

[13]. Sanchez-Santillan, Miguel, et al. "Predicting students' performance: Incremental interaction classifiers." Proceedings of the Third (2016) ACM Conference on Learning@ Scale. ACM, 2016.

[14]. Elbadrawy, Asmaa, R. Scott Studham, and George Karypis. "Collaborative multi-regression models for predicting students' performance in course activities." *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. ACM, 2015.

[15]. Manhães, Laci Mary Barbosa, Sérgio Manuel Serra da Cruz, and Geraldo Zimbrão. "WAVE: an architecture for predicting dropout in undergraduate courses using EDM." Proceedings of the 29th Annual ACM Symposium on Applied Computing. ACM, 2014

[16]. DeBoer, Jennifer, and Lori Breslow. "Tracking progress: predictors of students' weekly achievement during a circuits and electronics MOOC." *Proceedings of the first ACM conference on Learning@ scale conference*. ACM, 2014.

[17]. Accessed 14th Jan, 2019 Watershed Systems, watershedlrs.com/what-is-learning-analytics

[18]. Ellis, Ryann K. (2009), Field Guide to Learning Management, ASTD Learning Circuits

[19]. Rice, W. H. "9. E-Learning Course Development: A complete guide to successful learning using Moodle 1.9." (2008).

[20]. Boeren, Ellen. "The Blended Learning Environment in Higher Education: The Tutor and Student Perspective." *Handbook of Research on Learning Outcomes and Opportunities in the Digital Age*. IGI Global, 2016. 180-197.

[21]. Berg, Alan Mark, and Michael Korcuska. *Sakai courseware management: The official guide*. Packt Publishing Ltd, 2009

[22]. Picciano, Anthony G., Jeff Seaman, and I. Elaine Allen. "Educational transformation through online learning: To be or not to be." *Journal of Asynchronous Learning Networks* 14.4 (2010): 17-35.

[23]. Arrigo, Marco, Giovanni Fulantelli, and Davide Taibi. "Challenges of Using Learning Analytics Techniques to Support Mobile Learning." *International Association for Development of the Information Society* (2015).

[24]. Chen, Yu-ching. "Linking learning styles and learning on mobile Facebook." *The International Review of Research in Open and Distributed Learning* 16.2 (2015).

[25]. Hills, Andrew P., Donald R. Dengel, and David R. Lubans. "Supporting public health priorities: recommendations for physical education and physical activity promotion in schools." *Progress in cardiovascular diseases* 57.4 (2015): 368-374.

[26]. Wang, Rui, et al. "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones." *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 2014.

[27]. Baker, Ryan Shaun, and Paul Salvador Inventado. "Educational data mining and learning analytics." *Learning analytics*. Springer, New York, NY, 2014. 61-75.

[28]. Kitchenham, Barbara Ann, David Budgen, and Pearl Brereton. *Evidence-based software engineering and systematic reviews*. Vol. 4. CRC Press, 2015.

[29]. Auksorncherdchoo, Soontorn, and Anon Sukstrienwong. "Software development for student behavior tracking system." *Proceedings of the 9th International Conference on E-Education, E-Business, E-Management and E-Learning*. ACM, 2018.

[30]. Han, Kyungsik, Svitlana Volkova, and Courtney D. Corley. "Understanding roles of social media in academic engagement and satisfaction for graduate students." *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2016

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Technical Aspects and Challenges of Cloud Computing Environments

Roshan Jameel
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
roshijameel@gmail.com

Harleen Kaur *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

M. Afshar Alam
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract – *The cloud computing has seen enormous changes in itself in the past few years. Not just in terms of the services and resources provide but also in terms of crowding and the various types of data the users are creating and also the speed at which it is being generated. The main purpose of using the cloud is that, it increases the capacity and capability of the system without investing in development of the whole new physical infrastructure. This paper provides an overview about the cloud-computing environment along with the various models provided by the cloud infrastructures. The impact of the cloud on the every day users and society is also discussed in this paper that throws light on the various advancements that have affected the society for its betterment. Apart from all the benefits and features offered by the cloud there are still some challenges that will hinder the users and the performance of the cloud, some of them are discussed in this paper and also the research areas that one might consider while working on the cloud is briefed.*

Keywords *- Cloud Computing Environment, Big Data, Internet of Things, Self-Learning Systems, Virtual Machines*

I. INTRODUCTION

Cloud computing has become a platform of huge interest in communities including industries, research and academics. It is developing and evolving with time and includes the various technologies like distributed networks, storages, resources, infrastructures etc. [1]. The cloud-computing model upholds the characteristics like availability, scalability, resource pooling, broader network and pay as you go services. The availability means that when the resources are required by the users they can be accessed anytime form anywhere, while scalability ensures that the number of machines and resources can be added or deleted on the basis of the requirement. Moreover, the resources are pooled together and give the users the better opportunity for development of their business with the broader network availability [2]. The clouds have three types of deployment models including public, private and hybrid. The public cloud is the cloud that is widely used nowadays, in this model the applications and resources are hosted by the providers and are made available to

*Corresponding Author

the users on demand. The private cloud is solely dedicated to a single organization that can be either hosted on the site of provider or on the customer's site. While the hybrid cloud gives the mixture of both the clouds, it gives the facility to the organizations to have a private cloud for storing valuable data and also a public cloud for performing computations. This makes the use of best of both the worlds, and provides a faster and easier access to services.

The cloud endorses three service models, called IaaS (Infrastructure as a Service), PaaS (Platform as a Service) and SaaS (Software as a Service) as shown in Fig.1. The IaaS is the basic level of service models in which the providers provide the users with the hardware resources and virtual machines that are already configured. Basically it provides the infrastructure required for the development of the software mentioned in Fig. 2. The IaaS provides extra storages for the backups, better bandwidth of the network for servers and allows the higher level of computations. The examples of IaaS are Google's Compute Engine, Amazon EC2 and IBM SoftLayer. The next level is PaaS in which the provider offers a platform to the users on which the users can develop and deploy their applications. It provides the users with the platform that is ready to use and hence eliminates the requirement of building and maintain the infrastructure on their own. It comes between the two service models, IaaS and SaaS and therefore referred to as middleware. Google App Engine, Apache Stratos and IBM BlueMix are some of the examples of the PaaS that helps in easier development of the software. SaaS is the uppermost level of the cloud models, and is called 'on-demand software' because it delivers the software as the product to the users. The user can access the software after taking the license for the subscription and pay accordingly. The SaaS offers the end product that is retrieved by the using some interface like a web browser. SalesForce and Office365 are the most common SaaS products.

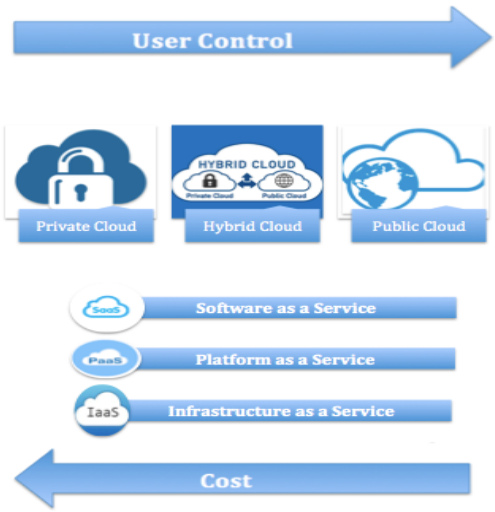


Fig. 1. The Cloud Architecture

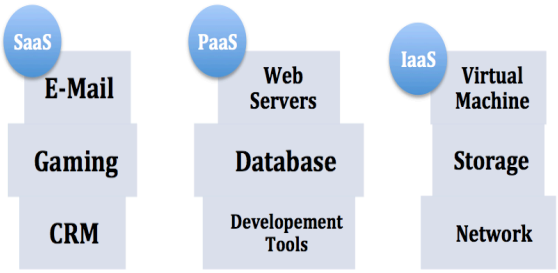


Fig. 2 Services provided by Cloud

The cloud computing has become so popular because of its benefits in terms of operations and economy. The services that clouds are offering are changing rapidly day by day, and there are still chances for the change and development in this area. The main focus of the industries is to make the cloud able to handle the heterogeneous data from various resources. Therefore, the architectures are evolving for handling this change and making the computations more efficient and secure. This paper throws the light on the brief introduction of the cloud-computing environment that is followed by the impact of the cloud computing on the society, next section discusses the issues and challenges that one might face while working in the cloud environment. The next section sets out the possible research areas in the field of cloud computing and the last section concludes the paper.

Your goal is to simulate the usual appearance of papers in *conference proceedings*. For items not addressed in these instructions, please refer to the last issue of your conference's proceedings for reference or ask your conference Publications Chair for instructions.

II. Societal Impact of Cloud Computing

As the cloud computing is becoming ubiquitous and pervasive, its impact on the society is also increasing. The advancements in the area aim at making the infrastructure able to handle the heterogeneous data on any type of cloud infrastructure and deal with the changes that are required in order to accommodate accordingly. The cloud will have impact on the following areas of today's society:

A. *Big Data on Cloud*

With the emergence of cloud computing models, the amount of data is increasing and the speed by which it is generating is very high as compared to the computing capabilities of the traditional devices and networks. This data volume is called 'Big Data'. The data that is generated by the users or companies needs to be transferred on the cloud so that whenever it is required, it can be accessed easily anytime from anywhere. Some of the data is never accessed or used, which is referred to as dark data. Performing the analysis or any type of processing on the data over the cloud environment is quite costly therefore; it's better to process it beforehand. The cloud infrastructure is usually decentralized but the current state of big data analytics says it works on the centralized infrastructure only and hence it becomes challenging to analyze the big data on cloud.

The processing and management of the resources on the cloud nodes is the primary concern. The cloud is supposed to be able to handle the heterogeneous and ad-hoc nature of the cloud and process the data in the distributed environment with minimal efforts [3]. Another concern is to build the models that have the capability to scale up and down and horizontally as well as vertically. Horizontal scaling means nodes can be added to the data center, while vertically means more data centers can be added to the model. One more

challenge that one might face while implementing big data on cloud is stacking the software for processing [4]. The data in the real world scenario is mostly unstructured and the challenge is to retrieve knowledge from such data. D2N2K is one of the approaches to convert this unstructured data into structured one and transform it to get useful information [5]. The last but not the least important issue related to big data on cloud is legal and governance issues related to the sensitive data of the organizations, that needs to be maintained in terms of access control, integrity, privacy and security.

B. *Cloud Computing and Internet of Things (I-o-T)*

The growth in the area of wearable including gadgets and sensors, that helps in bringing the users closer to the computing and devices. IoT is the trend that combines the sensors, devices, infrastructures, appliances and various environments together for developing smart world for the users [6,7]. The focus of IoT is improving the processing of systems and reducing the human intervention as much as possible. The things might refer to sensors like biometric sensors, mobile phone sensors, chips, gadgets depending upon the context it is being used. Cloud is required to connect all these devices and networks and helps in monitoring their functionalities and managing their functioning.

But there are few challenges in this area too, the first and the most important one is security. Various types of networks like wireless network, cloud data centers, sensors, clouds are integrated together to provide the IoT environment to the users, and hence securing all these networks becomes important. However encryption user authentication is being applied nowadays to prevent the attacks from the outsiders, but there is still threat from the insiders and requires protocols to monitor the actions of the users thereby securing the system from malicious attacks or installations. Moreover, sensing devices and decision making on the basis of the collected information also need some advancements, as of now the sensors are combined with the environment, in future the sensors should be people centric so that the actual users can give feedbacks about their environments [8,9].

C. *Self-Learning Systems using Cloud*

Nowadays every data that is generated by the users is transferred to the cloud, this data could be in the form of audio, video and images. Also the metadata including network, activity of the users etc. is being transferred to the cloud because of its low cost and easier availability. Machine learning is being implemented for carrying out research in the field of voice recognition, pattern recognition, text or image analysis and language processing [10]. All these researches come under the field of Deep Learning [11] that is a subset of artificial intelligence. With the development of Tensorflow [12] and Nervana [13] the implementation of deep learning has become very easy and economical. In the cloud environments the presence of GPUs have reduced the time of computation for the algorithms on big heterogeneous data [14]. The deep learning has gained so much popularity because of its capability of predicting the outcomes based on past events. All these mechanism constitutes the future generation of cloud.

D. *Services on Cloud*

Initially the service models that were offered by the cloud computing were IaaS, PaaS and SaaS. But with the advancement in the technologies the services provided by the cloud are increasing in variety and are becoming more powerful and richer. For example AaaS (Acceleration as a Service) was proposed for providing the acceleration in the capability of the GPUs [15]. However this is not widely used and hence is not that developed but with time it will become mature. CaaS (Container as a Service) is another service model that gives the benefits of development and deployment of the containers. However, these containers are not widely used nowadays but are slowly replacing the virtualization. One more service model that is gaining attention these days is FaaS

(Functions as a Service) in which the functions are executed on the cloud that are started by the events.

III. Cloud Computing Challenges and Issues

However, the cloud computing technology has emerged widely and several types of industries are moving towards it. But there are still some issues and challenges that are not resolved yet and the researches are still going on. This section summarizes some of the issues and challenges in this area.

A. *Cloud Computing Challenges*

Handling Traffic: The data traffic is examined nowadays to achieve better decisions making for the businesses. That's why the data traffic analysis is considered to be an important aspect. Many web applications and networking operators needs to know about the flow of the traffic for planning and management. However, it is not an easy task, there are some challenges like huge number of nodes, links, servers needs to be analyzed which is difficult for the existing systems. There is not much research done in this area for now, that makes it a challenge for the organizations to handle the traffic.

Integrating Servers: It is an effectual method for maximizing the utilization of resources and minimization of the consumption of energy. In this approach several virtual machines that are not being utilized efficiently are migrated on a single server, to make the other servers rest. It is often considered as a NP-hard problem called bin packing [16]. The challenge is to consolidate the servers without affecting the performance. However, sometimes when the resources that are migrated for server consolidation, result in congestion of the resources [17], that needs to monitor carefully and the system should also be able to react in case any type of congestion occurs [18].

Migrating Virtual Machines: Virtual machines are migrated in the cloud environment to balance the load between the data centers. Additionally, it provides better resource allocation and robust provisioning. The main benefit of migrating the virtual machine is, it prevents the hotspots in the environment, but it is not that simple to achieve this. The detection and response time for migration is still lacking. Moreover, the transfer of the state should also need to be maintained regularly.

Efficient Energy Management: Improving the efficiency of the cloud is one of the major concerns of the service providers. Approximately 53% of the total cost of the cloud is because of the power of the data centers [19]. Therefore, the providers are under huge anxiety in order to reduce the power consumption. The objective is not just the lesser energy consumption but also meeting the environmental regulations and standards. In order to make an efficient energy management in the cloud, several techniques can be used like turning off system when not in use [20], server consolidation [21], efficient job scheduling [22] etc. But the main challenge is to attain an efficient trade-off between the energy saving and performance.

B. *Cloud Computing Issues*

Lack of Expertise: It is considered as one of the biggest issue in cloud computing. However, there are IT experts that are trying to learn to work on cloud but still the workers with required knowledge is still lacking. As the technology is continuously improving with time, the IT workers need to polish their skills accordingly. And in order to match up with the requirements of the cloud environment, companies are providing training and learning facilities to their existing workers, but finding appropriate knowledgeable person still remains the issue.

Vendor Lock-In: Presently the cloud market is mostly dominated by a very few vendors like Amazon, Google, Microsoft, IBM etc. that gives rise to the issue of vendor lock-in, in which the customers who are availing the cloud facilities from a vendor become so much dependent on the providers that it becomes nearly impossible for them to move to some other provider. This increases the concern of the users about moving to the cloud, and in order to move towards any cloud provider the organizations need to explore the chances of moving to any other cloud if needed in future.

Lack of Customer Support: It is the most common complaint about the cloud services. When the organizations need to migrate or recover the resources, the providers still lacks in providing proper and timely customer support. It is important to analyze the providers first that whether or not the services will be flexible so that the main focus of the organization remains on the business development rather than concerning about the everyday operations. It is really significant to select the provider that provides continuous support, and is able to handle any issue that might occur.

Lack of Automation in Billing: Cloud computing environment is subscription based in which the user pays as per usage and has the full flexibility to scale the resources. There are various ways in which the cloud services can be bought like prepaid, reserved, pay per use etc. The billing of cloud services is quite complicated and it has different price of different models. Therefore, having automation in billing in cloud environment is still an issue that needs to be addressed.

IV. FUTURE RESEARCH AREAS

There are always chances of advancements in every technology, so is in the cloud computing. Some of the possible research areas are given in this section that could contribute to the development of the new trends in the cloud.

A. *Cloud Security*

Along with all the capabilities of the cloud and all the services that cloud provides, the security is still one of the major concerns of the users before moving on to the cloud. It is the responsibility of the service providers to ensure the guaranteed security of the users information and data [23]. The cloud has multiple risks related to the security of data and needs to isolate the users from one another, this becomes more complex when the advance cloud architectures are used like fog computing, volunteer computing etc. because in such architectures the nodes are accessed through the routed network that is easy to hack [24]. The affect on security and privacy of the data might also occur while scaling up and down the systems. While working in the distributed clouds the authentication of the users is also one of the major area of concern. Another concern is to develop the encryption mechanisms that are less resource consuming and more effective. DDoS attacks are also threats to the cloud that have huge impact on the systems overall security.

B. *Cloud Reliability*

The cloud is being adopted for various applications some of them require remote storages and computing, and for that the reliability is major concern. Because in case of remote resources requirements there are chances of failure and that affects large number of services. This has happened in past to renowned services like Amazon.com, Netflix and Dropbox that have affected the lost in terms of money as well as users. Since such outages are not predictable and unavoidable, the losses caused by them become a challenge for the organizations, and researches are carried out to develop systems that are more reliable. One way to deal with the failures associated with the hardware failures in case of natural disaster is to keep the replicas on multiple geographical areas [25]. Microsoft also offers an architecture called FailSafe that is disaster resistant. However, the distributed cloud does

not have any such disaster resistant application and hence requires to be dealt in future. Moreover, the disaster management is a costly process and the requirement is to recover the cloud with minimal efforts in terms of cost as well as time.

I.*Organized Management*

Typically there are two tasks of management in the cloud infrastructure one deals with the compliances between the users and the service providers in order to optimize the performances and the second task includes the providing of services and resources to the users and ensuring the quality of service being achieved. The compliances are usually ensured using the Service Level Agreements [26] but it is difficult in case of distributed cloud environments. There are several techniques available for measuring the performance of the applications, however the existing techniques face difficulties while monitoring the edge nodes, heterogeneous devices, multiple levels of hierarchy etc.

D. *Sustainable Infrastructure*

The data centers nowadays are consuming huge amount of energy and have impacts on the environment due to the emission of carbon footprints. It is not possible to totally eradicate the data centers from the organization but geographical distribution of the data centers can help in reducing the amount of problems caused and helps in achieving sustainability. This requires the development of the algorithms that supports the geographically distributed architecture, which will in turn lessen the energy consumed and provide better quality of service. Moreover, there is a tradeoff between the performance of the cloud and the energy consumption efficiency. This is an open area of research. Conserving the energy is not the only thing to be considered while implementing; a tradeoff that satisfies both the efficiency as well as sustainability is to be selected.

V. CONCLUSIONS

Cloud computing is showing great prospects and potential as it can be related to various applications and technologies related IT. Because of its dominance in the It industry, the organizations are moving towards the cloud and the whole market place has changed. Although the cloud is revolutionizing the current state of IT but the users must be aware of the issues and challenges that might come along with it. This paper gives a brief outline about the cloud environment and its architecture. It also discusses about the impact of cloud computing on the current societal situation. Some of the challenges and issues that one might face while implementing the cloud are also discussed. As the cloud is still in its development phase the possible research areas in the field of cloud computing are also highlighted in this paper.

REFERENCES

[1] Cloud Security Alliance Report,"Security Guidance for Critical Areas of Focus in Cloud Computing V2.1" (http://cloudsecurityalliance.org/)

[2] http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-def-v15.doc

[3] Assuno M.D., Calheiros R.N., Bianchi S., Netto M.A., Buyya R., Big data computing and clouds: trends and future directions, J. Parallel Distrib. Comput. 7980, pp. 3-15 (2015).

[4] Horey J., Begoli E., Gunasekaran R., Lim S.-H., Nutaro J., Big data platforms as a service: Challenges and approach, in: Proceedings of the 4th USENIX Conference on Hot Topics in Cloud Computing, (2012).

[5]https://www.nsf.gov/awardsearch/showAward?AWD_ID=1705169&HistoricalAwards=false.

[6] Mattern F., Floerkemeier C., From the internet of computers to the internet of things, in: Sachs K., Petrov I., Guerrero (Eds.) P. From Active Data Management to Event-based Systems and More, pp. 242–259, 2010.

[7] Gubbi J., Buyya R., Marusic S., Palaniswami M., Internet of Things (IoT): A vision, architectural elements, and future directions, Future Gener. Comput. Syst. 29 (7) pp. 1645–1660, 2013.

[8] Delmastro F., Arnaboldi V., Conti M., People-centric computing and communications in smart cities, IEEE Commun. Mag. 54 (7) pp. 122–128, 2016.

[9] Peixoto J.P.J., Costa D.G., Wireless visual sensor networks for smart city applications: A relevance-based approach for multiple sinks mobility, Future Gener. Comput. Syst. 76 pp. 51–62, 2017.

[10] Rendle S., Fetterly D., Shekita E.J., B.-y. Su, Robust large-scale machine learning in the cloud, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1125–1134, 2016.

[11] Dean J., Corrado G.S., Monga R., Chen K., Devin M., Le Q.V., Mao M.Z., Ranzato M., Senior A., Tucker P., Yang K., Ng A.Y., Large scale distributed deep networks, in: Proceedings of the 25th International Conference on Neural Information Processing Systems, pp. 1223–1231, 2012.

[12] https://www.tensorflow.org/.

[13] https://www.nervanasys.com/cloud/.

[14] Cui H., Zhang H., Ganger G.R., Gibbons P.B., Xing E.P., GeePS: Scalable deep learning on distributed GPUs with a GPU-specialized parameter server, in: Proceedings of the Eleventh European Conference on Computer Systems, pp. 4:1–4:16, 2016.

[15] Varghese B., Prades J., Reao C., Silla F., Acceleration-as-a-service: Exploiting virtualised GPUs for a financial application, in: 11th IEEE International Conference on E-Science, 2-15, pp. 47–56, 2018.

[16] Chekuri C., Khanna S., On multi-dimensional packing problems. SIAM J Comput. 33(4): pp. 837–851, 2004.

[17] Padala P., Hou K-Y., Shin K-G., Zhu X., Uysal M., Wang Z., Singhal S., Merchant A., Automated control of multiple virtualized resources. In: Proc of EuroSys, 2009.

[18] Wood T., Shenoy P., Yousif M., Black-box and gray-box strategies for virtual machine migration. In: Proc of NSDI, 2007.

