Universidad de Oviedo

CMR
INSTITUTE OF TECHNOLOGY
EXPLORE TO INVENT

International Journal of
MACHINE LEARNING AND NETWORKED COLLABORATIVE ENGINEERING

Vol.2 No.4 2018

PREFACE

The International Journal of Machine Learning and Networked Collaborative Engineering (IJMLNCE) with ISSN: 2581-3242is in a clear stage of expansion.The journal now appears in popular indexes such as BASE (Bielefeld Academic Search Engine), CrossRef, CiteFactor, DRJI, Google Scholar, Index Copernicus, J-Gate, Portico, PKP-Index, ROAD, Scilit and Socolar. After two years of hard work we are proud to present the sixth volume of the journal, Volume No-02 Issue No-04, with other five high quality works written by international authors and covering different aspects related to machine learning and collaborative engineering.

Chauhan et al. [1] published a work entitled "IoT Based Intelligent Vehicle Parking Solution System".In this paper, authors present a vehicle parking solution based on the Internet of Things through four different layers to compose the parking system: sensor, hardware, cloud and application. The main idea is that users are updated in real time on the available spaces near the destination allowing them to choose the one more suitable for their needs.

Mahmud et al. [2] published a work entitled "Domestic Mechanization System with IoT and Robotics". In this work, authors discuss home automation based on the Internet of Things focusing on three different projects: a smart window, a smart almirah and a smart bookshelf. They pay special attention to the smart window, which can be controlled in accordance with the weather conditions, the house temperature and the proper balance of gas in the air.

Sharma et al. [3] published a work entitled "DNA Based Storage: Introduction, Characteristics, Applications and Challenges". This study describes how the domain of knowledge of storage systems based on how the DNA works, since it is a viable alternative for conventional methods. They review the past, the current state of the art, with the advantages and drawbacks, and they also explore different challenges that would be interesting to overcome in the future.

Dash and Mohanty [4] published a work entitled "A Machine Learning Approach for Speech Detection in Modern Wireless Communication Environment". Authors propose a technique that improves the intelligibility of speech quality in noise environments. To that end, authors propose the use of different elements like an OFDM modulation based communication system, a neural network model of RBFN and different parameters such as energy and fundamental frequency.

Gupta et al. [5] published a work entitled "Study of Concurrency Control Techniques in Distributed DBMS". In this paper, authors present and discuss various lock-based concurrency control techniques for distributed data base management systems. They also show a comparative study of various two phase locking based concurrency control techniques. The focus is on proposing a proper concurrency control technique to maintain the integrity of database systems.

REFERENCES

[1] Chauhan, V., Sharma, S., & Manju Khari (2018). IoT Based Intelligent Vehicle Parking Solution System. *International Journal of Machine Learning and Networked Collaborative Engineering*, 2(04) pp 143-150.

https://doi.org/10.30991/IJMLNCE.2018v02i04.001

[2] Mahmud S., Ferdous M. F., Tasnim L., & Chakraborty D. (2018). Domestic Mechanization System with IoT and Robotics. *International Journal of Machine Learning and Networked Collaborative Engineering*, 2(04) pp 151-162.

https://doi.org/10.30991/IJMLNCE.2018v02i04.002

[3] Sharma, D. K., Kumar S., & Kumar A. (2018). DNA based storage: Introduction, Characteristics, Applications and Challenges. *International Journal of Machine Learning and Networked Collaborative Engineering*, 2(04) pp 163-169.

https://doi.org/10.30991/IJMLNCE.2018v02i04.003

[4] Dash S., & Mohanty M. N. (2018). A Machine Learning Approach for Speech Detection in Modern Wireless Communication Environment. International Journal of Machine Learning and Networked Collaborative Engineering, 2(04) pp 170-179.

https://doi.org/10.30991/IJMLNCE.2018v02i04.004

[5] Gupta M. K., Arora R. K.,& Bhati B. S. (2018). Study of Concurrency Control Techniques in Distributed DBMS. *International Journal of Machine Learning and Networked Collaborative Engineering*, 2(04) pp 180-187.

https://doi.org/10.30991/IJMLNCE.2018v02i04.005

Editor-in-Chief's

International Journal of Machine Learning and Networked Collaborative

Vijender Kumar Solanki, CMR Institute of Technology, Hdyerabad, TS, Insia

https://orcid.org/0000-0001-5784-1052

Vicente García-Díaz, University of Oviedo, Spain

https://orcid.org/0000-0003-2037-8548

Mihir Narayan Mohanty

- Ph.D, Professor,
- Department of ECE
- Siksha 'O' Anusandhan, (Deemed to be University), Odisha, India
- mihir.n.mohanty@gmail.com