[19] Hamilton J., Cooperative expendable micro-slice servers (CEMS): low cost, low power servers for Internet-scale services In: Proc of CIDR, 2009.

[20] Brooks D. M., Bose P., Schuster S.E., Jacobson H., Kudva P., Buyuktosunoglu, A., Wellman J-D., Zyuban, V.V., Gupta M., Cook P. W., Power-aware microarchitecture: design and modeling challenges for the next-generation microprocessors, IEEE Micro, 2000

[21] Vasic N., Barisits M., Salzgeber, V., Kostic D., Making cluster applications energy-aware. In: Proc of automated ctrl for datacenters and clouds, 2009.

[22] Srikantaiah, S., Kansal, A., F. Zhao, F., Energy aware consolidation for cloud computing. In: Proc of HotPower, 2008

[23] Chejerla, B.K., Madria, S.K., QoS guaranteeing robust scheduling in attack resilient cloud integrated cyber physical system, Future Gener. Comput. Syst. 75 pp. 145–157, 2017.

[24] Wang, Y., Uehara, T., Sasaki, R., Fog computing: Issues and challenges in security and forensics, in: Computer Software and Applications Conference (COMPSAC), 2015 IEEE 39th Annual, vol. 3, pp. 53–59, 2015.

[25] Ferdousi, S., Dikbiyik, F., Habib, M.F., Tornatore, M., Mukherjee, B., Disasteraware datacenter placement and dynamic content management in cloud networks, IEEE/OSA J. Opt. Commun. Networking 7 (7) pp. 681–694, 2015.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Failure prediction in IoT Environment using Support Vector Machine.

Ranjana Singh
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
ranjanaelectronics@gmail.com

Syed Imtiyaz Hassan *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract - *The advent of Internet of Things (IoT) technology brings an automation i n our day today life with an exponential growth in network size and huge increase in the n etwork devices like sensors, RFIDs, microprocessors/controllers etc. This give rise to an e ssential need for proactive efficient mechanism to maximize the network uptime. In IoT en vironment, the system performance may be increased if the accurate failure predictions ar e known the downtime may be reduced and the impact of network failure may be mitigate d in a specified time frame to manage the overall failure impact. Due to the large number of components in IoT environment the failure predictions strategies are essentially import ant. This paper proposed a method to use Support Vector Machine (SVM) approach for p roactive failure prediction which reduces the undesirable consequences generated by syst em failure in IoT environment.*

Keywords - Logging, SVM, Failure Prediction, Syslog, IoT.

I. INTRODUCTION

IEEE described the phrase "Internet of Things" as: "A network of items each embedded with sensors which are connected to the Internet." [1]. The Internet of Things (IoT) technology will bring a revolution in all aspects of our lives. The number of devices connected to IoT environment is expected to reach to 50 billion by 2020 and which will give rise to an enormous amount of valuable data [2]. The IoT system is supposed to work successfully if the transmission of real time data flows properly/accurately, however if the flow of information is slow or lags, the decision making of the IoT system may get incorrect. Checking the system logs give a user/system an idea of how healthy the system is running and the user can predict the futuristic system failures which deteriorate the system performance. Reading the System logs are best way to manage a computer system as they provide the history of the events happening in the system.

The logs of system are stored in the system log file, and by reading the logs it may be possible to determine reasons for the happened events and also know the criticality of the event. (Critical, Error, Alert or information). These log files contain messages that show a change of system state pertaining to the CPU utilization, the RAM utilization, the bandwidth utilization, the temperature of the device etc.

While a single message in the log file may not be sufficient for failure prediction, a sequence or pattern of log messages may be used to predict futuristic system failure. The future event can also be predicted after reading the information in the logs of the events. The system hardware components are monitored and the information such as temperature, CPU, RAM and bandwidth utilization is logged in the log files.

*Corresponding Author

The services offered by IoT can benefit many application area, in IoT heterogeneous devices are used [3]. The Logs of the IoT devices like Rasberry pi, Arduino, Intel Galileo Etc. may be recorded and pre-processed by filters and then analysed on the basis of threshold values to identify the coming failure indications and future trends. The system only appends the lines in the log file and never change the previous written text, the each line reports the occurrence of the specific event. Proper use of logs alongside powerful log analysis tools will improve the performance and efficiency of the systems in IoT environment.

II. PROBLEM STATEMENT:

With reference to the advent of Internet of Things (IoT), there has been an exponential growth [4] in network size and it is predicted that many more devices will be connected in future and device status monitoring became more demanding task. If accurate failure predictions are given, the impact of system failure may be mitigated in a time frame. However the challenge is to provide accurate predictions with sufficient lead time.

The IoT system is supposed to work successfully if the transmission of real time data flows properly/accurately, however if the flow of information is slow or lags, the decision making of the system may get wrong. This problem may be addressed by monitoring the logs generated by the IoT systems. These logs may be stored in its own file system or the system sends their logs to a remote machine using Syslog.

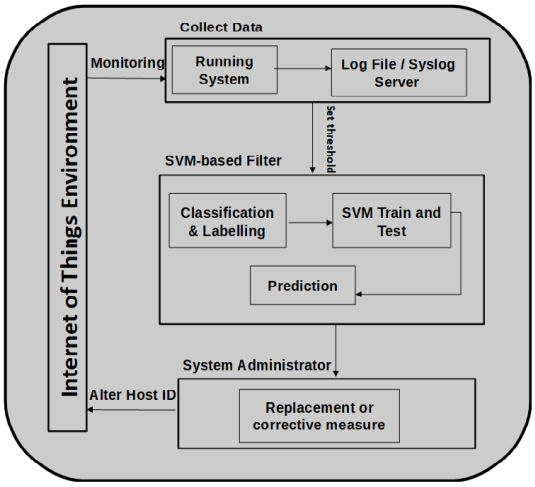
In order to increase the system performance, numerous proactive failure prediction techniques have been introduced which decreases undesirable effects, improves the fault tolerance and enhances the system performance. This paper describes a Support Vector Machine (SVM) approach to predict system failure based on the linear classification of system logs.

III. SVM-BASED FAILURE PREDICTION IN IOT ENVIRONMENT

The objective of this paper is to propose a model to predict failures of the devices in IoT runtime environment by using the logs of the hardware devices. The motive is to develop a proactive fault tolerance method, which monitors the system health using log files and predicts for the upcoming hardware failure based on the recorded threshold readings. SVM is supervised learning method and is mostly used for classification and predictions. SVM basically uses two sets of data viz. Training set data and testing set data, the training dataset enlists the class labels and features or variables to differentiate between the non-failed and failed systems. The Most relevant metrics for SVM classification performance evolution is mentioned in Table 1 [5].

Table 1: SVM Classification Metrics

Name	Syntax	Definition
True rate (		
Falsc rate		
Preci		
Accu		
F-sco		



TP: True positive classes. FP: False Positive classes.

TN: True negative classes. FN: False negative classes.

Fig 1: Proposed schema of SVM based failure prediction

IV. METHODOLOGY

As per the Fig.1 there are three sections in the schematic view of the proposed approach viz. the Section I: Collect Data, Section II: The SVM based filter and the Section III: System Administration

Section I: Data Collection

The logs are very important in the IoT systems, because of the nature of data however for the volume, variety and velocity of sensor data, the devices are not the right place to hold the data for analysis, which leads to the requirement of keeping the data at a remote location for a reliable backup and analysis purpose.

The Syslog protocol is supported by a wide range of devices and can be used to log different types of events occurring in the machine, it is a way for network devices to send event messages to a logging server and this logging server is known as Syslog server [6]. Syslog simply sends messages to a central location when specific events are triggered. Syslog is

an efficient way to consolidate logs from multiple sources into a single location.

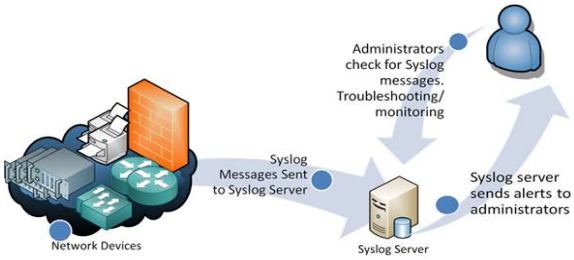


Fig 2: Centralized Logging Architecture

Section II: The SVM Based Filter

The input for the SVM based filters are the log files generated in the running host or the logs files stored at the remote locations the features like Temperature, RAM, Bandwidth utilization are recorded in the log files.

Algorithm 1 for Monitoring an IoT Host

```
Start
     int N; //number of IoT Host
     int i;
     for i = 0; i < N; i++ do {
          Get "IoT-Host-Temperature" and "IoT-   Host-
          CPU-utilization" and "IoT-Host-    bandwidth"
          and "IoT-Host-RAM"
          write IoT Host health status to output
     }
     Return output and send it to SVM-based Filter
End
```

Algorithm 2 for SVM-based Filter

```
Start
     int N; //number of IoT Host
     int i;
     for i = 0; i < N; i++ do {
          if (IoT-Host-Temperature) or (IoT-Host CPU-
                    utilization) or (IoT-Host-bandwidth) or
                         (IoT-    Host-RAM)>=(pre-specified
                    threshold)
          then
               set Host label to -1
     }
     else
          set Host label to 1
          test and train of SVM
     Return output to System Administration Section
 End
```

Section III:

The output of the SVM based filter is passed to the system administration section, and then the host ID of the failed host is identified accordingly, the corrective measures are taken.

V. RESULTS

Suppose that the threshold is chosen as 70%, the readings for host features like temperature, CPU utilization, Bandwidth utilization, RAM are collected in a log files and these log files which shows the health status of a host is then passed to the SVM-base filter where each host is labelled as failed or Non failed Host on the basis of selected threshold level. Finally the specific host number is sent to the system administration section, where actions of replacement or any corrective measures will be taken.

VI. CONCLUSION

The aim of this paper is to propose a method for failure prediction in IoT environment by using the information contained in log files through SVM based filters. The proposed approach uses the SVM which shows a satisfactory performance for the failed and non-failed host classification. The operation cost and productivity loss may be reduced if the accurate prediction of system failure is known to the user. The system failure information may be accessed by the log files of the system, the log files records the system events and contains the very useful information about the system failures the records accessed from the log files provides the system information which can be useful for determining the causes, predict the actions and rectify the problems.

REFERENCES

[1] Roberto Minerva, Abyi Biru, Domenico Rotondi: Towards a definition of the Internet of Things (IoT): IEEE (IEEE internet initiative): 27th May 2015.

[2] Udit Gupta: Monitoring in IoT enabled devices: Information Networking Institute Carnegie Mellon University, Pittsburgh – Pennsylvania, USA. Research Gate: July 2015.

[3] Somayya Madakam, R. Ramaswamy, Siddharth Tripathi: Internet of Things (IoT): A Literature Review: Journal of Computer and Communications, 2015, 3, 164-173: May 2015.

[4] Erric W Fulp, Glenn A. Fink, Jereme N. Haack: Predicting Computer System Failure Using Support Vector Machines

[5] Qi Wang, Wajih Ul Hassan, Adam Bates, Carl Gunter,: Fear and Logging in the Internet Of Things: February 2018

[6] Jonathan de Carvalho Silva, Joel J. P. C. Rodrigues, and Mario Lemes Proenc¸a Jr.: XXXV Brazilian Symposium on Telecommunications and Signal Processing: IoT Network Management: Content and Analysis: September 2017.

[7] Okorogu V.N, Obioma Peace C., Oranugo Charles O, Ezeribe Basil E. The Internet of Things: Insight to a Technological Revolution International Journal for Research in Applied Science & Engineering Technology (IJRASET), September 2017

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Vanadium pentoxide (V_2O_5) Doped Poly (aniline): Synthesis and Characterization

Shama Islam
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
shamaphysics786@gmail.com

S. S. Ashraf *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
shahabash@gmail.com

M. Afshar Alam
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
aalam@jamiahamdard.ac.in

Abstract **-** *Conducting polymer composites of poly (aniline)/vanadium pentoxide (PANI/V_2O_5) were synthesized by polymerization of aniline with V_2O_5 using $(NH_4)_2 S_2O_8$) as an oxidant. The V_2O_5 is varied in five different weight percentages of PANI in PANI/V2O5 composites. The synthesized polymer composites are characterized by dc conductivity; UV–Visible absorption spectroscopy and XRD techniques.*

Keywords **-** *PANI/V_2O_5 composites; UV–Visible absorption spectroscopy.*

I. INTRODUCTION

Polymers have been emerged as a very important class of materials. They have unique properties of optical, electrical and chemical. It leads to a large number of technological applications. These classes of materials provide wide scope for tuning to their electrical conductivity from semiconducting behaviour to metallic regime by method of doping [1, 2]. These unique properties of conducting polymers have encouraged to development of new models to study their observed properties, particularly charge transport mechanism [3, 4]. Therefore conducting polymers have been emerged as a promising class of materials due to their unique optical, electrical and chemical properties. By doping with appropriate class of materials, the conductivity of these materials can be enhanced from semiconducting to metallic values which give new concept of charge transport mechanism. In this research paper conductive polyaniline (PANI) have been studied as it can be synthesised in aqueous media. It also has environmental stability and special optical and other properties.

II. EXPERIMENTAL

Synthesized polyaniline (PANI) using $(NH_4)_2 S_2O_8$) as oxidant in aqueous medium with 0·1 mol of aniline was dissolved in 1 M HCL and its Vanadium pentoxide composites

*Corresponding Author

were obtained by oxidative polymerization .To this reaction mixture, 0·1 M solution of ammonium persulphate was added gradually with continuous stirring for 4–6 h at 0–5°C. As soon the aniline is mixed with the (NH_4) S_2O_8) solution, it turned to be characteristic dark green colour. This indicates that the organic polymerization reaction started immediately. The hybrid of PANI and solid oxide was filtered and then they were washed thoroughly with water. Finally the resultant precipitates were dried in an oven for 24 h so that constant weight is achieved

.

III. RSULT AND DISCUSSION

A. OPTICAL AND DIELECTRIC STUDIES

From UV-Visible absorption spectrum, the optical band gap has been determined using Tauc's relation [5],

$(\alpha h\nu) \propto (h\nu - E_g)^n$ where n=1,2,3,…. (1)

Where n is called index and it has discrete values like 1/2, 3/2, 2 or more depending on whether the transition is direct or indirect and allowed or forbidden respectively. In the direct and allowed cases, the value of index n is 1/2 whereas for the direct but forbidden cases it is 3/2. For the indirect and allowed cases n = 2 and for the forbidden cases it has value 3 or more. In the present case the photon energy (hυ) is plotted against αhν and for n = 2.A straight line is obtained which indicates that the samples obey indirect transition. α is given by,

$\alpha = O_d/t$, (2)

Where O_d is called optical density (OD) which is measured at a particular film thickness (t), and the extinction coefficient (k) is given by

$k = \alpha\lambda/ 4\pi$, (3)

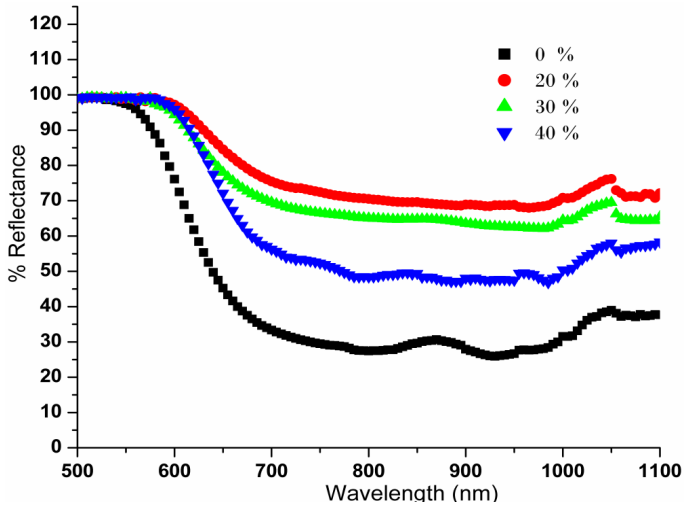


Fig. 1. Plot of reflectance at different weight percentage of V_2O_5 versus wavelength

Where λ is wavelength of incident photon of energy hυ. The optical band gap of doped and undoped samples have been determined by plotting a graph between $(\alpha h\nu)^{1/2}$ versus energy shown in Fig. 1 [5].

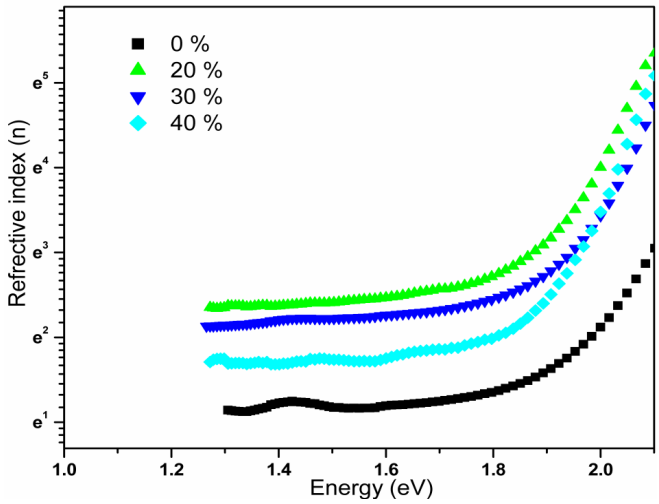


Fig. 2. Plot between refractive index at different weight percentage of V_2O_5 versus energy

From the above observation, the optical band gap has been found to be increasing with increase of weight percentage of V_2O_5. The UV-visible reflection spectrum is given in Fig. 2. From the reflection and absorption spectra in visible region, it has also been measured the refractive index, dielectric constant and dielectric loss .The refractive index (n) has been determined by using following relation

$$R = [(n-1)^2 + k^2] / [(n+1)^2 + k^2] \quad (4)$$

Since interface between the polymer and grain is weak, therefore it leads to decrease in conductivity and dielectric values of the polymer. It is observed that there is no change in the dielectric constant in visible region. The change observed only in UV region [5].

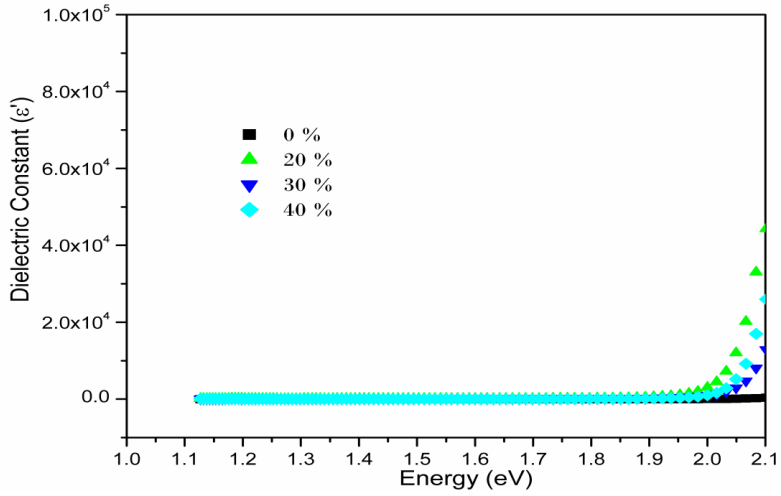


Fig. 3. Plot between dielectric constant at different weight percentage of V_2O_5 versus energy.

REFERENCES

[1]. A. G. MacDiarmid, A. Epstein, *Synthetic Metals.* pp. 69, vol. 85, 1995.

[2]. A. B. Kaiser, C. K. Subramanium, P. W. Gilberd, B. Wessling, *Synthetic Metals*, pp. 69, vol. 197, 1995.

[3]. Kaiser, Liu, C. J. Gilberd, P. W. Kemp, N. T. Wessling, B. Blaszkiewicz, M. McLahlan, D. S. Newnham R.;*Polymer Engineering* & *Science*and *Polymer* Composites. pp. 32, vol. 421, 199,.

[4]. Angelopoulos, M.; Ray, A.; MacDiarmid, A. G.; *J. Epstein, Synth Met.* pp. 21, vol. 12, 1987.

[5]. Kumar, N.; Vadera, S. R.; Singh, J.; Das, G.; Negi, S. C.; Aparna, P.; Tuli, A.; *Def. Sci. J.*, pp. 56, vol. 91,1996.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Development of IoT Based Smart Room using Google Assistant

Ramsha Suhail
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Raza Hasan
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Bhavya Alankar
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Vinita Kumari *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract- *Nowadays smart phones are gradually becoming powerful and could be utilized for development of several applications. This work aims towards development of IoT based smart room using Google assistant with the help of smart phones. In this work room appliances could be controlled using voice command through Google assistant from anywhere so that there is convenience for people at workplace and home and their productivity could be enhanced. For this purpose, a model has been developed in this work to control all the room appliances like light and fan with the help of voice commands using IoT technology.*

Keywords- IoT, Internet, Smart room, Google assistant

I. INTRODUCTION

In the modern stressful life, particularly in metropolitan cities where work pressure is very high, comfortable working and household environment has importance to sustain and enhance the productivity of manpower. For this purpose a working model for voice command enabled smart room has been developed [1]. To design this smart room, we have utilized IoT based smart room using Google assistant [2, 3]. The Internet of Things (IoT) is a system of interrelated computing devices where data transfer occurs over internet without any human interaction [4, 5]. While addressing IoT, our main requirement is a good internet connection. Google assistant is a virtual assistant developed by Google and powered by artificial intelligence which is available primarily on mobile devices. Here, we have used Google assistant to integrate functionalities of a room which could be used to develop smart room with the help of voice control [6].

II. DEVELOPMENT OF MODEL FOR IOT BASED SMART ROOM USING GOOGLE ASSISTANT

We have developed a model for providing smart room monitoring using IoT for comfortable and convenient lifestyle (Fig.1). To implement this model BLYNK application is chosen as it is open source and freely available [7]. Using the BLYNK App we can add different widgets. BLYNK Server acts as a communication link between our Smartphone and hardware both. By BLYNK Libraries we can communicate with the server and process all the incoming and outgoing commands [8].

*Corresponding Author

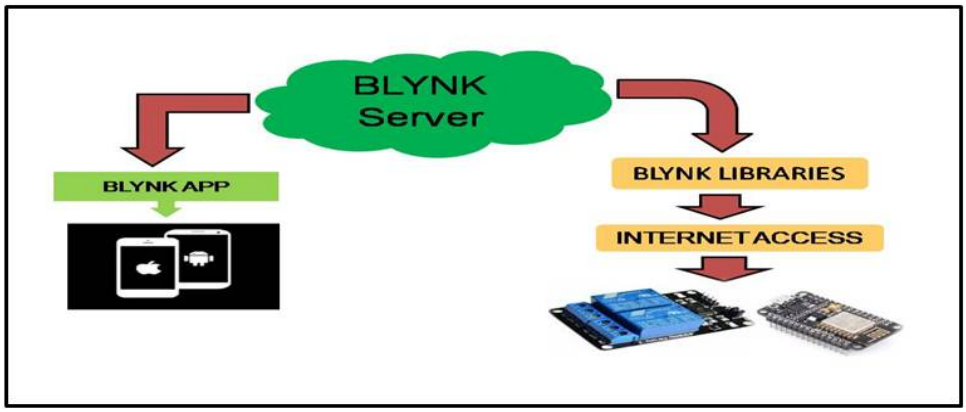


Fig. 1. Model for IoT based smart room using Google assistant

A. EMBEDDED CONTROL UNIT

The embedded control unit in our developed model consists of open source platforms: NodeMCU and Arduino IDE. NodeMCU is an open source platform with ESP8266-12 chips used for developing IOT applications. NodeMCU firmware comes with ESP8266 Development board i.e. NodeMCU Development board [9]. We have used a black colored PCB for NodeMCU Development Board v1.0 (Version2). NodeMCU Dev board has Arduino like Analog (i.e. A0) and Digital (D0-D8) pins, has Wi-Fi capability and it supports serial communication protocols i.e. UART, SPI, I2C etc., through which we can connect it with serial devices like, touch screen displays and SD cards[10]. We have used a well-known IDE known Arduino IDE which uses C/C++ language for developing applications on NodeMCU [11][12]. When we use C/C++ for NodeMCU, it builds binary firmware file of the code we wrote and it writes the complete firmware. We found it better as it makes it easy for the arduino developers than learning a new language and IDE for NodeMCU[13].

B. IOT PLATFORM

BLYNK is an application program available for android and iOS systems which is capable of controlling the hardware devices like Raspberry Pi and Arduino over the internet for IoT applications. It can visualize, display and store sensor data and control hardware remotely. BLYNK has been used in this work to control the room appliances from anywhere using user's voice with the help of Google assistant and IFTTT [14]. In this work, we have added five button as widgets in the BLYNK application to control the fan, light, lamp1 and lamp2 and one slider to control the door. IFTTT is a web tool that puts the internet to work for us and is capable of automating all our tasks for all the internet-connected things. IFTTT is used in this work to create a trigger and connect it to Google assistant (Fig. 2)[15].

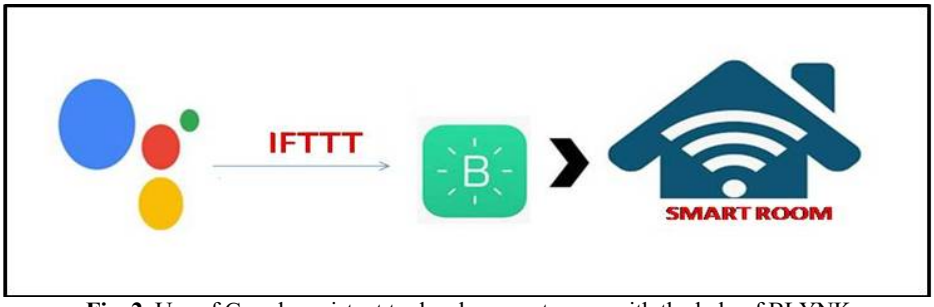


Fig. 2. Use of Google assistant to develop smart room with the help of BLYNK

III. WORKING EXPLANATION

In the developed model, working of IoT based smart room is demonstrated using fan, bulb, lamps and door (Fig.3). These appliances are connected to the NodeMCU to turn on and off the home appliances. Commands are send through BLYNK. For this purpose, voice commands are used which are given to Google assistant which is connected to BLYNK through IFTTT.

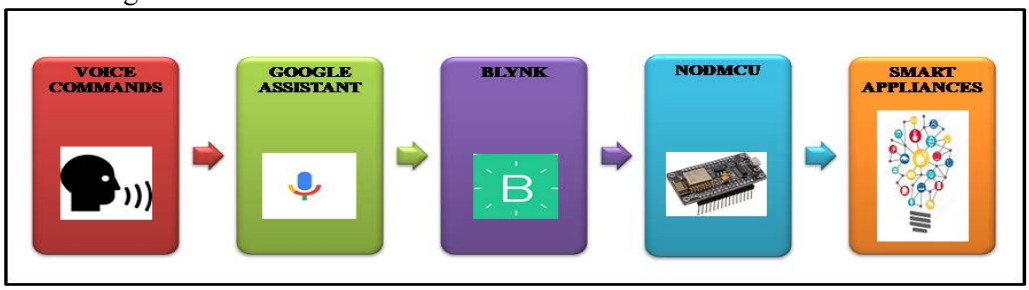


Fig. 3. Block diagram for model of developed system

IV. RESULT AND DISCUSSION

We have successfully developed a working model to control all the appliances in a smart room using voice commands from anywhere. For this purpose Google assistant has been used, which is connected to BLYNK application and further to NodeMCU. We achieved the control the room's appliances remotely using the Wi-Fi technology (Fig.4).

Fig. 4. Developed model

V. CONCLUSION

We have developed a model for controlling the appliances in smart room with the help of voice commands so that our lives are much better and productivity is enhanced. We can conclude that the required goals and objectives of controlling smart appliances like lights, fans, doors, etc using our voice commands through WiFi have been achieved. This work can be used further to control the entire office and house using voice commands.

REFERENCES

[1] Castro R.G, Prez A. G, and Corcho O.: Ready4Smartcities: ICT roadmap and data interoperability for energy systems in smart cities. In: 11th Extended Semantic Web Conference (ESWC14). 2014.

[2] Manjusha C. H., Mounika K.: IOT Based Agriculture System Using Google Assistant. Int. J. Eng. Trends and Appl. 5(2) 225-236 (2018).
[3] Pavithra D. and Balakrishnan R., IoT based monitoring and control system for home automation. In: Proceedings of Global Conference on Communication Technologies (GCCT), pp. 169-173 (2015).
[4] Lecu´e F. , Tallevi-Diotallevi S., Hayes J., Tucker R., Bicer V., Sbodio M., Tommasi P. Star-city: semantic traffic analytics and reasoning for city. In: Proceedings of the 19th International Conference on Intelligent Users Interfaces. ACM, pp. 179-188 (2014)
[5] Greengard S.,The Internet of Things. MIT Press, USA (2015).
[6] C. Floerkemeier et al. (eds.), "The Internet of Things" in First International Conference, IOT 2008, Zurich, Switzerland, March 26-28, 2008, Proceedings. Vol.4952. Springer, 2008
[7] www.BLYNK.cc
[8]. Doshi H. S., Shah M.S., Shaikh U.S.A.: Internet of Things (Io T):Integration of BLYNK for domestic usability. Vishwakarma J. Eng. Res. 1(4) 149-157 (2017).
[9] https://nodemcu.readthedocs.io/en/master/
[10] Kodali R. K., Mandal S., IoT based weather station. In: International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 16-17 Dec. 2016, Kumaracoil, India
[11] https://www.arduino.cc/en/main/software
[12] Plaza P., Sancristobal E. , Carro G., Castro M., and Ruiz E.: Wireless Development Boards to Connect the World. Online Engineering & Internet of Things Lecture Notes in Networks and Systems, Springer International Publishing AG 19-27(2018).
[13] https://devzone.iot-ignite.com/knowledge-base/nodemcu-and-arduino-ide/
[14] Purushottam, Dubey C. K.: Automation by Voice Commands. Int. J. Manag. Tech. And Engg, 8 (V), 285-288 (2018).
[15] www.IFTTT.com

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Importance of IoT in Healthcare: A Review

Jawed Ahmed *
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi
Javed2047@gmail.com

M Afshar Alam
Jamia Hamdard University
New Delhi
aalam@jamiahamdard.ac.in

Bhavya Alankar
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Akshay Chamoli
Department of Computer Science & Engineering,
Jamia Hamdard, New Delhi

Abstract - *In this modern era as technology is growing very fast. We need healthcare system must also improve to the best. The term internet of things (IOT) is a technology focusing on connectivity and data sharing and it has grown up at a large scale. Healthcare systems on the other hand deals to provide a better and quick health for people. Our goal is to combine IOT technology with healthcare systems to provide a better evolution to mankind. These days researches have grown up to a vast extent as now wearable healthcare devices like smart watches, rings and other healthcare related smart devices are available in the market which can instantly monitor one's health and notify about the related health, as for example monitoring one's heart beat or pulse rate or calories burnt during workout sessions etc. These technology are helping in a way as a person can monitor its own health in a smart way. The advantage of these IOT based devices is more to an old person, as in case of emergency the smart devices can signal the hospital or nearby healthcare team which can provide quick aid and any casualty can be saved. In this paper we have surveyed and analyzed the opportunities and challenges hidden behind this technology and the best possible solutions which can upgrade this technology.*

Keywords: *IoT, Healthcare, Cloud Computing.*

I. INTRODUCTION

The things or devices which uses the internet as a medium to communicate and share data with other devices defines IOT (Internet of Things) [1, 3]. This new technology Internet of things has emerged more in these recent years. The logic behind Internet of things is that, this technology will help the IOT inbuilt featured devices to communicate among them and share data as well. The known famous projects like smart city, smart homes are fully based on the concept of this technology [23]. The IOT in these recent years is providing a wide range of solutions for waste management, smart cities, traffic control, healthcare systems and much more.

Healthcare on the other hand is the basic need of a person's life. Good health always keeps a person better in every way. But now days due to unhealthy environment, diseases are growing at a massive rate and medical assistance has become necessary for everyone. But due to busy schedule in a common person's life it is really difficult for a person to take

*Corresponding Author

care of its own health. In this busy schedule life no one has time for a proper food or for a health checkup at regular intervals, so to overcome this a smart way for daily routine body checkup, internet of things concept has evolved to a great success with healthcare. These days devices like fit bit or smart watches [6] have taken the technology to a next level. With these devices people can monitor their own pulse rate or heart rate or how much calories burnt a day and much more of things related to their own body. The advantage is within instant these devices are monitoring a person and is providing a continuous regular second by second body checkup. Likewise, we need many such other devices. Like glucose level sensing or body temperature or heart rate monitoring and prediction of any severe disease like sugar increase at instant or heart attack sudden so precaution could be taken as early as possible [3,5].

The future is like, these healthcare wearable devices inbuilt with IOT feature will be used to analyze and monitor a patient data continuously, and the data collected will get stored in database and will get monitored continuously by experts or healthcare analysts [4]. It will be useful as a patient can get the cure of his/her illness without visiting a doctor or hospital and in the case or urgency or emergency this can be the best help for technology as well as for people.

II. RELATED SURVEY WORK

Wireless sensor network plays an important role in healthcare systems as for the device to communicate or to fetching data so for this technology to hold the energy consumption is always less while the communication coverage improves [13]. Due to smart healthcare facilities the technology is improving as well as health care is available at any time reducing casualty [14]. Many devices like wearable are present to monitor health of a person. This paper deals with device for heart with a feature inbuilt that as it reaches a specific threshold value set or marked and predefined in the device it will automatically switch on the alarm notifying about the heart problem for a person [12]. IOT enabled feature with healthcare for smart homes with the help of IOT layered approach [15]. Sensors created to monitor old patients body activities continuously so that for any urgent emergency the cure should be given instantly [16, 17]. In this IOT architecture with healthcare using temperature sensor, ECG Sensor so that remote monitoring becomes easy [18].

In this paper BSN termed as body sensor network is said to be one of the important sensor for health care management. The concept is as the patient is monitored with this BSN. This BSN is a collection of low powered and light weight and is used for monitoring body parameters [19]. Cloud is used for sharing and storing data virtually and as data generated by healthcare is massive as well as the information of a patient must be kept secure so that no data leak or alteration on original data must not take place. The new technology object hyper linking is discussed in this paper which aims as internet getting extended and a large amount of data always flows in a network but will reside in the cloud safe and fetching of data from cloud also requires high processing power in a reliable environment [20].

The aim to make IOT compatible with e-healthcare so that the e-healthcare applications work fine with IOT technology. The e-healthcare application consist of medical health information which must be transferred to cloud or to the data storage center but this can only work smooth if IOT technology gets compatible with e-healthcare [21].

As IOT with healthcare technology is present data analysis plays an important part as to predict what next or after continuous health monitoring by data analysis one can predict about the health of a patient and this can lead to quick casualty handing. for example if an elderly person's heart rate pulse is not normal by checking the results we can predict the disease and treat that person before any major casualty take place. In these cases only real time monitoring helps but the major security challenge as no data should go undetected and get alter [22].

The papers suggests that the monitoring of IOT enabled systems is controlled by some unit which use to take decision for IOT monitoring devices. As medical data is under

continuous monitoring and even a slight or major error can get detected according to the data generated and in case of any critical situation the device needs to send alert signals or should communicate with the emergency [23].

Wearable devices like smart watches or fitness bands have a feature to monitor human body like counting pulse rate or burning calories or heart rate monitor. Now as virtualization has begun these smart watches also have the capability to monitor a human body with generating a health report for that person. Higher technologies deals with eating pills having sensors which can possibly generate one's sugar level or blood pressure [24].

III. IOT HEALTHCARE DESIGN

The IOT healthcare design is the backbone of this IOT enabled technology and the architecture of this design deals with three layers

- **Topology**: it deals with the physical configurations and the term topology refers to arrangement of different IOT elements.
- **Architecture**: the hierarchical model or the actual design of the system.
- **Platform**: the framework, environment used to implement this feature.

Communication standards:

As communication is needed so that devices are able to interact among them.

The communication protocols are of short range and long range.

Short range communication: as wearable healthcare devices present, short range communication is mostly used among devices nearby. The main feature of these devices they are of low energy so communication as well as energy efficient.

- **Bluetooth low energy**: for small communication among devices these are best. Generally used in star topology. As analyzed the range is 150m in an open area. Can give upto 1mbps and low latency of around 3ms. The consumption of power within this technology is very minimal making it efficient to use [8].

- **Zigbee**: it is also a low cost, low power technology for short range communication. The standard for this is IEEE 802.15.4 and generally used with mesh topology. The range is up to 30m but the modified or can say the new zigbee technology termed as zigbee pro 900 xsc has range up to 90m. The feature of zigbee as it can operate within a range of frequencies like 868 MHz, 900 MHz or 2.4 GHz depending on the chosen module [7].

Long range communication:

These are wide area networks (LPWANS) with high range so can cover distance without interference in communication. Can go up to several kilometer. These are longer in range and more power with respect to Bluetooth or Wi-Fi.

- **SIGFOX**: it is simple with limited feature. It generally works with star topology the best part as it has a feature can detect weak signals making signals to travel at longest distances. The maximum range it has is around 9.5k with bandwidth of 100bits per second. In free or can say rural areas the distance can go up to 50km. As SIGFOX deals with high network capacity it can support around 50000 nodes. Research say can go up to 40 devices per home or within one connectivity.

- **LORA** and **LORWAN**: in this the physical layer deals with LORA and network layer deals with LORAWAN. As LORAWAN is a network layer protocol provides low power, long range communication. Generally star topology suits this technology. Each gateway supports up to 40000 nodes. The range may go up to 7.2km with the bandwidth .25-5.5kbps [9].

- **NB-IOT**: it deals within the licensed bands of LTE or GSM with the feature of low power consumption with long range communication. The feature as it can go up to a distance of 15km with around 250kbps with a max connectivity of 50.000 above nodes [10].

IV. IOT HEALTHCARE WITH CLOUD COMPUTING FEATURE

Cloud technology is useful for high storage of data and in case of IOT with healthcare data flow is continuous among devices so cloud can be stated as a secure area for data storage as well as for data fetching. The data processing and data storage are the key main feature of cloud.

Healthcare with cloud computing

These days ongoing research on how cloud can be more beneficial with healthcare. The basic three services provided within cloud for healthcare environment:

Software as a Service (Saas): it is also termed as application service provider which provides an application on which healthcare system can run, configure and end users can check, analyze the data within the software provided for communication devices.

Platform as a Service (Paas): in this the tools are provided for management of data or virtualization of data or any other feature that can be handled in Saas. .

Infrastructure as a Service (Iaas): in this the hardware is provided for storage of data or devices like servers and much more.

Internet of things healthcare benefits

- Communication: IOT helps to communicate among devices as we can say machine to machine communication and due to this thing, the devices are able to communicate among each other and with this the full transparency is maintained with efficient qualities.

- Automation and control: As IOT deals with communicating with devices among themselves without human intervention so this technology is faster, quicker and leading to timely output.

- Remote health monitoring: with the analysis of monitoring we can provide a good direct help to patient's in need as we know due to lack of care many patient's won't get attention or care as it is not possible to take care of every patient at every time but with the help of iot devices inbuilt with sensors each and every data of a particular person will be shared by healthcare providers so that any emergency can get cured without any delay.

- Remote medical assistance: with the remote assistance any emergency can be delayed as the patient can directly communicate with the expert via mobile application provided and can get assistance about do's and don'ts to follow as in case of health issues. As the patient continuous health report is being checked, the patients are continuously monitored by experts so in case of emergency remote medical assistance can also play a vital role in saving a person's life.

- End to end connectivity: with the new healthcare monitoring solution IOT can manage and analyze patient care workflow. The feature of IOT enables for interoperability, data exchange, and machine to machine communication.

- Data management and analysis: as a vast amount of data continuous will get shared by the wearable devices in a short time so to manage and analyze and then after generate a proper report about a person's health. IOT device can manage, collect and generate the report of the data fetched to it and they are less prone to errors.

V. INTERNET OF THINGS HEALTHCARE CHALLENGES

- Data privacy is the key component
  The IOT inbuilt devices must capture vital data and transmit it in real time so there must be a way to secure data while transferring so that the cyber-criminal must not misuse patient's data.

- High amount of data and accuracy maintained:
  As a very Hugh amount of data must get collected to the database so the maintenance of data as well as the accuracy of data both must be maintained. No replica of data should be present.

- Cost effective: as to monitor a person's health iot based healthcare devices are available in market but it is not that much cost efficient so the devices made must be cost efficient so that everyone can get the benefit of this technology at a large scale.

- Standardization: set of rules present and globally certified so with different iot devices having different feature of different type but out of all the complexities available within they are able to communicate with each other.

- Platform or domain free: this means the application running on android interface or mac operating system or Linux or any other but have the feature to interact among each other. Generally devices are platform specific and communicate with the same platform devices only but in case of iot healthcare devices the technology should be platform free.

- Scalability: IOT networks as well as services related to application and database must be scalable because as technology gets modify complexity increases and creates a trouble for end users.

- The model related to business: The IOT healthcare model must be robust as always new requirement, new infrastructure new policies and as ever time the continuous flow of data would be present or data sharing within devices so every time some new data or information will surely flow so to understand and analyze accordingly.

- Quality of service: as the data sharing among devices will flow continuously and will be highly sensitive so the reliability factor must be strongly taken care of as the data sharing must involve privacy with a sufficient bandwidth to interact anytime.

VI. INTERNET OF THINGS HEALTHCARE APPLICATIONS

- Fit bit wearable's: like smart watches which can analyze a person's heart rate, pulse rate, calories burnt and much more, with these devices people can monitor their own health.

- Ingestible sensors: They are sensors in a shape of a pill which can purely manage the medication in the body and notify if detects any irregularities inside our body.

- Moodables: These are head wearable devices like a head phone or ear phone which can change your mood according to your body need. If for example your body temperature is getting high that device will try to calm a person down by giving low intensity current waves to the body to make a person relax.

- Hearable: The modern age hearing aid which can connects with the Bluetooth and syncs with the smartphone. The best part is it allows you to filter, analyze sounds of the real world.

VII. CONCLUSION

As we survey the papers with the combined technology of Heath care with IOT. We observed that this technology can have a great impact on mankind. Health related issues will not get resolved faster and due to data analysis and with data results prediction a person can be cured before getting a disease.

Some facts like we can standardize some of the IOT technologies which we feel work well with these technologies like for short range communication. Zigbee technology is the best while for long range communication NB-IOT must get standardize. As fit bit band as well as smart watches works well but are not cost effective so the new healthcare devices must be cheaper, easily available so that even a common person can also be able to avail this technology.

The security aspect is also one of the major issues in this technology and it is in continuous process to make this technology a fully secure hub.

REFERENCES

[1] y. Yin, y. Zeng, x. Chen, and y. Fan, "the internet of things in healthcare: an overview," journal of industrial information integration, vol. 1, pp. 3–13, 3 2016.
[2] l. Atzori , a. Iera , g. Morabito , the internet of things: a survey, comput. Netw. 54 (2010) 2787–2805
[3] s. M. R. Islam, d. Kwak, h. Kabir, m. Hossain, and k.-s. Kwak, "the internet of things for health care : a comprehensive survey," ieee access, vol. 3, pp. 678 – 708, 2015.
[4] d. V. Dimitrov, "medical internet of things and big data in healthcare," healthcare informatics research, vol. 22, no. 3, pp. 156–163, 7 2016.
[5] t. C. Arcadius, b. Gao, g. Tian, and y. Yan, "structural health monitoring framework based on internet of things: a survey," ieee internet of things journal, vol. Pp, no. 99, p. 1, 2017.
[6] fitbit, "fitbit pure pulse," 2017. [online]. Available: https://www.fitbit.com/au/purepulse
[7] digi, "zigbee wireless standard," 2017. [online].available: https://www.digi.com/resources/standards-and-technologies/rfmodems/zigbee-wireless-standard
[8] bluetooth, "bluetooth low energy security (white paper)," 2017.
[9] lora alliance, "lorawan: what is it? A technical overview of lora and lorawan (white paper)," 2015. [99] r. Miller, "lora security - building a secure lora
[10] huawei, "nb-iot (white paper)," 2017.
[11] ashton, that ''internet of things'' thing, rfid journal (2009).
[12] ms. Supriya chandrakant padwal m. E. (vlsi & es), dept. Of e&tc sahyadri valley college of engg. & tech. Pune, india. Mr. Suraj vishnu kurde department of computer engg. Sahyadri valley college of engg. & tech. Pune, india "longterm environment monitoring for iot applications using wireless sensor network", february 2016.
[13] rashmi singh "a proposal for mobile e-care health service system using iot for indian scenario", 1 january 2016

[14] ashly mathew, farha amreen s. A, pooja h. N, aakriti verma "smart disease surveillance based on internet of things (iot)" international journal of advanced research in computer and communication engineering vol. 4, issue 5, may 2015.
[15] lei yu school of computer and information, hefei university of technology, hefei, china 2 school of medical information technology, anhui university of traditional chinese medicine, hefei, china email: fishstonehfut1006@163.com yang lu, xiao juan zhu school of computer and information, hefei university of technology, hefei, china email: luyang.hf@126.com, xjzhu@aust.edu.cn "smart hospital based on internet of things", 10 october 2012
[16] s. M. Riazul islam, (member, ieee), daehan kwak, md. Humaun kabir, mahmud hossain, and kyung-sup kwak, (member, ieee) uwb wireless communications research center, inha university, incheon 402-751, korea networked systems and security laboratory, rutgers university, newark, nj 07102, usa 3 secure and trustworthy computing laboratory, university of alabama at birmingham, al 35233, usa "the internet of things for health care: a comprehensive survey", 1 june 2015.
[17] ashok khanna business head uk, eu and major accounts, life sciences (eis) global head, presales and solutions for life sciences (eis) prateep misra research area manager, tcs innovation labs, "the internet of things for medical devices prospects, challenges and the way forward", march 2017
[18] sreekanth k u, nitha k p "a study on health care in internet of things" international journal on recent and innovation trends in computing and communication issn: 2321-8169 volume: 4 issue: 2 044 – 047, february 2016.
[19] prosanta gope and tzonelih hwang "bsn-care: a secure iot – based modern healthcare system using body sensor network" ieee sensors journal, vol. 16, no. 5, march 2016.
[20] haobijam basanta, yo-ping huang and tsu-tian lee,"intuitive iot – based h2u healthcare system for elderly peopleieee 13th international conference on networking, sensing and control mexico, april 2016.
[21] amna abdullah, asma ismael and aisha rashid "real time wireless health monitoring application using mobile devices", international journal of computer networks and communications, vol. 7 no.3, united arab emirates, may 2015
[22] arijit ukeil, soma bandyoapdhyay and chetanya puri, "iot healthcare analytics: the importance of anomaly detection" ieee 30th international conference on advanced information networking and applications, india 2016
[23] punit gupta, deepika agrawal and jasmeet chhabra, "iot based smart healthcare kit" ieee international conference on computational techniques in information and communication technologies, india 2016
[24] sapna tyagi, amith agarwal and piyush maheshwari "conceptual framework for iot- based healthcare system using cloud computing", ieee 6th international conference on cloud system and big data engineering, india 2016

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Anticipating Depression and Stress Exploration using Regression Analysis with High Order Artificial Neural Networks

Ramachandran VG
Jamia Hamdard
Department of Computer Science & Engineering
New Delhi, INDIA
Siddhartha Sankar Biswas *
Jamia Hamdard
Department of Computer Science & Engineering
New Delhi, INDIA
Hareendran VG
ELCOME International
Dubai Investments Park 598-1121
PO Box 1788, Dubai, UAE
Safdar Tanweer
Department of Computer Science & Engineering
Jamia Hamdard
New Delhi, INDIA
Sherin Zafar
Jamia Hamdard
Department of Computer Science & Engineering
New Delhi, INDIA
Satarupa Biswas
Department of Computer Science & Engineering
Jamia Hamdard
New Delhi, INDIA

Abstract - *Nevertheless, present high-tech epoch is building natural life more contented and gifted us with a lot of irrevocable health complications. With the development of advancing medical field, the digital confirmations have made Medic-Technicians so easy for confirming medical complications with in no time rather than predicting the same. Early detection of mental disorders became very much essential as the issues related to Stress and Depression with appropriate and correct controlling by quelling the convolutions. The same can be analysed by using various techniques like Regression Analysis for the Knowledge Discovery and clustered with the help High Ordered Artificial Neural Networks (HONN). Using software installed over mobile phones, we can capture the vocal clips and the same can be ported to a custom databases using Big data, a growing stint that terms a huge capacity of structured, semi-structured and unstructured records that has the potential to be extracted by Datamining using High Ordered Artificial Neural Networks. Further, it can be forecasted by reverting back to users indicating that he or she might have chances of falling in Stress/ Depression. Also, it can also be implemented for parents to monitor kids using parenting-control.*

.

Keywords*–Data Mining, High Order Neural Networks.*

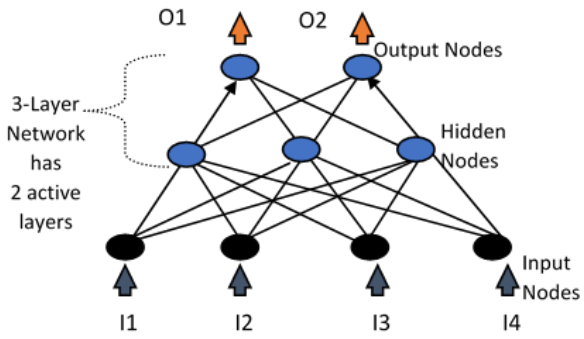
I. INTRODUCTION

*Corresponding Author

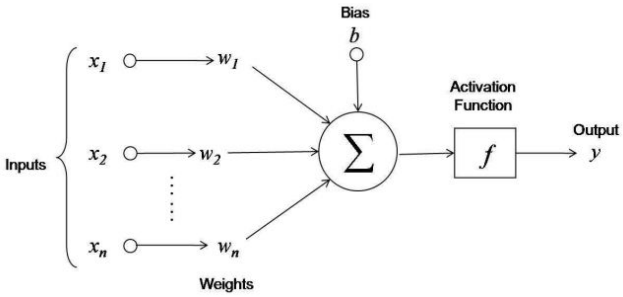
Data mining (DM) is the process trusting on discerns and machineries from the intersection of database management, indicators, and machine learning which is required to have a process and draw conclusions from a large raw data. Using 'Tracking patterns', patterns of the datasets can be recognized which can be aberration in the information ensued at even intermissions, or recede and current of a definite variable over period. 'Classification' is a more compound data mining technique that services enables to assemble various attributes together into distinct groups, which can use to draw further assumptions, or assist some purpose [1, 2, 4] . 'Association' is correlated to tracking patterns, but is more specific to feebly allied variables and further procedures or characteristics are to be correlated with another event or trait. 'Outlier detection' at several cases, just distinguishing the overarching pattern can't give a clear thoughtful of the data set and identify irregularities, or outliers in the data. 'Clustering' is very analogous to classification, but encompasses grouping chunks of data together based on their resemblances. 'Prediction' is one of the most appreciated data mining techniques as now a days that is used to venture the types of data in the imminent. In many cases, just diagnosing and accepting antique trends is enough to chart a somewhat accurate forecast of what will happen in the upcoming. 'Regression'used mainly as a form of preparation and moulding and that is used to recognise the possibility of a certain variable in the existence of other variables. Regression analysis is an easy process of crafting resolutions or a traditional from the input variables to outcome in favour of retort variables set which is incessant in nature [1, 3].

II. DATAMINING TECHNIQUES

The most common practises: Artificial neural networks: Non-linear analytical reproductions that learn over training and look like biological neural networks in building. Decision tree: Tree-shaped constructions that epitomise sets of choices. Decisions engender rules for cataloguing of datasets. Exact approaches contain Classification, Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID) [5, 6]. Genetic algorithms: Optimizing the procedures which use process such as genetic mixture, transmutation and natural assortment in a design based on the notions of progression.



[Figure- 1 : Schematic Diagram of ANN]



[Figure- 2 : ANN Diagram with Mathematical Expression]

III. APPLICATION OF ARTIFICIAL NEURON NETWORK

An **ANN** is a computational design to simulate the way the human brain analyzes and processes information and is a way of finding Outputs from various Hidden layers **[Figure 1]** through some given Inputs. ANN is the capability to derive gist from a problematical or vague data that can be used for pattern extraction and analyses trends those are too multifaceted to be noticed by either humans or other usual machines [7, 8].

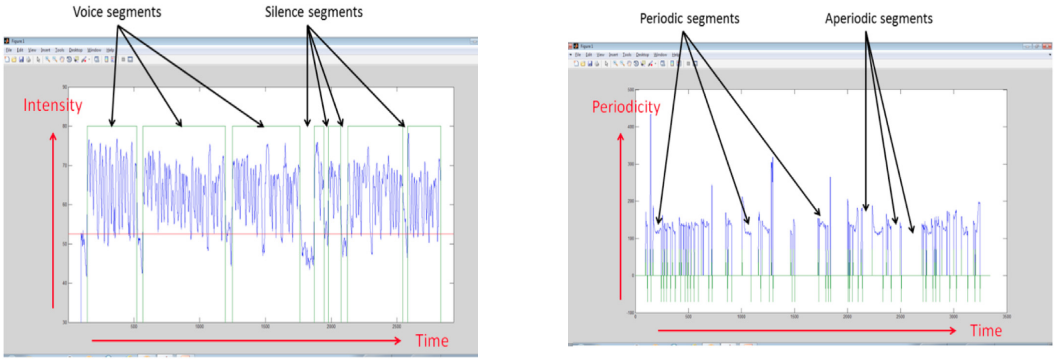
In order to analyse a process patterns which may hidden in nature of a available raw data conferring to diverse perspectives for tagging into useful facts, which are actually required is collected and lump together in common areas like data warehouses. In order to have an proficient data mining algorithms, facilitating intelligent business decision making and other information necessities to ultimately cut total costs & increase revenue **[Figure 2]** ,'Neural networks' can be used as a complex data mining technique. Using a **Neural Network (NN),** it can crack a major range of teething troubles, most of which includes in finding trends of large amounts of raw data. At many occasions, a human brain is very noble at for cracking highly complex teething troubles when paralleled with traditional computer architecture. Natural **Neurons** Cells are the handling elements of a biological neural network where **Nucleus** acts as the vital processing portion of a neuron. **Dendrite** is a part of that biological neuron that be responsible for inputs to the cell processing through Neurons [5, 9].

IV. PREDICTION OF DEPRESSION/ STRESS/ DEMENTIA

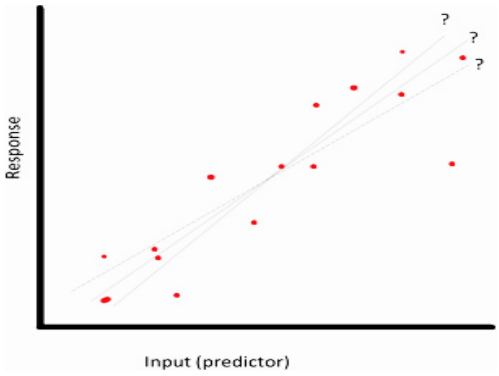
Depression is a form of decline in mental illness to perform everyday activities andcauses enlightened brain disorder which slowly extinguishes memory and thinking abilities, and can be of irreversible in nature. It gradually reduces the ability to carry out the simplest tasks and symptoms appear initially in the mid-30. In recent clinical evidences of past evaluations indicate that the disorder will leads to Heart disease and Cancer, further can take away the life of older people very easily. A common loss of cognitive functioning such as thinking, remembering, and reasoning that affects behavioural abilities. In the course of the pre-clinical stage, people give the impression to be symptoms- free, but harmful fluctuations were taking place in the brain. The certain drop in non-memory physiognomies of perception issues like word-finding, three-dimensional trepidations, and meet halfway reasoning or decision and vision problems that can point to the very primary phases of dementia. Using the biological signs of disease found in brain images, cerebrospinal fluid, and blood to see if it can be perceived at primary changes in the brains of people with MCI. The damage transpires in areas of the brain which is responsible to instrument language, reasoning, carnal processing, and cognizant- thought. Also, Memory loss and muddles grow worse, and people initiate to have snags to recognize family and colleagues friends etc.

V. CONCLUSIONS

Recording protocol: Persons can have a custom application installed on their mobile. After successful installation, he can have a registration with his personal data such as age, sex, profession, address and selection of language with a complete speech recognition Test. Further, according to the choice, voice will be recorded either during calls or on a regular basis.



[Figure 3 a and 3 b : Recorded Vocal Sound]



[Figure- 4 : Variation Analysis of facts]

With the help of voice recorded [Figure 3a and 3b], we can have many tasks like Voice Modulation Test, Speech Continuity, counting various backward task, various sentence repeating tasks, description of an image, and fluency test over verbal etc. can be achieved. People with similar age range, similar profession, living in similar areas, similar type of food habit, having similar type of diseases, similar reactions due to similar type of medication etc. can be analysed by implementing ANN. Various graphs can be generated. Audio features and analysis: Plentiful vocal features can be mined from each spoken chore and it can include the specific prediction of Dementia or other Neurological disorders [Figure.4] shown below.

Presently, language- sovereign technology development, speech recognition was not included, and only non- verbal features were aimed. Duration of (a) Voice segment, (b) Silence segment length, (c) Periodic segment length, (d) aperiodic segment length are measured in seconds.

The numerous data produced can be collected into database with the synchronizing of the application installed mobile phone/ handheld set. The mean of the intervals have to be derived as vocal sound may for longer intervals and periodic segment spans and squatter spells for the silence and aperiodic segment lengths. Further, we can calculate the vocal sound topographies of the whole sentence(s), reiterating chore for each person to figure the mean and standard Deviations (SD) of the verbal sound trials across the diverse stretch pairs. These structures and the use of vibrant time Distorting to derive the vocal sorts from pronounced cognitive tasks are also unique to the field of talking scrutiny.
Voice against hush segments and episodic versus aperiodic segments of a classic obvious speech of soundtrack. The horizontal axis labels time frames of 10 ms; the vertical axis on the left labels the signal concentration and that on the precise labels the signal periodicity. Voice against silence and periodic counter to aperiodic were stubborn from the trampled intensity and periodicity delineations, harmoniously.

VI. REFERENCES

[1]. To assess the patients Automatic Speech analysis who seems to have a chances of pre-dementia disease. Monitoring, 1(1), 112–124. https://www.elsevier.com/ Authors: König, A., Satt, A., Sorin, A. R., A. David Toledo-Ronen, O., Derreumaux, R. Hoory, (2015). DOI: 10.1016/ J.dadm. 2014.11.012. Audio classification and segmentation Content analysis, IEEE Volume: 10, Transactions on Speech and Audio Processing Issue: 7, Oct 2002), Page(s): 504 – 516, DOI: 10.1109/TSA.2002.804546.

[2]. Data Mining, 2008 International Conference on Computer Science and Software Engineering - Application of New Adaptive Higher Order Neural Networks, 2-14 Dec. 2008, IEEE DOI: 10.1109/CSSE.2008.897.

[3]. Automated financial data modelling with Neuron-adaptive higher order neural-network models, IEEE Transactions on Neural Networks (DOI: 10.1109/72.977 07 August 2002 Volume: 13, Issue: 1, Jan 2002), Page(s): 188 - 204 (Volume: 6, Issue: 2, Mar 1995), Page(s): 422 – 431, Mar 1995 of IEEE Transactions on Neural Networks DOI: 10.1109/72.363477.

[4]. Identification of dynamical systems with High-order neural network structures, DOI: 10.1109/72.363477, (Volume: 6, Issue: 2, Mar 1995), Page(s): 422 – 431, Mar 1995, IEEE Transactions on Neural Networks

[5]. Data mining through Neural networks, 2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing DOI: 10.1109/ICIEAM.2016.7911596 (ICIEAM), 19-20 May 2016, IEEE.

[6]. Artificial Neural Network by using photovoltaic and atmospheric data with the performance comparison of Multiple Linear Regression, Random Forest. IEEE DOI: 10.1109/EMES.2017.7980368., 2017 14th International Conference on Engineering of Modern Electric Systems (EMES),

[7]. Dataset containing Prostate Cancer details with Data Mining through Artificial Neural Networks, IEEE DOI: 10.1109/NEUREL.2006.341201 (8th Seminar on Neural Network Applications in Electrical Engineering Kenneth Revett, 2006), 25-27 Sept. 2006,

[8]. https://www.alz.org/.

[9]. Using High-dimensional clinical trial data by Jim O' Donoghue A Configurable Deep Network; Martin Van Boxtel Mark IEEEDOI.10.1109@ IJCNN. 2015. 7280841.Roantree.2015 International Joint Conference on Neural Networks (IJCNN) Year: 2015 Page: 1–8.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Internet of Things: A Review from Cloud Perspective

Tabish Mufti *
Department of Computer Science and Engineering,
Jamia Hamdard, New Delhi, India
tabishmufi@jamiahamdard.ac.in

Shahab Saquib Sohail
Department of Computer Science and Engineering,
Jamia Hamdard, New Delhi, India
shahabsaquibsohail@gmail.com

Deepak Kumar
Amity institute of Information Technology
Sector-125, Noida, India
deepakgupta_du@rediffmail.com

Abstract — *In today's world of advance technology, we are coming across various new paradigms of technology. IoT is one of the most talked about among them in the industry. Internet of Things is influencing our lifestyle and is becoming fastest growing technology.IoT provide infrastructure for real time objects and also help in keeping track about these objects. In IoT devices are connected smartly so that they can share data, resources with other machines. IoT uses various types of sensors embedded in various devices which emit data. These sensors share data using IoT common platform. These platforms collects data from various sources and then further analytics are performed on data and essential information is extracted finally the result is shared. This research article comprises of meaning of IoT, characteristics, basic requirements of IOT and its applications. The main objective of this paper is to provide overview of the evolution and usage of Internet of Things (IoT), its architectures and benefits as well as disadvantages.*

Keywords — *Internet of things, sensors, Architecture, smart devices, Radio frequency identification, web services*

I. INTRODUCTION

Internet of Things (IoT) is the buzz word in all academic and industry quarters of sciences and technology. In general sense, it represents the capacity of network devices to logically sense and systematically collect data from various sources around the world and then share this data across internet[1].Then the shared data is further processed and utilized for other useful purposes. The IoT is a gamut of smart machines communicating with other smart machines, objects, environments and infrastructures. In today's modern digital world each individual is connected with every other individual using sundry connecting and communication devices, wherein the most popular mode of communication is Internet. Therefore it is the internet that connects people around the world and thus IoT [2] becomes the focus for deriving the underlining behaviours, information, trends as well as patterns through the usage of internet. The basic idea of IoT has prevailed from around two decades.

*Corresponding Author

It has attracted many academicians, researchers as well as industrialists because of its huge impact in improving the daily life and society. When things like smart household appliances are connected to a network they get improvised to provide ideal service as a whole. Some mundane things which is possible with IoT is a smart residence with automatic windows that can be opened and closed and respond when the gas burner is turned on by automatically opening. The air Conditioner can be switched from the car and lights could be controlled by using internet. This kind of an environment is especially more useful for persons with disability and moreover the ultimate arrangement of devices as a system instead of individual units.

The theoretical basis of network of smart devices was first applied in 1982 to a coke vending machine at Carneige Mellon University as the first digital appliance reporting its stock of bottles and the temperature condition of the drinks. After reading various sources[3] ,it was found that the book "The Computer of the 21st Century" written by Mark Weiser in 1991 as well as Academic quarters like Ubicomp and Percom designed a contemporary vision of IoT. Reza Raji, a researcher in 1994, described the concept of IEEE Spectrum as "moving small packets of data to a large set of nodes to integrate and automate everything from home appliances to entire factories." From 1993 to 1997, several companies like Microsoft at Work (MaW) and Novell's proposed solutions based on a similar platform. MaW was a small project promoted by Microsoft to bring together common business machinery, like fax machines and photocopiers, using a common communications protocol permitting control and status information to be pooled with computers running Microsoft Windows. The idea gained popularity when Bill Joy envisaged D2D (Device to Device) communication at the World Economic Forum at Davos, Switzerland in 1999[4].

Manpower requirements are urgent in every organization for the information desk to each and every department. To provide information, advertisements, messages and other notifications for the customers and the staff the information desk plays a crucial part. Due to IoT this function and manpower role has been cut down and replaced by smart devices. This has been a major achievement especially in cost cutting, updating of information for prompt services and better and efficient utilization of resources.

II. LITERATURE SURVEY

Kevin Ashton, a British technological pioneer coined the term 'Internet of Things' to support the idea of supply chain management in 1999. However, in the past few years the term has become more comprehensive and now includes wider spectrum of services like Healthcare, Transport, Utilities, Consumer goods etc[5]. The connotation of the word "Things" has changed due to changes in technology but the purpose and goal of computer sensing information without any human aid has remained the same though.

Some technologies that supplements and promotes 'Internet of Things' are:

i. Near-field communication and Radio Frequency Identification (RFID) - Near Field Communication short range connectivity protocol that enable communication between two devices during 2010 NFC became more popular on the other hand In the 2000s, RFID technology uses radio waves to spot the objects.

ii. Quick response codes and Optical tags - QR code consist of data. This is low cost tagging technique. Phone cameras decipher QR code using image processing techniques.

iii. Bluetooth and low energy - This is the latest high speed, low powered wireless technology which is designed to unite smart devices or gadgets with other portable apparatus together.

The authors in [6] describes the concept of Internet of things along with the architecture of IoT, protocols to develop IoT Architecture and challenges for developing Intelligent system for real time environment. The authors [7] describes smart urban Ecosystem which includes smart cities environment ,applications and infrastructure .Integration of cyber and physical component to control and monitor urban environment. In literature present in [8], [9] describes the concept of automatic smart parking system by using IoT. Smart parking will be using cloud services for storing information about various vehicles along with their IN-OUT time, number

of parking slot, number of parking slots available. Components for smart parking will include Raspberry Pi, Camera, IR sensors, Display device, User device etc

In [10], [11] the author describes the IoT along with new paradigms like fog computing and edge computing.fog computing is decentralized computing infrastructure which means processing is done closer to the node where data is created[12].IoT technology has great potential; it can help in cost reduction and supports new business models. IoT is channelling itself in all the developed and emerging markets globally. Companies like Samsung, LG, Qualcomm, Intel etc. The Industrial Internet of Things (IIoT) market is predicted to reach $123B in 2021 reach a CAGR of 7.3% by 2020.Top 3 IoT projects in progress are Smart Cities (23%), Connected Industry (17%) and Connected Buildings (12%) according to Forbes 2018 report. Various IoT analytics have set up half of smart cities projects in Europe with 45% in America and 55% of global projects[13].

Major advantages of this technology:

- Access Information - data can be accessed from remote locations
- Communication - effective communication is possible via connected devices
- Automation - task done without human intervention

Major disadvantages of this technology:

- Complexity- A diverse devices connected to a network single loophole can effect entire network
- Privacy/Security - In today's tech world where all the devices are connected to internet, Loss of data is possible
- Loss of Jobs - automation leads to loss of jobs

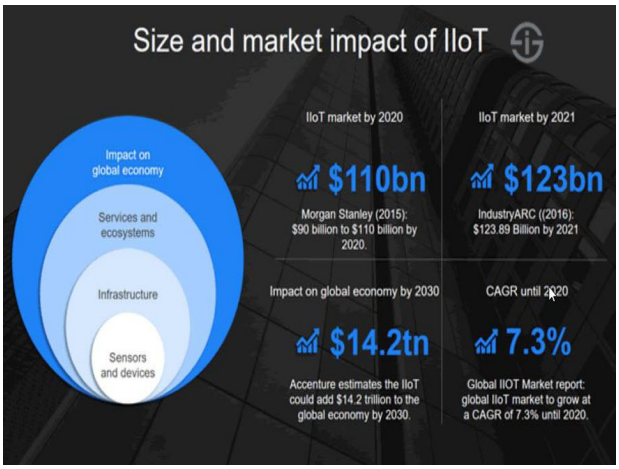
Applications of IoT technology:

As Internet of Things (IoT) in 2019 is ready to rule world, its cost efficient feature has enabled new business models. There are various areas where IoT is being used. Some of them are listed below:

- **Smart Home**
- **Smart Cities**
- **Wearable's**
- **Connected Cars**
- **Industrial Internet**
- **IoT in agriculture**
- **Smart Retail**
- **Energy Engagement**
- **IoT in Healthcare**
- **IoT in Poultry and Farming**
- **Ground water detection and water reservation**

III. FUTURE OF IOT:

The future for IOT is very scalable and bright. Most of the developed countries are investing billion dollars to convert the existing infrastructure in Smart Infrastructure. The Industrial Internet of Things (IIoT) market is predicted to reach $123B in 2021, attaining a CAGR of 7.3% through 2020 according to a recent Forrester survey of 2018.The following figures show the year on rise of the use of Industrial IoT worldwide as predicted by state of the art analytics software Statistic.



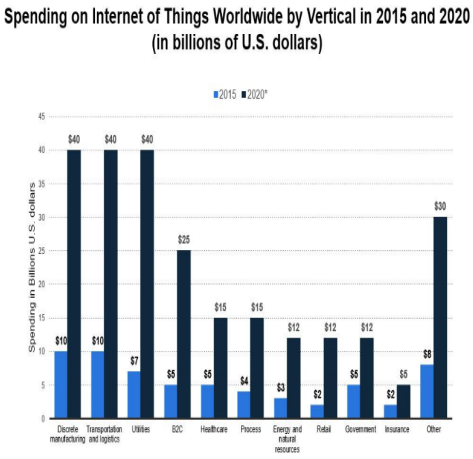


Fig.1 Source: The Industrial Internet Of Things Fig. 2 Source: Statista The Business Guide To Industrial Iot (IIoT)

IV CONCLUSION

IoT promises of an improved quality of human life and productivity of enterprises. It has the potential to enable extension and advancements of fundamental services in health care, transportation, logistics, security, education through widely distributed and locally intelligent networks of smart devices and robust ecosystem of application development. Although, substantial efforts are required to mobilize the industry to move beyond the early stages of market development towards market maturity by unleashing the hidden opportunity offered by IoT. The market can place differing demands on the mobile networks with regard to service distribution, customer-charging model and capacity to deliver IoT services etc. which can pose a challenge to the mobile service providers. The pieces of technology puzzle are coming together to welcome IoT sooner than most conservatives expect. Just as it was not very long ago the internet became a household name within few years and www became a necessity, the Internet of Things will also touch every aspect of human life sooner than we can imagine.

REFERENCES:

[1] P. Asghari, A. M. Rahmani, and H. H. S. Javadi, "Internet of Things applications: A systematic review," *Comput. Networks*, vol. 148, pp. 241–261, 2019.

[2] X. Cui, "The internet of things," in *Ethical Ripples of Creativity and Innovation*, Springer, 2016, pp. 61–68.

[3] S. S. Sohail, J. Siddiqui, and R. Ali, "An OWA-based ranking approach for university books recommendation," *Int. J. Intell. Syst.*, vol. 33, no. 2, pp. 396–416, 2018.

[4] K. Ashton, "That 'internet of things' thing," *RFID J.*, vol. 22, no. 7, pp. 97–114, 2009.

[5] F. Bader and S. Jagtap, "Internet of Things Linked Wearable Devices for Managing Food Safety in the Healthcare Sector," 2019.

[6] P. Aswale, A. Shukla, P. Bharati, S. Bharambe, and S. Palve, "An Overview of Internet of Things: Architecture, Protocols and Challenges," in *Information and Communication Technology for Intelligent Systems*, Springer, 2019, pp. 299–308.

[7] F. Cicirelli, A. Guerrieri, C. Mastroianni, G. Spezzano, and A. Vinci, *The Internet of Things for Smart Urban Ecosystems*. Springer, 2019.

[8] M. B. SR, "Automatic smart parking system using Internet of Things (IOT)," *Int. J. Sci. Res. Publ.*, p. 628, 2015.

[9] S. S. Sohail, J. Siddiqui, and R. Ali, "Classifications of Recommender Systems: A review.," *J. Eng. Sci. Technol. Rev.*, vol. 10, no. 4, 2017.

[10] C. Chang, S. N. Srirama, and R. Buyya, "Internet of Things (IoT) and New Computing Paradigms," *Fog Edge Comput. Princ. Paradig.*, pp. 1–23, 2019.

[11] S. S. Sohail, J. Siddiqui, and R. Ali, "Ordered ranked weighted aggregation based book recommendation technique: A link mining approach," in *2014 14th International Conference on Hybrid Intelligent Systems*, 2014, pp. 309–314.
[12] J. Granjal, E. Monteiro, and J. S. Silva, "Security for the internet of things: a survey of existing protocols and open research issues," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 3, pp. 1294–1312, 2015.
[13] W. Ejaz and A. Anpalagan, "Internet of Things for Smart Cities: Overview and Key Challenges," in *Internet of Things for Smart Cities*, Springer, 2019, pp. 1–15.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Data Science: Introduction, Concepts, Challenges & Future in News Classification

Feroz Ahmed
Jamia Hamdard
Hamdard Nagar, New Delhi, India
feroz.78600@gmail.com

Shabina Ghafir
Jamia Hamdard
Hamdard Nagar, New Delhi, India

Abstract - *Data Science has become a buzz in this era of Artificial Intelligence. The applicability of Data Science can be seen in marketing where it has been used in targeted advertising, in banking where it has been used in fraud & risk detection, in both speech and image recognition, in healthcare, in social media to analyze the sentiment of users and even in the field of document classification. This paper is discussing & exploring the role of data science in news documents classification. The motivation for this work arises from the need of Automatic Document Classification which is necessary when the task involves business specific contexts which cannot be fulfilled via querying on any search engine. Since a large number of websites are available over the internet nowadays therefore users generally search information from different websites via search engines now. But search engines require appropriate keywords from users in order to give relevant information from the web and sometimes users have obtained irrelevant results if he or she is not sound enough to provide keywords correctly. Thus, we need a proper document classification for the material of our wish so that the one which is required can be obtained easily instead of wasting time in searching. To understand this we have discussed ADC in news domain and mentioned a brief detail of major techniques related to it. Further, the paper also shed light on the model for classification using Stochastic Gradient Descent.*

Keywords *– Data Science, Automatic Document classification, Stochastic Gradient Descent.*

I. INTRODUCTION

The age of internet has never reached its zenith like today. Now, the accessibility of internet to a vast majority of people across the globe has given birth to platforms like e-governance, Internet of Things (IoT), Cloud Computing and others. But all this has only become possible because of the easy availability of internet and because of that, now almost every other individual is connected to it either by android smart phones, or by laptops, or by tablets. According to Mr. Ralph Jacobson (IBM Worldwide Retail Industry Analytics Portfolio Marketing Manager) the data generated in the last two years holds for about 90% of the total data [1]. The data generated via online searches, social media, e-commerce websites, online transactions, messages sent across the net, billions of photos & videos uploaded and downloaded over the net has led to an explosion of data. And with this enormous data (also known as Big Data) available with us, we are now in a situation where we need only queries or questions to ask from the system and the system will answer or reveal us interesting results and patterns and also a prediction will be made by the system based on the results and patterns through a predictive analysis. In simple words, this is what

*Corresponding Author

data science is. Thus in a single line definition it can also be defined as a systematic study of data that utilizes scientific methods, statistical skills and algorithms to extract knowledge from insights of data (both structured & unstructured) [2].

Automatic document classification [3] [4] is the task of assigning predefined categories to text documents so that they can be easily retrieved while searching. As the size of online information increasing exponentially, this task is of great significance. Even searching information via search engines will not yield totally satisfactory result unless users do not enter the exact keywords in search engines to get the relevant information. Also, simply searching will be a waste of time and in this age of digital world we cannot afford manual classification of documents over the internet. Thus, in place of manual classification of documents, machine learning algorithms can be trained to classify documents based on human-labeled training documents. Thus, here we are exploring data science in news domain to classify the news articles.

II. BACKGROUND

The various concepts of Data Science in news domain have been extensively explored in the past. Taeho Jo [5] proposed a modified version of K Nearest Neighbor (KNN) for text classification. He defined the similarity measure which was based on the semantic relations among words to consider both attribute and attribute values between representations of texts. Using the similarity measure he modified the traditional KNN. Shuo Xu, Yan Li and Zheng Wang [6] proposed a Bayesian version Naïve Bayes (NB) classifier for text classification. They applied that Bayesian version on 20 newsgroup dataset with appropriate Dirichlet hyper-parameters. Abdelaali Hassaine, Souad Mecheter and Ali Jaoua [7] represented a corpus of documents by a binary relation linking each document to the word it contains. They made use of the Hyper Rectangular Algorithm to extract the list of the most representative words in a hierarchical way from the relation being made. They then fed the extracted keywords into the random forest classifier in order to foretell the category of each document. Ari Arulia Hakim, Alva Erwin, et al. [8] performed the classification of news articles in Bahasa Indonesia using Term Frequency Inverse Document Frequency (TF – IDF) algorithm. Even, the concept of deep learning has been applied in news classification. Shuang Qiu, Mingyang Jiang, et al. [9] proposed a new kind of Stacked Denoising Auto Encoder (SDAE) algorithm which they termed as LMSDAE algorithm. They applied this newly version for classification of Chinese news articles and found it better than other three algorithms – SDAE, Sparse Denoising Auto Encoder (SPDAE) and Deep Belief Nets (DBN) respectively because of the reduced training times and increased convergence rate. Similarly, Chenbin Li, Guohua Zhan, et al. purposed an improved Bi-LSTM-CNN [10] in news text classification. Also, the classification of Chinese text has also been discussed in [11] using semantic kernel in SVM. Ankit Dilip Patel and Yogesh Kumar Sharma [12] proposed Web Page Classification (WPC) on newsfeeds to recognize and allocate them into categories of news like sports, business, world, health in order to enhance user's accessibility towards relevant news by pass over suitable category as per user's choice. This has been done by choosing hybrid technique of URL analysis and content context analysis. Fasihul Kabir, Sabbir Siddique, et al. [13] performed Bangla text document categorization using SGD classifier. They performed their experiment on BDNews24 documents.

Thus, after exploring the above work, we observed that majority of the work has been done in mainly Naïve Bayes, SVM, Neural Networks, k-NN, Decision trees and Regression-based classifiers which has also been mentioned in [4] and [14] besides few exceptions. Thus, in the next section we discussed these notable techniques.

III. TECHNIQUES

A. K- Nearest Neighbor

K-Nearest Neighbor (K-NN) is the simplest algorithm in machine learning which is widely used in ADC. The basic idea of it is to predict the label of a data point by looking at the 'k' closest labeled data point and getting them to vote on what label the unlabeled point should have. The closeness of unlabeled data point is defined in terms of distance metric like Euclidean distance. Some examples related to it are mentioned in [15, 16, 17]. The preprocessing of the dataset is being done before rather than applying the classifier on the original dataset directly, in order to improve the performance of the classifier. Some examples of such techniques are discussed in [18, 19, 20]. Since feature extraction and representation of document is an important part of ADC as majority of the terms in large collection of texts might not be part of interested topic. Hence, in [17] a lot of ways has been described to understand the coalition among words and categories.

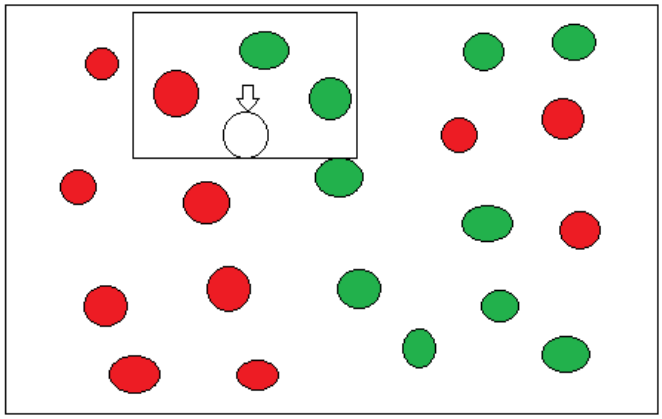


Fig. 2 K-NN in two dimensions

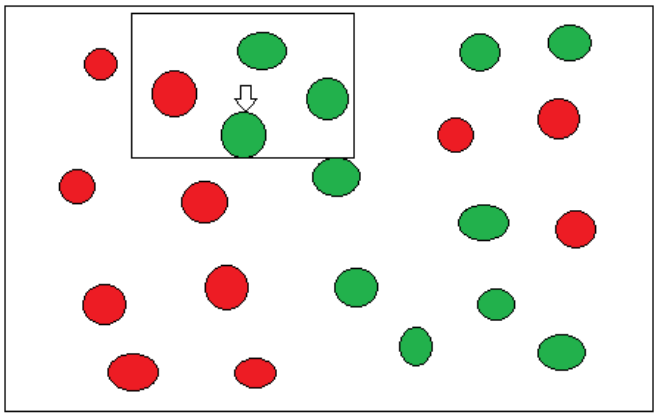


Fig. 3 unlabeled data point will be green when k=3.

There are some limitations of traditional k-NN which are as follows:

i. Problem of calculation because all training samples used in classification.
ii. The presentation is solely dependent on the training set.
iii. Samples having same weight.

B. Naïve Bayes

It is one of the most exploited algorithms in the field of data science for text classification [21]. It is based on the Bayes theorem. According to it a class's feature is free

from any other's feature present. And even if they depend on each other, their contribution to probability will remain independent due to which they are being called as 'Naive' [22].

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{1}$$

$P(c|x)$ is the posterior probability where 'c' is target class and 'x' is an attribute on which prediction is based. $P(c)$ is the prior probability of class. $P(x|c)$ denotes the likelihood. $P(x)$ is the prior probability of the predictor. The multinomial naïve bayes is another enhanced version of naïve bayes algorithm [23]. The multinomial NB classifier is appropriate for classification with discrete features. It normally requires integer feature counts however fractional count like TF-IDF also work.

C. Support Vector Machines (SVM)

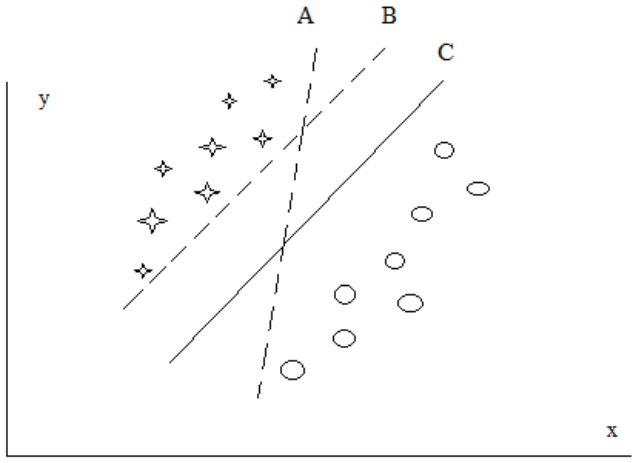


Figure 4

The main principle in SVM is to determine the best hyperplane that separate different classes perfectly. For example in the figure 3, we can see that the line 'C' is acting as the best hyperplane that segregates the two different classes rather than 'A' and 'B'. Hence, the line 'C' is maintaining the maximum margin of segregation. Here the hyperplane is a straight line because it is a case of two-dimension. However, theoretically there can be n number of hyperplanes for the separation of training data. Because of sporadic high dimensional nature of text data, normally SVM classification is used as shown in [24]. Linear SVM is used commonly because it is simple and ease to interpret. In [24, 25] the first set of SVM classifiers which were transformed for text field were suggested. A detailed work related to SVM algorithm has been discussed in [26]. It gives a reason to choose SVM as an option to work well in a diverse variety. This has also been shown in [27] where it has been applied in e-mail for classification of e-mails into two categories i.e. either spam or not spam. The findings reveal a clear result giving SVM an upper hand than the rest of the methods. Because of its flexibility it can be combined with interactive user-feedback methods [28]. The SVM has also been efficiently used where unlabeled data is in large quantity and labeled data is in small quantity [29].

D. Decision trees

As the name denotes, it is a tree like structure where internal nodes of the tree represents a test on features, every branch of the tree depict an outcome of the test and the class label is held by terminal node. The methods of decision trees involves the recreation of manual classification of training documents in a tree form where nodes represent queries and leaf nodes represent the corresponding class for the document. After the creation of tree, the categorization of a new document can be done by placing it in the root node. Some of the advantages of decision trees include generation of rules which are easy to understand, require less computation, able to deal with both categorical and continuous variables. On the other hand, some of its disadvantages are: less appropriate in case of predicting the value of a continuous attribute, chances of errors in classification when there are small number of training instances with relatively many classes and last that they are computationally costly to train.

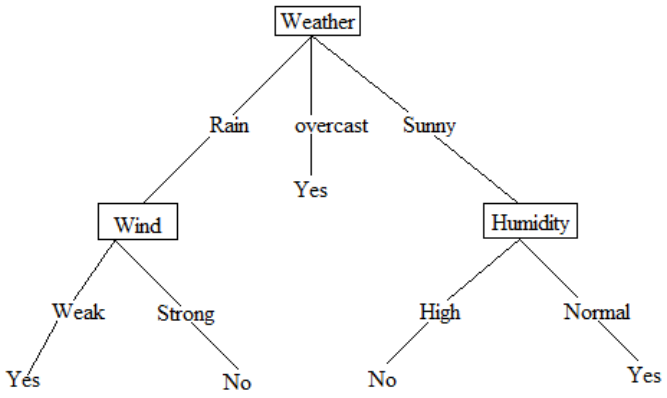


Fig. 5 Decision Tree for playing cricket

E. Regression

Regression is a type of predictive modeling which is used in learning the relationship between attributes of real world i.e. it finds out the relationship between a target variable and a predictor variable. Here, we try to maintain a curve or a line which passes through all data points in such a way that the difference between the distances of data points from the curve or line remains minimum. As opposed to binary attribute the regression methods are designed for attributes of real world. But, this does not become an obstacle in classification, since the binary value of a class may be considered as an elementary case of real value; methods of regression technique for example logistic regression are able to design discrete response variables. The learning of high dimensionality of data by logistic regression which is like text of natural language poses a statistical and computational challenge.

IV. CONCLUSIONS & FUTURE SCOPE

After having discussed the work in section 2 and section 3 above, we can conclude that Data Science has been widely applied and used in the field of news classification. While discussing the techniques in section 3 we came across the strengths & weakness of some techniques. However, all techniques are generally suited for the classification of text data although each having its pros and cons related to it. We noticed that the Stochastic Gradient Descent has used less in the context of large scale learning although it has been in the community of learning techniques for a long time which is also mentioned in [30]. Thus, we decide to take this work further with an aim to explore and build a better model for news classification.

REFERENCES

[1] R. Jacobson, "2.5 quintillion bytes of data created every day. How does CPG Retail manage it?," *Industry Insights*, April 24, 2013. [Online]. Available: https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/ [Accessed: Nov. 8, 2018].

[2] V. Dhar, "Data Science and Prediction," *Communications of the ACM*, vol. 56, no. 12, pp. 64-73, Dec. 2013.

[3] "What is automatic document classification?," *Expert System,* Mar. 17, 2016. [Online]. Available: https://www.expertsystem.com/what-is-automatic-document-classification/. [Accessed: Jan. 1, 2019].

[4] C.C. Aggarwal and C. Zhai, "A Survey of Text Classification Algorithms," in *Mining Text Data,* C. C. Aggarwal and C. Zhai, Eds. Boston, MA: Springer, 2012, pp. 163-222. [Online] Available: SpringerLink

[5] T. Jo, "Classifying News Articles Using Feature Similarity K Nearest Neighbor," in *Proceedings of the Sixth International Conference on Green and Human Information Technology : Proc. of the 6th Int. Conf. on Green and Human Information Technology, ICGHIT 2018, Chiang Mai, Thailand, Jan 31-Feb 2, 2018,* S.O. Hwang, S.Y. Tan, F. Bien, Eds. Singapore: Springer, 2018. pp. 73-8.

[6] S. Xu, Y. Li and Z. Wang, "Bayesian Multinomial Naïve Bayes Classifier to Text Classification," in *Advanced Multimedia and Ubiquitous Engineering: Proc. of the 11th Int. Conf. on Multimedia and Ubiquitous Engineering, MUE 2017, Proc. of the 12th Int. Conf. on Future Information Technology, FutureTech 2017, Seoul, South Korea, May 22-24, 2017,* J.J. Park, S.C. Chen and K.K.R. Choo, Eds. Singapore: Springer, 2017. pp. 347-52.

[7] A. Hassaine, S. Mecheter and A. Jaoua, " Text Categorization Using Hyper Rectangular Keyword Extraction : Application to News Articles Classification," in *Relational and Algebraic Methods in Computer Science : Proc. of the 15th Int. Conf. on Relational and Algebraic methods in Computer Science, RAMiCS 2015, Braga, Portugal, Sep 28 – Oct 1, 2015,* W. Kahl, M. Winter, J. Oliveira, Eds. Cham: Springer, 2015. pp. 312-25.

[8] A. A. Hakim, A. Erwin, K. I. Eng, M. Galinium and W. Muliady, " Automated document classification for news article in Bahasa Indonesia based on term frequency inverse document frequency (TF – IDF) approach," in *Proc. of the 6th Int. Conf. on Information Technology and Electrical Engineering, ICITEE 2014, Yogyakarta, Indonesia, Oct 7 – 8, 2014,* pp. 1-4.

[9] S. Qiu, M. Jiang, Z. Zhang, Y. Lu and Z. Pei, "Chinese News Text Classification of the Stacked Denoising Auto Encoder Based on Adaptive Learning Rate and Additional Momentum Item," in *Advances in Neural Networks – ISNN 2018 : Proc. of the 15th Int. Symposium on Neural Networks, ISNN 2018, Minsk, Belarus, Jun 25 – 28, 2018,* T. Huang, J. Lv, C. Sun, A. V. Tuzikov, Eds. Cham : Springer, 2018. pp. 578-84.

[10] C. Li, G. Zhan and Z. Li, " News Text Classification Based on Improved Bi-LSTM-CNN," in *Proc. of the 9th Int. Conf. on Information Technology in Medicine and Education, ITME 2018, Hangzhou, China, Oct 19 – 21, 2018,* pp. 890-93.

[11] M. Fanjin, H. Ling and T.J.W. Xinzheng, " The Research of Semantic Kernel in SVM for Chinese Text Classifiction" in *Proc. of the 2nd Int. Conf. on Intelligent Information Processing, IIP'17 2017, Bangkok, Thailand, July 17 – 18, 2017*.

[12] A.D. Patel and Y.K. Sharma, "Web Page Classification on News Feeds Using Hybrid Technique for Extraction," in *Proc. of the 3rd Int. Conf. on Information and Communication Technology for Intelligent Systems, ICTIS 2018, Ahmedabad, India, April 6 – 7, 2018,* S.C. Satapathy, A. Joshi, Eds. Singapore: Springer, 2018. pp. 399-05.

[13] F. Kabir, S. Siddique, M. R. A. Kotwal and M. N. Huda, "Bangla text document categorization using Stochastic Gradient Descent (SGD) classifier," in *the Int. Conf. on*

Cognitive Computing and Information Processing, CCIP 2015, Noida, India, March 3 – 4, 2015, pp. 1-4.

[14] S. Brindha, K. Prabha and S. Sukumaran, "A survey on classification techniques for text mining," in Proc. of the 3rd Int. Conf. on Advanced Computing and Communication Systems, ICACCS 2016, Coimbatore, India, Jan 22-23, 2016, pp. 1-5.

[15] W. Cohen, H. Hirsh. Joins that generalize: text classification using Whirl. *ACM KDD Conference*, 1998.

[16] E.-H. Han, G. Karypis, V. Kumar. Text Categorization using Weighted-Adjusted *k*-nearest neighbor classification, *PAKDD Conference*,2001.

[17] Y. Yang, C.G. Chute. An example-based mapping method for text categorization and retrieval. *ACM Transactions on Information Systems*, 12(3), 1994.

[18] E.-H. Han, G. Karypis. Centroid-based Document Classification: Analysis and Experimental Results, *PKDD Conference*, 2000.

[19] W. Lam, C. Y. Ho. Using a generalized instance set for automatic text categorization. *ACM SIGIR Conference*, 1998.

[20] J. Rocchio. Relevance feedback information retrieval. *The Smart Retrieval System-Experiments in Automatic Document Processing*, G. Salton, Ed. Prentice Hall, Englewood Cliffs, NJ, pp 313–323,1971.

[21] D. Chickering, D. Heckerman, C. Meek. A Bayesian approach for learning Bayesian networks with local structure. *Thirteenth Conference on Uncertainty in Artificial Intelligence*, 1997.

[22] S. Ray, "6 Easy Steps to Learn Naïve Bayes Algorithm (with codes in Python and R),"*analyticsvidhya.com*, Sep. 11, 2017. [Online]. Available: https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/ [Accessed: Jan 31, 2019].

[23] Mansi Goyal, Ankita Sharma "An Efficient Malicious Email Detection Using Multi Naive Bayes Classifier", Vol.5, Issue5, Pp39-58,May 2015.

[24] T. Joachims. A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization. *ICML Conference*, 1997.

[25] T. Joachims. Transductive inference for text classification using support vector machines. *ICML Conference*, 1999.

[26] T. Joachims. A Statistical Learning Model of Text Classification for Support Vector Machines. *ACM SIGIR Conference*, 2001.

[27] H. Drucker, D. Wu, V. Vapnik. Support Vector Machines for Spam Categorization. *IEEE Transactions on Neural Networks*, 10(5), pp.1048–1054, 1999.

[28] H. Raghavan, J. Allan. An interactive algorithm for asking and incorporating feature feedback into support vector machines. *ACM SIGIR Conference*, 2007.

[29] V. Sindhwani, S. S. Keerthi. Large scale semi-supervised linear SVMs. *ACM SIGIR Conference*, 2006

[30] "Stochastic Gradient Descent". [Online]. Available: https://scikit-learn.org/stable/modules/sgd.html [Accessed : Jan 27, 2019]

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

Developing CNN Model for Image Classification

Asim Jamal
Jamia Hamdard
New Delhi
jamalasim081@gmail.com

Nazia Siddiqui
Jamia Hamdard
New Delhi
nazias0913@gmail.com

Farheen Siddiqui
Jamia Hamdard
New Delhi

Richa Gupta *
Jamia Hamdard
New Delhi

Abstract – *The idea behind computer vision techniques is to make day to day life easier and more convenient, keeping this in mind, we develop an approach using CNN to classify Images to respective labels. In The paper we discuss the process of detection, training and recognition of hand-gestures. We develop a 3 layer fully-connected neural network to and use back-propagation for each epoch.*

There is no special hardware used anywhere in the process. We have compared each label with the prediction probability to test our model and we observe how much iterations of training is required for the model to reach desired accuracy.

Keywords *– Convolution neural network, Label, Epoch, Layers*

I. INTRODUCTION

In this paper, we study machine learning algorithms and use convolutional neural network layers to train the dataset to and be able to classify images. We detect different gestures; we classify each gesture with a particular label to and try recognizing it. A current approach on Computer vision techniques has a vital role of machine language algorithm especially when it comes to detection of any object. We develop our approach by convolutional neural network and work on avoiding over fitting to ensure accuracy and precision during activation of our neural nets.

Weights and biases are randomly dispensed to each cell initially and analysed further for a output value after each epoch the process is repeated until we get the output same as the desired output.

II. RELATED WORK

Machine learning enables our system to make decisions by operating in a self-learning manner. We feed the data and analyse the trends, so when presented with new data, it provides us with explanations according to experience i.e. training of dataset. In this paper we talk about classifying hand gestures which are captured as HSV images to their individual interpretation through labels for each gesture [1].

*Corresponding Author

In "Research Paper on Basic of Artificial Neural Network" we study that the neurons must be arranged on to different layers, and then the connections between them are made to successfully make a network like structure. The input layer is fed with data; the output is what gives us a value which is interpreted between them can be any number of hidden layers containing neurons in conjugated layout.

The data comprises of noise and errors too, when all the values from the neural net are appended for evaluation of output value the model complies with errors. To avoid this a method called Dropout discussed in [2],which is introduced to reduce over fitting by randomly removing or not considering values of some neurons from the network and the rest are trained by back-propagation.

For the dataset we capture images in HSV format is used for a clearer segregation of colours and background subtraction. [3] Puts forward a procedure to analyse and recognize hand gestures. The methodology enables us to detect hand gestures without any extra equipment.

III. METHODOLOGY

The system consists of 3 parts that are making of dataset, training of dataset and recognition. All the modules are implemented and executed using python.

For training of dataset, we use python libraries like Keras, TensorFlow, open CV for training and hand detection [4].The captured data is fed to the neural network and trained.

Our label or class set comprises of 12 emoticons and will be referred as class or label further in this paper. We capture different gesture according to each label or class. This image dataset is used as input to our neural network model.

A. Capturing Dataset

To create dataset we capture hand gestures using primary camera. We define coordinates for contour frame on the input screen while capturing so it ignores the other components in the camera frame. We capture 1200 images for each label then flip the image about y-axis and convert it from RGB to HSV simultaneously. we use the Gaussian distribution algorithm to eliminate the background and shadow by only recognizing only the colour light brown .We then perform basic open cv Image analysis operation like masking, bitwise and Gaussian blur to remove noise and smoothen the resultant captured image. [5].

We capture 1200 images of dataset for each class or label and store in separate folders which makes pointing to data easier

Our computer analyses images as a dataset of numbers where each pixel represents a value. The values are different for different formats. As our dataset is captured in HSV format each cell value represents pixel intensity ranging from 0-255.0 being absolute black, 255 being white and grey values range in-between.

These values for whole dataset are stored in a file of CSV format. For our computer to analyse a gesture or an object is analyses its curves and edges, which for a particular item will be consistent in every image. In this case for hand gesture, area surrounding the hand shape will be analysed by the extreme difference in pixel intensity values.

We achieve this through ImageIO library of Python and convert these HSV format images in their pixel intensity representation and store in CSV file. Each and every pixel is successfully extracted from the images before converting it to CSV format. This is the final Dataset which will be fed to our neural network.

B. Model Description & Training

We start with a convolution layer followed by non-linear layer and pooling layers. Data is passed on through the layers for output evaluation.

The convolution layer works on a small matrix of dataset from whole picture and the values are evaluated and brought down to single value which is fed to a neuron, this is done through filter. The filter goes on evaluating the next set of matrix and feeding it to the network along the whole dataset. The convolution scans the whole dataset this way starting

from the upper left corner. As the computer analyses images in terms of edges and peripheries it is important to scan through whole dataset. The values are passed through all the layers. After evaluation in first layer the output values from previous layer become input to the next layer.

A non-linear function to each convolutional layer is added for non-linear properties without which the model won't be able to classify and provide output efficiently. we then add pooling layer which works on width and height of the image and apply down sampling operation on them to reduce the overall size of the data or images in this case. Already evaluated features need not be evaluated again.

After layers we add fully connected layers, this evaluates the whole convolutional network. The whole network is attached to results as a dimensional vector. This points the resulting classes from which the model predicts the appropriate class.

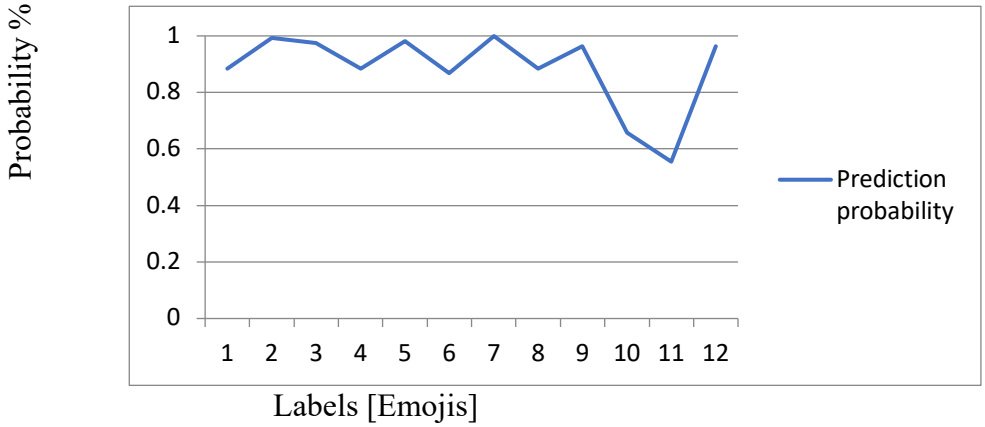
In this approach we use 2 convolution and 2 pool layers, flatten them to make a network of 3 fully connected layers. We use the activation function "relu" and "sigmoid for non-linearity. Flatten performs the input role and next we add dense layer to connect layer with activation function "relu". Dropout is used to avoid Over fitting of model which may make prediction erroneous. Finally, we compile our model.[2]

For training we divide the data into testing and training data. For our dataset we use 90% data for testing and 10% for testing. We train the data with "fit" function.

For each epoch we analyse the accuracy and error %.

C. Recognition

We now need to test our model. We take input through similar manner we capture images for dataset. This data is imposed on to the model and predict the class or label accordingly.



In the above graph, We observed the accuracy for each label.

We have compared all the labels according to their prediction probability. As we can observe for most of the labels the prediction probability is more than 80%.external factors such as proper lightening proved the result to be better.

The application is independent of any factors irrespective of size, colour of the hand while we take gesture as input. The accuracy rate after training is 1.

IV. CONCLUSION

With the result observed we can conclude that our model when trained for the very first time was effective with a accuracy rate of 0.99833 and in identifying the correct label but after 2 or 3 iteration in training the model proved to be much precise and accurate as it became 1 eventually, thus we can say that out convolutional neural network is capable of achieving better result with challenging dataset.

The dataset is separated into training and testing data, a particular percentage is used for both the process in our application 90% and 10% respectively. The accuracy and probability of prediction of each class is average but are not similar the accuracy level depends on how much we have trained the model, with each epoch the accuracy rate is observed to increase.

V. Future Scope

Increase in the computational power of the system it is predicted that the overall training time will decrease and there can be an upgraded system where the number of labels can increase gradually. Further with advancement in technology we aim to implement such systems in day to day life as well can these system can be used for Defence development as well. Also if we can further reduce the number of modules required for the implementation of the system it would be much easier. Other improvements can be the versatility of the gesture; the system can be more flexible in recognizing the gesture by creating bigger dataset which holds more information about gesture for particular labels.

References

[1] Mufarroha, Fifin A., Fitri Utaminingrum, and Wayan F. Mahmudy. "Segmentation Algorithm to Determine Group for Hand Gesture Recognition." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 9, no. 2-8,pp.13-17, 2017.

[2] Baldi, P. and Sadowski, P., 2014. The dropout learning algorithm. *Artificial intelligence*, *210*, pp.78-122.

[3] Suganya, P.,Sathya, R.and Vijayalakshmi,K."Detection and Recognition of Hand Gestures to control the system application by Neural Networks"International Journal of Pure and Applied Mathematics 118 No. 10, pp.399-405, 2018.

[4] Meduim Homepage,Image Classification with Convolutional Neural Networks,Ksenia Sorokina,last accesed:- 04/03/2019

[5] Git-Hub Homepage, Akshay Bahadur profile,last accesed:- 03/03/2019

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

RADAR IMAGING SYSTEM USING WIRELESS VISION (Wi-Vi) TECHNOLOGY

Harleen Kaur
Department of Computer Science and Engineering
School of Engineering Sciences and Technology
Jamia Hamdard
New Delhi, India

Ritu Chauhan
Centre for Computational Biology and Bioinformatics
Amity University
Noida,U.P., India

Kisa Zehra
Department of Computer Science and Engineering
School of Engineering Sciences and Technology
Jamia Hamdard
New Delhi, India

Abstract- *Wireless Fidelity(Wi-Fi) is a wireless mechanization that broadcast the data through air using radio frequency. Wi-Fi signals are information carrier from transmitter to receiver. Wireless Vision(Wi-Vi) is machinery that is invented by scientists that uses Wi-Fi signals which helps a user to observe the traversing objects through the walls and trailing closed doors. Wi-Vi technology is based upon the concept of radar imaging, but instead of using the high power it uses Wi-Fi signals that helps in tracking human movement behind the wall. This paper has introduced the use of Multiple Input, Multiple output (MIMO) interface to focus the receiver on moving object and nullify the effect of the static object.*

Keywords- *Wireless technology, Multiple Input, Multiple output (MIMO), Seeing Through Walls, Wireless Vision (Wi-Vi), Transmitter*

I. INTRODUCTION

This paper is exploring the concepts of Wireless Fidelity (Wi-Fi) signals and Multiple Input, Multiple output (MIMO) communication that helps us to construct a machine to capture the gestures of humans behind the walls. A low-cost, X-ray vision. Wireless Vision (Wi-Vi) technology is based upon the concept of radar imaging but instead of using high power it uses Wi-Fi signals that helps in tracking human movement behind the wall and closed doors. The paper uses the recent advances in Multiple Input, Multiple output (MIMO) communications to through-wall imaging. In Multiple Input, Multiple output (MIMO), the transmission is encoded by multiple antenna systems so that signal at particular antenna is null using this capability the Multiple Input, Multiple output (MIMO) systems eliminate interface to unwanted receivers. This concept is used to minimize the causalities in deadlock and surety conditions, emergency responder can use this device to see through the wall, rubber, collapsed structure or any opaque object. To track the moving object behind the wall we transmit Wi-Fi signal in the direction of wall, resulting in two main problems (I) Flash effect, (II)reduction in the magnitude of signal by three to five times caused after traversing the wall [1]. To resolve these problems experiments are done on those system for getting the desired result. A system named Ultra-wideband was developed by radar community which detects the human behind wall and it also does show the moving blobs in the output screen. Since this system have some drawbacks two new

*Corresponding Author

systems were developed which were Wi-Vi and WISEE which work on Wi-Fi signals to detect the human movement behind the wall. WISEE works on the concept of Doppler Shift Principle. Whereas, Wi-Vi uses Wi-Fi signals to detect the human movement behind the wall, this technology uses nulling technique for removing the flash effect. And it does not require large array of antenna it only requires smaller three- antenna MIMO radio [2]. The term Flash effect actually means the entire reflection of an object behind the wall not just the wall. The Wi-Vi device sends Wi-Fi signals to any opaque barrier which is to be examined and calculate the time of the returned signal. The researchers need to discover out a way by which unwanted interferences is cancelled, as only fraction of signal penetrates the wall [3]. The solution that came up was that transmitting two signals, in which one was inverse of another. Identical reflections are created by all static object, so when a stationary object is hit by one signal, the other cancels out [4]. Nevertheless, between 2 signals the moving object change and so the receiver picks it up.

II. RELATED WORKS

The three major areas in which Wi-Vi is related are:

(i) Through Wall Radar

For many decades human has fantasized about seeing through the wall. From many years modeling and simulation was the only focus for the inventors. Lately, some implementations were tested on human movement. Formerly the flash effect of the system was removed by removing reflected signal of the targeted wall from reflected signal of the object which is present behind wall or closed door. These isolations are made by time domain with very short-term pulses (lesser than 1ns)[5]. The time taken to arrive by a pulse reflected off of a wall is lesser than that of a pulse reflected from a moving body/object that was behind the wall or closed door. Another method of achieving these isolations is by frequency domains that use a liner frequency chip in which the tone is analog filtered which corresponds to the wall and helps in removing the flash effect[6].

(ii) Gesture Based Interfaces

Due to the advanced modern day technologies, gesture identification/recognition systems (like the ones used in gaming consoles like Microsoft Xbox and Nintendo Wii) can easily identify various gestures. With the application of cameras or sensors on human body, the academic community was able to develop various systems/software that are capable of recognizing human gestures[7].

Leveraging narrowband signals in ranges of 2.4 GHz for identifying human activities/gestures in line of sight using micro Doppler Signatures has been made possible by the recent technological advancements[8]. However, Wi-Vi is the first gesture based interface that does not require line of sight sensors. It can detect gestures through a wall and humans are not required to carry wireless devices or wear any gear[9].

(iii) Infrared and Thermal Imaging

Infrared and Thermal imaging technologies, similar to the Wi-Vi, help extend the human vision beyond the visible range of the electromagnetic spectrum, allowing us to detect various object even in dark conditions or smoke [10]. These devices capture thermal/infrared energy that is reflected off the objects kept in line of sight of the sensors. The cameras that use similar technologies have a short wavelength (micrometer to sub millimeter) and hence cannot see through walls, unlike Wi-Vi that employs signals having wavelengths having wavelengths approximately equal to 12.5 cm^3 .

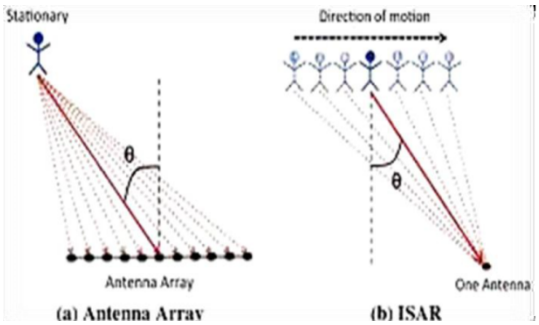


Fig.1 (a) Antenna Array (b) Inverse Synthetic Aperture Radar (ISAR)

III. IDENTIFYING AND TRACKING HUMANS

Tracking a single human

Nowadays, wall systems use antenna array to track human motion. They do so by moving the array beams to the specified direction where the maximum energy exists[11]. The approach direction will be the signal's spatial arrival angle (fig.1). They speculate the object's motion in space by tracking the particular apex in time.

Wi-Vi avoids antenna array mainly for two reasons: -

(i) To achieve a good resolution with a narrow beam, large sized antenna is required with many antenna elements, hence the cost of the device increases.

(ii) The flash effect is taken care of with the utilization of MIMO nulling by the Wi-Vi. MIMO nulling adds multiple receiving antennas which require signal nulling at every one of them. Hence, more transmit antennas are added to the machine making it larger[12].

Multiple Human Tracking

Let's perform an activity to learn how Wi-Vi tracks multiple humans. We assume that each human emulates a different antenna ray. The Wi-Vi comprises of one single antenna only and hence the signal that is received will be a superposition of antenna rays of the moving individuals [13]. Hence, we will obtain a number of curved lines (as moving humans) as opposed to one curved line at that point in time (Fig.2). As a result of multiple humans, the noise will increase automatically. The different human body parts can be considered as different objects. The reflected signal from all humans is correlated in time as the transmitted signal is reflected from them[14]. The presence of multiple humans causes problems as the reflections combine and can diminish each other over sometime[15].

IV. THROUGH WALL GESTURE-BASED COMMUNICATION

V.

To send message to a computer wirelessly, a human should carry a wireless device. Wi-Vi helps human by representing "1" but and "0" but gestures who won't carry any wireless device for communication[16]. Then, the gesture for creating the messages which have different interpretations can be composed. Additionally, Wi-Vi evolved itself by borrowing other known principles namely to add easy code which is reliable or to reserve '0's and '1's for packet preambles[17]

Gesture Encoding

0 and 1 bits are encoded with the help of a modulation scheme at the transmitter. Then, Wi-Vi can enable this type of encoding by using gesture, variety of gestures can be used to represent the bits[18].

Conditions for Gesture Encoding

i) At the start of the gesture a human should be in the initial state itself.
ii) Gesture should be simple in nature.
iii) Gesture should be easy in order to detect and decode it.

0 bit indicate- a step forward by a step backward.

1 bit indicate- a step backward by step forward.

The angle which connects a human to the Wi-Vi will be positive when the human goes towards the Wi-Vi and when the human goes away from the Wi-Vi the angle will become negative [19].

Gesture Decoding

Gesture decoding uses standard communication techniques and is simple to use. The Wi-Vi is used to input the data. Wi-Fi is used to apply two filters which are matched, one is used for step backward and the other one is used for step forward. The Wi-Vi then uses the matched filter to apply it to the received signal and then to get the output [20].

CONCLUSION

Wi-Vi is a device that uses Wi-Fi signals to observe the moving human and objects behind the wall and also in closed doors. Comparing Wi-Vi with previous systems, this technology (operated in ISM band) allows cheaper devices to see through walls. The device can build the wall or door without any transmitting device. Wi-Fi can be used in many other services like indoor sensing, localization and control. Wi-Vi device only uses one receiver, which calculates the time taken for the signal to be reflected that helps in accurately finding the location of the moving object or human. Since Wi-Fi is a low-cost technology the Wi-Vi devices can also be used in disaster recovery and gaming activities. This device can also be used as gesture-based interface so it can be used in controlling appliances or lighting since it has the capability to track all the actions that are made behind any opaque object and also it does not demand any line of sight among the user and the Wi-Vi machinery.

REFERENCES

[1] M.Murugan, Mr.G.Sathish, "A Study on Future Scope of Wi-Vi Technology", International Journal of Scientific & Engineering Research Vol. 9, No. 4, April 2018.

[2] Rohan H. Patel, Arvind G. Thosare and Kaustubh U. Pathak, "A Review on Wi-Vi Technology", International Journal of Trend in Research and Development, Volume 4(5), Sep-Oct 2017.

[3]Prerna Garg, Shikha, "Wi-Vi TECHNOLOGY", International Journal Of Advance Research In Science And Engineering, Vol. 2, No.9, September 2013.

[4]Munipalli Uma, Maheswara rao, Bala Brahmeswara Kadaru, "Seeing Through Walls Using Wi-Vi", International Research Journal of Engineering and Technology, Vol..4, No. 6, June 2017.

[5]Sudarshan Adeppa,"Detection of Objects across the Walls with Wi-Fi Technology", International Journal on Emerging Technologies, Vol. No.6(2),2015.

[6] Mayuri Dharma Shinde, Shailesh Jadhav," Review on Wi-Vi Technology", International Journal of Advanced Research in Computer and Communication Engineering ,Vol. 5, Issue 5, May 2016

[7]Seeing through walls - MIT's Lincoln Laboratory. http://www.youtube.com/watch?v=H5xmo7iJ7KA2013

[8] Fadel Adib and Dina Katabi,"See through walls with wi-fi," Proceedings of the ACM SIGCOMM Conference, 2013.

[9] Kevin Chetty," Through-the-Wall Sensing of Personnel Using Passive Bistatic WiFi Radar at Standoff Distances", IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. 50, NO. 4, APRIL 2012

[10] H. Yacoub and T. K. Sarkar, "A homomorphic approach for through-wall sensing," IEEE Trans. Geosci. Remote Sens., vol. 47, no. 5, pp. 1318– 1327, May 2009.

[11]Y.YunqiangandA.E.Fathy,"Developmentandimplementationofarealtimesee-through-wall raar system based on FPGA, "IEEE Trans. Geosci. Remote Sens., Vol. 47, No. 5, pp. 1270–1280, May 2009.

[12] H. Estephan, M. G. Amin, and K. M. Yemelyanov, "Optimal waveform design for improved indoor target detection in sensing through-the-wall applications,"IEEETrans.Geosci.RemoteSens.,Vol.48, No.7, pp.2930– 2941, Jul. 2010.

[13] D. Poullin, "Passive detection using digital broadcasters (DAB, DVB) with COFDM modulation," Proc. Inst. Elect. Eng.—Radar, Sonar Navig., vol. 152, no. 3, pp. 143–152, Jun. 2005.

[14]M.C.Jackson,"Thegeometryofbistaticradarsystems,"Proc.Inst.Elect. Eng. F-Commun., Radar Signal Process., vol. 133, no. 7, pp. 604–612, Dec. 1986.

[15] F. Colone, D. W. O'Hagan, P. Lombardo, and C. J. Baker, "A multistage processing algorithm for disturbance removal and target detection in passive bistatic radar," IEEE Trans. Aerosp. Electron. Syst., Vol. 45, No. 2, pp. 698–722, April 2009.

[16] J. A. Gomez and G. Brooker, "Ultra-wideband radars for throughwall imaging in robotics," in Recent Advances in Sensing Technology. New York: Springer-Verlag, 2009, pp. 271–282, ed.

[17] R. J. Fontana, "Recent system applications of short-pulse ultra-wideband (UWB) technology," IEEE Trans. Microw. Theory Tech., Vol. 52, No. 9, pp. 2087–2104, Sep. 2004.

[18] P. Howland, "Editorial: Passive radar systems," IEE Proc. Radar, Sonar Navig., Vol. 152, No. 3, pp. 105–106, Jun. 2005.

[19] M. Amin and K. Sarabandi, "Special issue on remote sensing of building interior," IEEE Trans. Geosci. Remote Sens., Vol. 47, No. 5, pp. 1267– 1268, May 2009.

[20] O. Sisma, A. Gaugue, C. Liebe, and J.-M. Ogier, "UWB radar: Vision through a wall," Telecommun. Syst., Vol. 38, No. 1/2, pp. 53–59, Jun. 2008.

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

RSALocker: Design and Implementation of Ransomware based on Public Key Cryptosystem

Aqeel Khalique *
Deptt. of CSE, SEST, Jamia Hamdard
New Delhi, India
aqeelkhalique@gmail.com

Hirra Sultan
Deptt. of CSE, SEST, Jamia Hamdard
New Delhi, India

Imran Hussain
Deptt. of CSE, SEST, Jamia Hamdard
New Delhi, India

Md. Omair Ahmad
Deptt. of CSE, SEST, Jamia Hamdard
New Delhi, India

Abstract - *Ransomware is a malicious software program designed to encrypt data on a victim device. The data remains hostage unless a ransom demand has been fulfilled. The data can be recovered by paying the ransom or by cleaning the device and reinstalling the operating system and data from a backup. There is an increase in the ransomware attacks but no viable solution to decrypt the data has been designed yet. In this paper, weimplement a ransomware to encrypt data and ransom was demanded for decryption of the data. Brute force attack was used to decrypt the data but produced negative results. Decryption of the data without knowledge of key is not possible and key generation is not possible without knowledge of the prime numbers. The paper also present algorithms employed in the development of the ransomware and the difficulty to recover encrypted data without a key. All the possible reasons and factors have been analyzed to understand the threat of malware any connected device to the internet.*

Keywords - *RSALocker, Public Key Cryptography, Information Security, Malware, Ransomware.*

I. INTRODUCTION

Information and communication technology (ICT) is an inseparable part of the current world. It is embedded in so many fields that without it many things would not be the same. It is easily available and highly vulnerable technology [1].

Vulnerability in ICT can be attributed to the rapid growth of the field and the fact that all codes have some exploitable bugs. The discovery of a bug by a cyber-criminal can result in exploitation and security breach. The security challenges are of two types. Either the attacker directly attacks device making it an active attack or inactively steals data or monitors traffic generation and usage patterns of a user,thus executing a passive attack. Such attacks are executed to steal data, identity or resource by using various methods and mechanisms. Examples are snooping, masquerading, replaying, DOS etc.[2].

The various security challenges can be overcome by the use of certain security protocols. All the security protocols are at some point governed by cryptography. Cryptography helps

*Corresponding Author

in making sure the data remains safe, it is accessed by the right person and is always available to the authenticated person. Cryptography is the current field which faces all security challenges [3].

Apart from the vulnerability of acquiring data and hacking into devices, there are many software programs that can harm a device. Such software programs are called "Malware". A malware is specifically designed with the intent of harming the victim device. It can be used in many malicious ways. A lot of malware is created today for profit through forced advertising (adware), stealing information (spyware), spreading spam emails or child pornography (zombie computer) or to extort money (ransomware)[4].
Ransomware is a type of malicious software that hinders working of a device or user access to data and certain programs till a demand is fulfilled. The demand is a ransom amount to be paid for data decryption. Some malware programs may merely scare the user into paying the attackers by creating a pop-up on screen, while others may go to the extent of making changes to the boot sector of the operating system and encrypting everything. There has been an increase in ransomware attacks since 2012 since crypto-currency provides anonymity to the user and enables hackers to escape prosecution[5].
In recent times ransomware has become a major security threat. It has already affected millions of devices and continues to do so. Organisations and individuals, both are equally affected. Hospitals, security agencies, schools, nobody has been spared by ransomware attacks [6].

Crypto-ransomware is an interesting kind of new crime, one enabled by asymmetric cryptography, block-chaining systems, a large network of botnets, and the fact that the software that drives our computing devices always has exploitable bugs[7].

In this paper, we have designed and implemented a ransomware based on public key cryptosystem and named it RSALocker. We analyzed our RSALocker to ascertain the strength of encryption algorithms when it comes to ransomware. Also, the difficulty of recovering data after a ransomware attack was also analyzed. The usability of brute force attack and how long it would take to break an encryption was also checked.In Section II, we present principles behind ransomware.In Section III, we present working methodology of ransomware infecting a device and section IV describes the work flow of RSALocker. In Section V, we present the design and implementation of RSALocker. Further, in Section VI, we present the results obtained during implementation. And, in section VII, we present the analysis and observations from the results obtained. Finally, we conclude the paper in section VIII.

II. PRINCIPLES BEHIND RANSOMWARE

Ransomware is an implementation of cryptography to harm the network instead of protecting it. Ransomware uses various encryption technologies which are either the standard cryptosystems or their variations. Use of multiple standard and customized cryptosystems provides the malware security because new variants are not studied and hence not detected by malware analysis techniques [8].

It was Adam Young and Moti Yung who proposed the use of cryptography for ransomware in early 1990's [9]. Ransomware was not commercially used to extort money because financial systems could point to the person using the amount. Thus the cybercriminal could be easily traced and executed. The roadblock was cleared when crypto-currencies hit the market. They made the user anonymous and transactions could not be traced back to a real user. The anonymity provided by the crypto-currency gave attackers the perfect way to earn by using ransomware to extort money [10].

Ransomware is spread mainly by fake emails, mal-advertisements, pen-drives, and auto downloads on visiting certain websites etc. Once the file is downloaded on victim device, it starts auto-execution and encrypts data present on the device.

III. INFECTION PROCEDURE OF A RANSOMWARE

As a ransomware is downloaded, it has to follow a certain process to infect the system. To effectively do so, a crypto-ransomware is unpacked to disk and executed. Once the execution is complete, it is unpacked to memory. The payload is extracted from JPEG image. A suspended process is started in the background. The payload extracted from the JPEG resource is swapped with the image. As a suspended process is working, the payload is injected into explorer.exe. From there the payload is injected into svchost.exe and all shadow copies and system restore files are deleted [11]. Figure 1 depicts this procedure in a flow chart.

This is the process used by the most dangerous crypto-ransomware programs. They do not leave the victim with any option other than to pay the ransom. As the industry is evolving, stronger ransomware programs are getting most of the market share whichimplies higher chances of getting ransom from the victim.

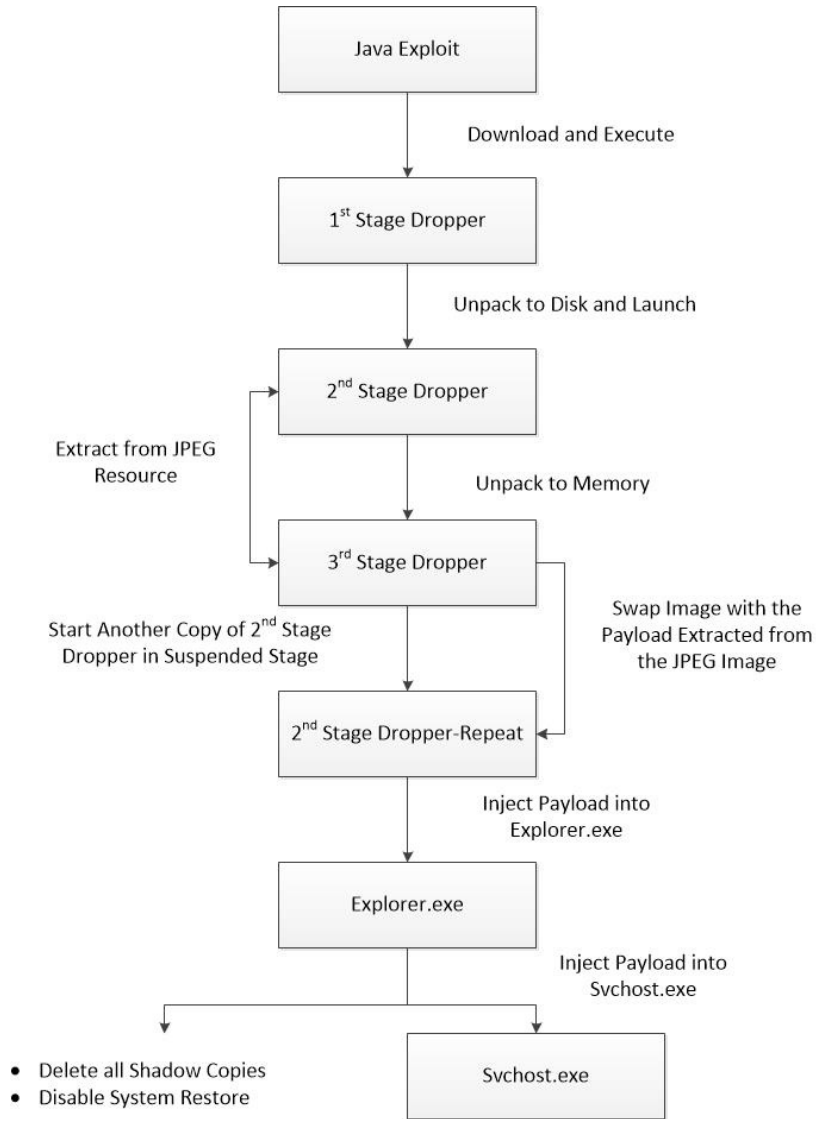


Fig. 1 CryptoWall Infection Workflow [11]

IV. WORK FLOW OF RSALOCKER

Figure 2 describes the work flow of RSALocker. The malware is unknowingly downloaded by the victim and its execution is started by a trigger activity. As the

installation is complete, keys are generated by the algorithm for data encryption. Once keys are generated, data is encrypted and the victim is informed by displaying a ransom message. The user is given the details of how to send the ransom and how much time he has to do it. If the ransom is paid, data is decrypted else all data is wiped out.

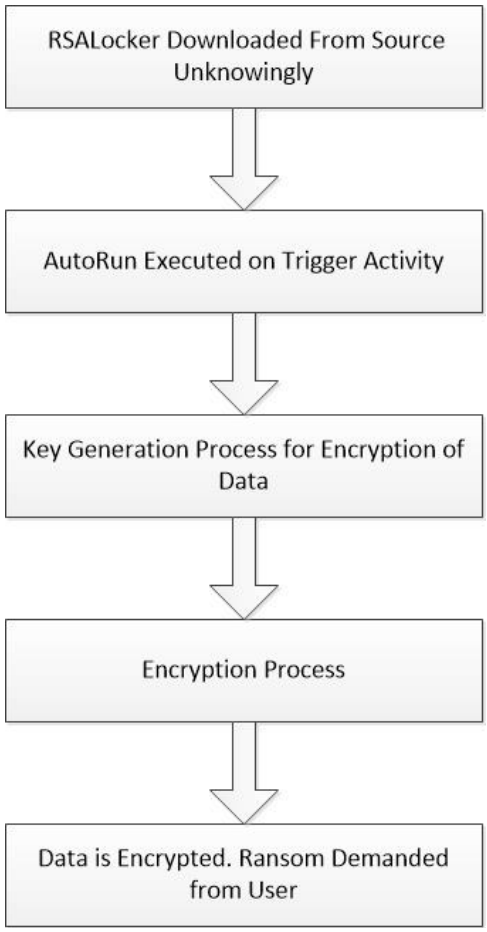


Fig. 2 Work Flow Diagram of RSALocker

V. IMPLEMENTATION OF RSALOCKER

The implementation of RSALocker was done using two different algorithms. One is the asymmetric public key cryptography algorithm, Rivest-Shamir-Adleman (RSA). The other algorithm is AutoRun.

RSA algorithm is used for encryption and decryption of data. The algorithm uses two random prime numbers and creates two keys, public key, and private key. One of the keys is used to encrypt the data and other to decrypt it. In this manner, it becomes difficult to get access to data. Brute force attack is possible but that would take a long time to execute and the costs incurred would be infeasible [12].

Ransomware programs are executable files. Any user does not download a file that could potentially harm his device. Hence the files have to be downloaded on victim devices without the knowledge of the user which is achieved by the use of AutoRun algorithm. AutoRun is an important feature of ransomware. Files are disguised as PDF's, documents or music files and sent as email attachments, or downloaded along with or instead of some required file. They are also downloaded due to trigger activities like clicking on an ad or visiting a website [13].

These algorithms are essential for the implementing RSALocker. Hardware and Software Requirements For implementation, Eclipse IDEusing Java Development Kit (JDK) and CrypTool were used.

Eclipse is an integrated development environment (IDE) used in computer programming and is the most widely used Java IDE. It is open source software developed by the Eclipse Foundation. In the current paper, Eclipse Luna was used. The development was done in Java language. Java was chosen because it has many inbuilt functions that proved helpful in easier and faster implementation of RSALocker.
The Java packages used were:

- java.math.*

It provides classes for performing integer and decimal arithmetic. It provides high precision [14].

- java.io.*It provides the system with input and output data streams, serialization and file system [15].
- java.util.*

It contains the collection framework, legacy connection classes, event model, date and time facilities, and miscellaneous utility classes [16].

- javax.swing.*

Provides a set of all-Java language components that work the same on all platforms [17].

CrypTool is also open source software. It is widely used to learn about cryptosystems and lets users analyze various cryptographic algorithms (cryptanalysis). In the implementation of RSALocker, version 2.1 of the tool was used to test the strength of the key and the time a brute force attack would take to break the encryption.
In hardware, an Intel Core i5-3230M was used with CPU running at 2.60 GHz with a primary memory of 4 GB and virtual memory of 7 GB.
The algorithms used are presented below.

Algorithm 1: RSA Algorithm

RSA_KeyGeneration()

INPUT:
Two large prime numbers, p and q

OUTPUT:
Public key components: {e, n}
Private key components: {d, n}

PROCEDURE:
$n \leftarrow p * q$

/ Compute Euler phi value of n*/*
$\phi(n) \leftarrow (p - 1) * (q - 1)$

Find a random number e, satisfying1 $<e<\phi(n)$ *and gcd* $(e, \phi(n)) = 1$

Compute a random number d, such that, $d \leftarrow e\text{-}1 \bmod (\phi(n))$

RSA_Encryption()

INPUT:
Plain text message, M $(<n)$

OUTPUT:
Cipher text, C

PROCEDURE:

- *A transfers a message (or plain text) M (<n) along with the public key to B.*
- *B encrypts the message using A's public key e to generate a cipher text, C.*

Cipher text, C ← Me mod n

RSA_Decryption( )
INPUT:
Cipher text, C
OUTPUT:
Decrypted plain text, P
PROCEDURE:
- *B transfers the cipher text C, to A.*
- *A decrypts the cipher text using its private key d to derive the plain text P.*

Plain text, P ← Cd mod n

Algorithm 2: Autorun Algorithm
The algorithm is critical in working of a ransomware. Ransomware is an unwanted file on a user device. Hence its installation has to be self-driven. Once the download is complete, the file sheds its extension to make it an executable file. The file then installs itself onto the device either immediately or after some trigger activity.

AutoRun()
INPUT:
executable RSALocker file
OUTPUT:
RSALocker file installed to trigger encryption activity.
PROCEDURE:
- *Infection of a device by download of malware*
- *Remove file extension (the file name changes from guide.exe.doc to guide.exe)*
- *Open guide.exe*
- *Install the malware*

VI. RESULTS

The implementation of RSALocker did not take into account the distribution of the malware or how it would reach the victim device. It only focused on the encryption part. In the current paper, we have used RSA algorithm to encrypt data. The data path is automatically set to default public folders. The path can be changed to include all drives or any specific drive.

Screenshots of the implementation phase are enclosed below.

Results of Iteration 1: Generation of Encryption and Decryption Key using RSA
In the first iteration, RSA algorithm was implemented to generate a public and private key. These keys are used to encrypt and decrypt the data. In a ransomware, the key generation is done on online communication servers and only the public key is sent to the victim device. The private key remains with the defaulters who use it to decrypt data when the ransom is received. Figure 3 shows the output screen displaying the keys generated.

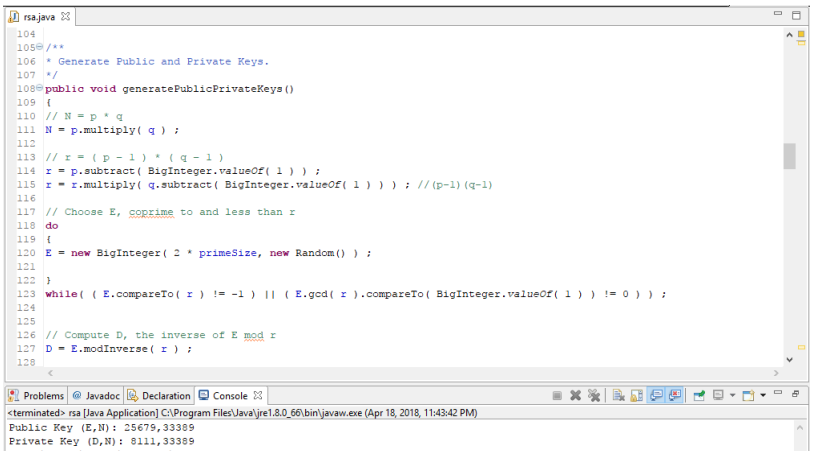


Fig. 3 Output Screen displaying the public and private key generated

Results of Iteration 2: Encryption of a Message with user input

Second iteration was done to encrypt plaintext to cipher text. The user was prompted to input the plaintext to be encrypted using a pop-up dialog box. The encrypted message was also displayed using a pop-up dialog box. Figure 4 and 5 display the input message and encrypted message generated.

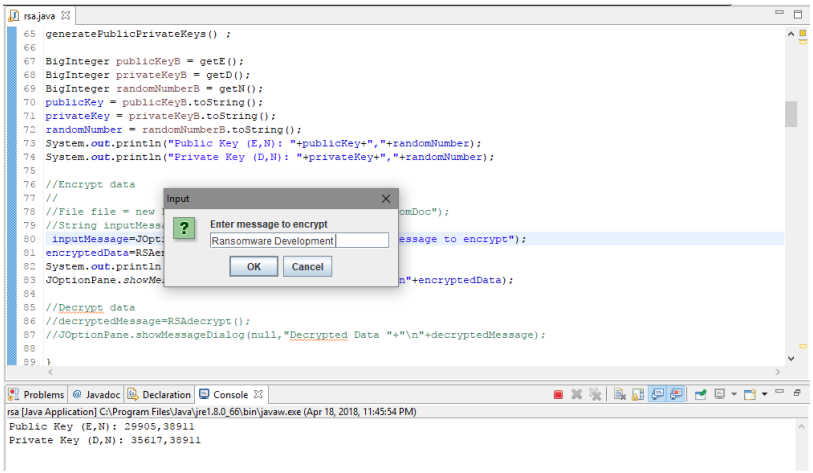


Fig. 4 Pop-up asking for plain text to be encrypted.

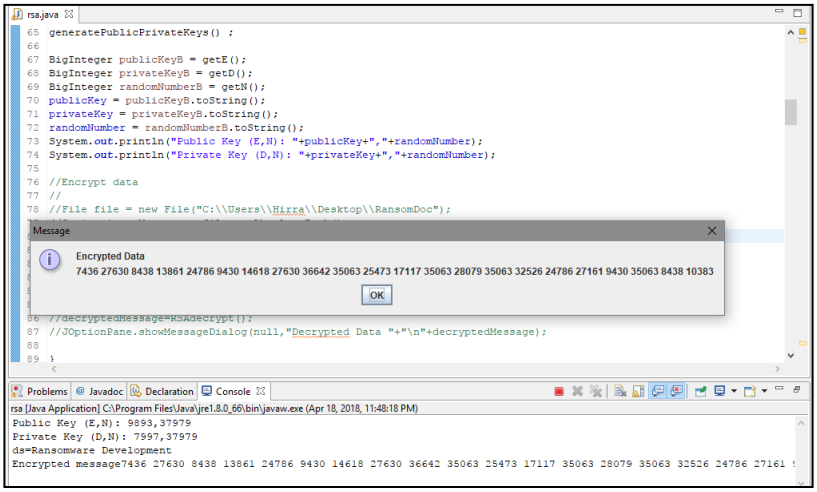


Fig. 5 Pop-up displaying the encrypted message.

Results of Iteration 3: Encryption of a file when file path is given

The third iteration was done to automatically read a file from a given path. The file path is read, the file is opened and the data is encrypted. Once the file was encrypted, an image was displayed as output informing the victim of his encrypted data. A demand for ransom was made and the account details were given out. Figure 6 shows the encryption of file and figure 7 shows the ransom demand screen.

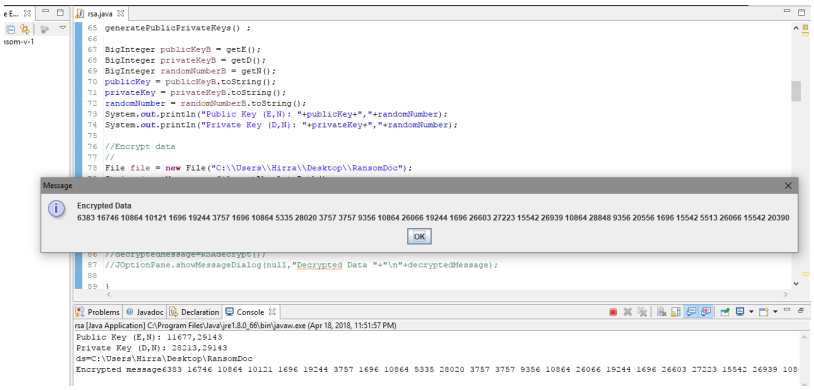


Fig. 6 Output showing encrypted data when file path was given as input.

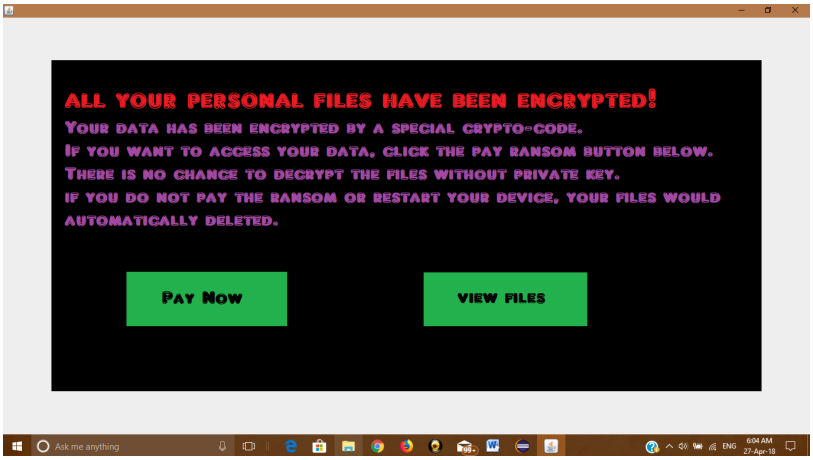


Fig. 7 Ransom Demand: Screenshot of the Output Screen after Encryption is complete.

VII. ANALYSIS AND DISCUSSION

In this paper, RSA was implemented by using only two prime numbers. It made RSALocker a basic ransomware using a known cryptosystem which has minimum security and can be recognized by malware detection algorithms. We could thus analyze how long brute force attack would take to break the encryption and recover data which could give us an idea of how difficult data recovery can be. As ransomware attacks are on the rise, data decryption is a challenge for the researchers of information security. The paper was an experiment on whether decryption without knowing the key can be done and the difficulty associated with it.

To better understand the ease oftaking a device hostage, I have evaluated the time taken by the RSALocker to generate keys encrypt data. Figure 8 (a) and (b) show that it took the malware only nanoseconds to generate keys and proceed to data encryption. Table 1 shows the time taken by RSALocker to generate keys and encrypt data at variable key sizes. The time taken to encrypt is small as the data to be encrypted is a single file. The time would increase as malware encrypts all the drives present on victim device. It is reported that a typical ransomware attack takes only 15 minutes to demand ransom from the time it has been downloaded on the victim computer [18].

TABLE I
Time taken to Generate Keys and Encrypt Data (nanoseconds)

Key Size (in bits)	Key Generation Time (in sec)	Encryption Time (in sec)
1024	490661	1306981
2048	451581	1041717
4096	453950	1460535
8192	490266	1494483

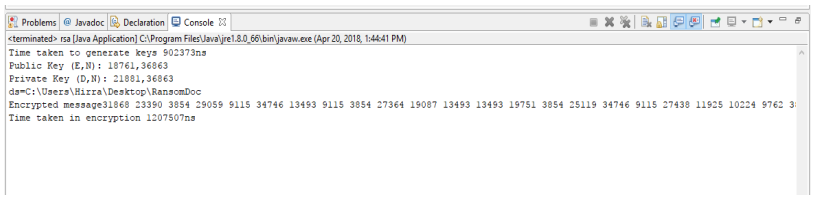


Fig. 8(a) Time taken to encrypt data

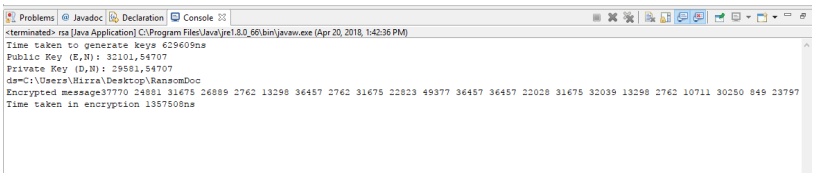


Fig. 8(b) Time taken to encrypt data

Figure 9 shows that the same encryption while done using CrypTool took a longer time (1s).

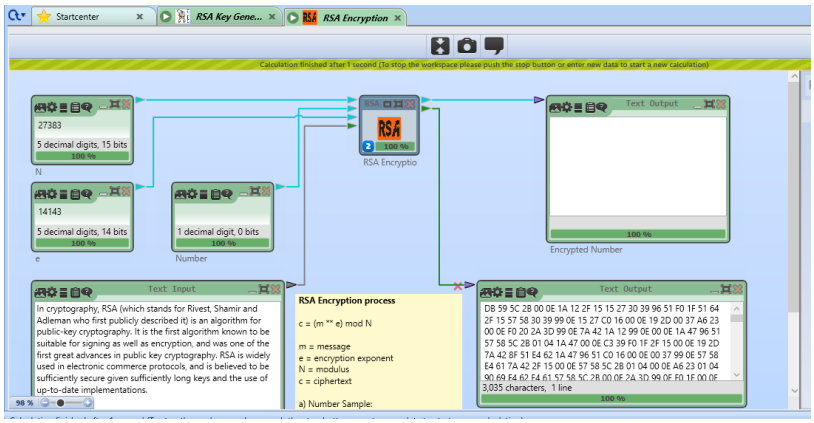


Fig. 9 Encryption time as calculated in CrypTool

Free tools for decryption of data have been made available by various governments over the internet. These tools are open access and can be used by anyone. Use of these tools does not guarantee data recovery. Hence in the advent of a ransomware attack, these tools might not prove helpful and the data would remain hostage unless a backup is available or ransom is paid. To analyze this, CrypTool was used to decrypt the data encrypted by RSALocker. The tool was not provided with the key but all other factors were made available. The software could still not decrypt the data nor could it generate the required key. It was because encryption is not easy to break using brute force and also because the device did not have enough resources available.

Figure 10 shows a screenshot of failed decryption of cipher text. The software could not work beyond 88.89% whichdepictsthe difficulty in data decryption. It was also a result of low resources available to the software.

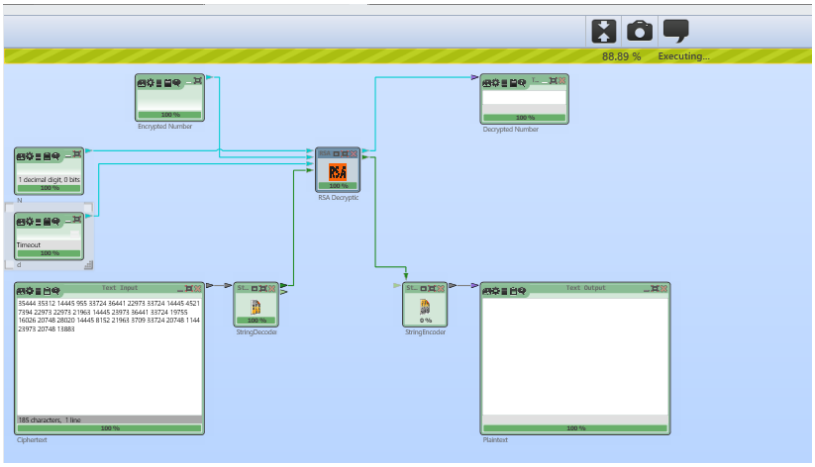


Fig. 10 Screenshot of CrypTool trying to decrypt a ciphertext without key

The results depict the difficulty faced by a victim when data is held hostage. Even though tools are available, they may not be successful in decryption of data. The scope of using cryptographic algorithms and functions for developing Ransomware is much huge. Due to the extensive possibility of using any cryptographic algorithms or functions, authors of Ransomware can develop multiple variants of Ransomware which may lead to difficulty in breaking it. Cryptanalysts have very less possibility of succeeding by usingany available predesigned decryption tool.

VIII. CONCLUSION

Information assets are prone to loss, theft or damage by various entities in today's digital world. There are always people looking for exploitable bugs to cash on. Malware is one such way in which the vulnerability is exploited and monetary benefits are gained. Previously, malware was found in the network as Virus, Trojan, Adware, Spyware, Key logger etc. Ransomware as a threat has been registered very recently which is mainly due to the widespread use of crypto-currencies which keep the user/account holder anonymous. Ransomware has created havoc because there is no sustainable solution currently available. A victim can be forced to pay a huge amount by holding critical data hostage like that of a hospital. Even though decryption tools are available online and can be accessed by anyone free of cost, they are not always effective. The reason is encryption algorithms used in ransomware are changed in every variant. The condition is also aggravated by the use of multiple algorithms and change in the sequence of algorithms used as they createhigh complexity which the decryption tools are unable to handle.

In the current paper, the complexity of ransomware was analysed by the implementation of RSALocker. The working, infection and the difficulty in decryption of user file was observed. The user file could not be decrypted by brute force attack even when basic algorithms were used for encryption. Hence the possibility of recovering the data by attacking was closed. Another possible solution is to create a decryption tool that is so strong that anyencrypted file can be decrypted by brute force attack. The other solution that is currently applicable is maintaining robust backups of data. In event of a ransomware attack, the victim can wipe the device clean and reinstall the operating system without paying any ransom.

REFERENCES

[1] W. Stallings, "Cryptography and network security:principles and practices", Fifth Edition, 2011, Prentice Hall, pp. 13-22, 257-287.

[2] B. Forouzan, "Cryptography and network security", 4thEdition, Tata McGraw Hills, 2007, pp. 931-960.

[3] B. Schneier, "Applied cryptography: protocols, algorithmsand source code in C", second edition, 1996, John Wiley andSons, pp. 1-10.

[4] [Online] "What is malware and how can we prevent it",Norton Security Center, https://in.norton.com/internetsecurity-malware.html [Last accessed 2018/04/18].

[5] H. Sultan, A. Khalique, S. Tanweer, S. I. Alam, "A Survey on Ransomware: Evolution, Growth and Impact", [Accepted, to be published in International Journal of Advanced Research in Computer Science, vol. 9, issue: March-April, 2018]

[6] "Understanding ransomware and ways to defeat it, WhitePaper", McAfee Labs, 2017.

[7] "Ransomware and phishing: how to avoid falling victim to these threats", White Paper, 19thJanuary 2017, Barracuda.

[8] Hilarie Orman, "Evil Offspring- Ransomware and Crypto Technology", White Paper, Purple Streak (2016), IEEE

[9] A. Young and M. Yung, "Cryptovirology: Extortion-Based Security Threats and Countermeasures," Proc. IEEE Symp. Security and Privacy, 1996.

[10] "Understanding ransomware and ways to defeat it, WhitePaper", McAfee Labs, 2017.

[11] VadimKotov, Mantej Singh Rajpal "Understanding Crypto-Ransomware: In-Depth Analysis of Most Popular Malware Families", White Paper, Bromium.

[12] Rivest, R.; Shamir, A.; Adleman, L. (February 1978). "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems". Communications of the ACM. vol.: 21, issue 2, pp. 120–126.

[13] Q. Chen, R. A. Bridges, "Automated behavioral analysis of malware: A case study of WannaCry ransomware", 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 2017.

[14] [Online] "Package java.math", https://docs.oracle.com/javase/7/docs/api/java/math/package-summary.html [Last accessed 2018/04/23]

[15] [Online] "Package java.io", https://docs.oracle.com/javase/7/docs/api/java/io/package-summary.html [Last accessed 2018/04/23]

[16] [Online] "Package java.util", https://docs.oracle.com/javase/7/docs/api/java/util/package-summary.html [Last accessed 2018/04/23]

[17] [Online] "Package javax.swing", https://docs.oracle.com/javase/7/docs/api/javax/swing/package-summary.html [Last accessed 2018/04/23]

[18] J. Crowe, "Ransomware Growth by the Numbers: Ransomware Statistics 2017", Barkley, June 2017

Proceedings of the 1st International Conference on ICT for Digital, Smart, and Sustainable Development
Department of CSE, Jamia Hamdard, March 15-16, 2019
ISBN. 978-93-5351-608-6

A Review: Implementation of Internet of Things for Smart Healthcare Systems

Garima Ahuja
Department of Computer Science and Engineering
Jamia Hamdard
New Delhi

Ibrahim Haleem Khan
Department of Computer Science and Engineering
Jamia Hamdard
New Delhi
ibrahimhaleemkhan.ihk@gmail.com

Jawed Ahmed *
Department of Computer Science and Engineering
Jamia Hamdard
New Delhi

Asfia Aziz
Department of Computer Science and Engineering
Jamia Hamdard
New Delhi

Abstract - *With an exponential increase in the network by the interaction of the Internet-enabled devices, data-sensing, information pro- cessing, and networking technologies have been boosted to over- come the challenges of human society. There has been a substan- tial increase in the amount of data that have been produced over the years as the number of the Internet of things (IoT) devices increases. There are multifarious applications of the Internet of things including healthcare. The idea of enhancing the healthcare facilities with the IOT can stimulate potential benefits prime con- cern being the increased patients' satisfaction. People across the world have begun to adopt such solutions like biosensors etc. This can empower the quality of care with daily examining the data collected by the patients. Iot can have many intriguing applica- tions as the devices become much smarter. IoT is a combination of advance and simple sensors with enough computational power to provide and analysis tasks. It has seen the much of the infor- mation generated on internet is because of the humans. The infor- mation provided by IOT will have a lot of applications, including healthcare, which is the primary focus of this work. These smarter healthcare systems will automatically gather the data and provide reports and in some cases identifies where a health worker is needed.*

IoT feeds the development of smart healthcare due to its cost, easy access and programmability. IoT is highly configurable, and due to its nature makes it more interesting to implement and eval- uate it. The decision tree for implemental smart healthcare using IoT involves three main parts defining the characteristics of the devices needed then the viability of the device like XBEE/PI/INTEL then the final steps of designing the system from how the information is processed and parsed with security in mind. The longevity is also the main factor for frequent updating for anywhere, anytime, anybody services. The most sig- nificant challenge is to make this information readily available to consumers regarding their personal health in the most efficient and accurate way. Some Iot platforms developed are Google Fit, Apple healthkit and Samsung healthcare systems.

*Corresponding Author

The health monitoring system implemented by Google is Google Fit. It is a set of application program interfaces that ag- gregates data from various sources. Now the data that has been collected is due to the modern sensors or M2M sensors that pro- vide the data to be analyzed in real time. The bio optic sensors have been easily handled by the growing progress of the sensor technology. The data analysis module performs machine learning. Machine learning is the science of developing computers to act without being explicitly programmed. Therefore, with the ad- vancement in machine learning many problems have been catered by the implementation of its algorithms of gesture recognition, voice detection etc.

It is cited that that billions of sensors and actuators will be con- nected to the Internet in just few years for collection and analysis of various types of data i.e. unstructured, semi-structured and structured data from various multiple heterogeneous sources. IoT is applied in areas such as smart city, healthcare, manufacturing industry, Automation (Home), energy and utilities, smart electric grids, transportation system, and traffic & tech management. A digital identity for each patient will ensure the personalization of healthcare

Keywords *- IoT,Smart IoT, Smart Healthcare, Automation, IoT person- alizationkey index terms here.*

I. INTRODUCTION

Healthcare and technology have always been closely related. Data produced from various biosensors adhered to users is made accessible to doctors to examine the vital health issues from any- where to help them in decision making to assist various healthcare workers. Various sensors like body-worn or implanted sensors do this task. These will deliver a continuous data stream for better decision making.

This can save the lives of many people in emergencies such as asthma, heart failure, diabetes etc. with the help of real-time mon- itoring. Some activities like monitoring persons through their sleep and many more examinations that people barely know they can be examined from a health perspective from such activities can also help them maintain their health in the long run. Connec- tivity protocols such as Bluetooth LE, Wi-Fi etc. can bring revo- lution in the way data is sensed from the health angle. IOT devices cut the need to store raw data in any data centers. It allows the users to use the data just after it is generated (called as in-memory analytics). IOT also help in real-time alerting, tracking, and mon- itoring, which permits hands-on treatments, better accuracy, and improve complete patient care delivery results. This will inevita- bly lower the cost expended of healthcare. IoT mends the way the provisions are delivered to the healthcare industry.

These technologies improve the product and can bring a revo- lution. Various stakeholders are stepping the efforts in this to ben- efit others and get benefitted by the same. It can increase work- force productivity, save costs, create new business models and improve collaboration between doctors and their parents. These IoT devices empowers every individual as cited by following the basic healthcare principle of "the suitable care for the right person at the right time", which can lead to increased customers' satis- faction and making healthcare cost effective.

An efficient healthcare service should deal with prevention, homecare, early ailment detection with an objective of cost effec- tive solutions. Healthcare systems make use of interconnected smart and uniquely identifiable devices to establish an IoT based utilitarian network for healthcare analysis, patient monitoring and automatically identifying situations where a physician's guidance is needed.

Healthcare is one of the basic needs of any person. The lack of effective medical facilities due to physicians' inappropriate be- havior in terms of care and money with every patient, this has become the need to provide customers with some technology aid so that they can cut the high clinical costs.

M- healthcare or mobile healthcare is also suggested .The dif- ferent technologies used in m-health such as GPRS, Wi-Fi, WLAN, ZigBeeX, Bluetooth+ and addresses future implementa- tions issues in these areas from the healthcare point of view.

II. RELATED WORKS

Much research has been happening since the past few years in the fields which aim at finding the best solutions that can be used to aid various problems of the society. There have been many buzz words these days out of which ranking on the top are big data. IoT is a significant contribution that has resulted in a large amount of big data. IoT has helped in enabling consumer self- service [1]. IoT can furnish a variety of benefits for effectiveness in real-time [1]. Some researchers have also presented the IT healthcare networks which serve a backbone for developing in- sights [4]. A list of various disease-specific sensors has been pre- sented. For diseases like diabetes, heartbeat monitoring, BP mon- itoring, Wheelchair management, cough detection, ECG self- monitoring, Eyecare Plus, Skin vision, and asthma tracker, fall detector and the list go on. Iot based healthcare devices include Fitbit bands, baby monitor, halo band, sync smart band, health Bp etc., [4]. Primary nursing care has also taken the assistance of IoT based devices [6]. It has also been presented in terms of three lay- ers namely perception layer, gateway layer and cloud layer. Gate- way layer has sensors connected to it. Providing more compre- hensive survey is the prime objective of the cloud layer. Wearable neck-cuff sleep detection tool is used for diagnosis of sleep in the process of health examination. Some sensors can also monitor se- cretions, analyses decisions and report the healthcare workers and nurses about the medical conditions [6]. Fall monitoring system provides instant positioning information. A hand hygiene device regularly notifies about the hygiene of the hands. Nursing calling systems have also been installed in the clinics [6]. The biggest challenge is to have accurate reports for the safety and health of the patients to enhance the quality of care.

The field of smart healthcare has evolved drastically with the implementation of iotHealthcare with smart discusses the ra- tionale of implementation of IoT smart devices in the medical field with advanced algorithms and data analysis techniques to- ward smart healthcare. The key points and finding majorly are dependent on statistical analysis with a multivariable approach. The use of Data analysis, Cloud Computing, IoT & sensors tech- nology are the enriching variables for IoT in smart healthcare building. [5]

Integration of the IoT technology with E-Health or smart Health solutions is the main focus here, that is, to show how IoT's central technology (RFID) is incorporated with patients' Elec- tronic Medical Records to allow & enable each patient have se- cured and easy access to his or her medical records in various health centers [2]. Distinguished characters for iotHealthcare with its logical methodology and analyzed architecture are stated as the Stability, continuity, confidentiality, reliability and efficiency. IoT is soon going to become a major part of the healthcare sys- tem.[5] The Tech like RFID[2] and Sensor technology[5] creating such devices involved different engineering levels of develop- ment like the dominant Raspberry, Intel Galileo board [3] and fre- quently adapted sensors like XB24-B[3], RFID[5,2] , LM35[3] and servers for intermediate connection through a channel [5].The discussing for creating stable interconnected devices with a pre- dominant aim of establishing the IoT in the healthcare field with accurate and acute results Lessing the complications with the cur- rently established man housed systems. The details can easily be fetched from the client or the doctor's computer without any has- sle after the housing of the smart IoT healthcare. The in-hospital patient monitoring system [3], the Data delivery and device [2] and design characteristics [5] solves the challenges or gives us a unique standpoint of fully developing a successfully IoT system as IoT is soon going to become a significant part of the healthcare system. The multifaceted approach towards the

implementation is quite fascinating and early detection of diseases and its preven- tion. The details finding are applied and systematically analysed with detailed data from sensors talking or sharing data over the network. [2] The safe integration of these components with five designated features can be used to build smart and exciting healthcare systems. [5]

REFERENCES

[1] Paul bronus, marlin janssen, (2015) effects of the internet of things (iot): a systematic review of the benefits and risks
[2] Mohd Ibrahim Bhat, Shariq Ahmad, Asif Amin, Suhail Ashraf (2017) E-Health with Internet of Things
[3] Ravi Kishore Kodall, Lakshmi Boppana (2015), An implemen- tation of IoT for healthcare
[4] S. M. Riazul islam, daehan kwak, md. Humaun kabir, mahmud hossain, and kyung-sup kwak (2015), The Internet of Things for Health Care: A Comprehensive Survey
[5] Joon-Soo Jeong, Oakyoung Han and Yen -You You (2016), A Design Characteristics of Smart Healthcare System as the IoT Application
[6] RiittaMieronkoskia, ImanAzimib, Amir M.Rahmanicd, Riku- Aantaaef, PasiLiljebergb, Sanna Salanteräaf (2017), The Inter- net of Things for Basic Nursing Care

Author Index