Nguyen Cuong

- Ph.D, Lecturer,
- Department of Information Technology
- Quang Nam University, Vietnam
- nguyenhahuycuong@gmail.com

Nikhil Kumar Rajput

- Ph.D, Assistant Professor,
- Department of Information Technology
- Ramanujan College,University of Delhi, Delhi, India
- nikhilrajput@gmail.com

Prakash Singh Tanwar

- Head, Department of CS & CSE
- Madhav University,Pindwara
- Abu Road(Sirohi),
- Rajasthan, India
- pst.online@gmail.com

Rashmi Agarwal

- Ph.D, Professor, Department of Computer Applications
- Manav Rachna International Institute of Research And Studies, Haryana, India
- rashmi.fca@mriu.edu.in

Ruben Gonzalez Crepo

- Ph.D
- Director & Professor,
- Universidad Internacional de La Rioja (UNIR), Spain
- ruben.gonzalez@unir.net

Tamane S.Chandrashekhar

- Ph.D, Professor & Head ,
- Department of Information Technology
- MGM CoE, Aurangabad, MS, India
- sharvaree73@yahoo.com

Son Doan Trung

- Lecturer, Network & Security Department,
- University of People's Security, Hanoi, Vietnam
- son.doantrung@gmail.com

Tran Duc Tan

- Ph.D, Assciate Professor,
- Faculty of Electronics & Communications, VNU
- University of Engg. & Tech., Vietnam National University, Hanoi, tantd@vnu.edu.vn

INDEX

EDITORIAL OFFICE

Vicente Garcia Diaz, Ph.D.
Associate Professor
Department of Computer Science
University of Oviedo, Spain
garciavicente@uniovi.es

Vijender Kumar Solanki, Ph.D
Associate Professor
Department of Computer Science & Engineering
CMR Institute of Technology (Autonomous), Hyderabad, TS, India
vijendersolanki@cmritonline.ac.in,spesinfo@yahoo.com

CALL FOR PAPER

International Journal of Machine Learning and Networked Collaborative Engineering (IJMLNCE) with ISSN **2581-3242**, is a quarterly published, open access, peer-reviewed, international journal, focuses on publishing authentic and unpublished quality research papers in the emerging field of Machine Learning and Collaborative Engineering.

This is a scientific journal aims to promote highly qualitative and innovative research works, not limited to but focuses on the area of machine learning, collaborative

engineering, and allied areas. We publish only original research articles, review articles, and technical notes.

The journal is registered with CrossRef **10.30991/IJMLNCE**. Each accepted manuscript shall be assigned a unique digital object identifier for identification. The abbreviated key title for our journal allocated by ISSN is **Int. j. mach. learn. Networkedcollab.eng**.

The journal appears in popular indexes like BASE (Bielefeld Academic Search Engine), CNKI Scholar, CrossRef, CiteFactor, Dimensions, DRJI, Google Scholar, Index Copernicus, J-Gate, PKP-Index, Portico , ROAD, Scilit, Semantic Scholar, Socolar, WorldCat. Our journals manuscripts can also be found in the directory of ReserachGate & Kudos.

In this present technological era, the areas like machine intelligence, machine learning, and its associated domains are one of the most popular and demanding choices for the researchers as well as the industry personnel.

In last few years, numerous uses of machine learning and its related domain, has drawn ample attention of the people, that has generated a large number of applications in this field, making machine learning and collaborative engineering highly admired one.

Machine intelligence or machine learning is not a new concept. In terms of Artificial Intelligence, we were familiar with several aspects of the field, but nowadays with the introduction of machine learning, the use of this has been highly evolving, especially for improving the lifestyle of the human being.

There are numerous application areas of machine learning or machine intelligence, irrespective of any famous sector, from Healthcare, Space, Automation, Aviation industries etc. to entertainment industry and even in academia.

With this insight, IJMLNCE invites quality and authentic research works in form of manuscripts, for our forthcoming version. The Journal solicits latest research work carried

out by the researchers, academicians and industry practitioners, in the domain of machine learning and collaborative engineering.

Plagiarism is considered as a serious ethical and professional malpractice in case of IJMLNCE. So similarity check should be taken care at its highest level. The originality of the work will be highly appreciated. Each paper will be peer-reviewed by the expert reviewers and in a multilayer approach. The editorial board of the journal believes in high moral and ethical practices.

The authors are invited to submit their authentic and unpublished manuscripts online only. They will get final decision about the acceptance or rejection of the manuscript within 06-12 weeks from the date of submission of manuscripts. Once the manuscript gets accepted, author will be asked to complete the copyright agreement. An e-mail communication will also be sent to the author's registered mail ID, after verification of incorporation of all reviewers' comments, in term of their final camera-ready-paper; it will be published to the upcoming journal's volumes.

Please note that there is **no article processing cost** for publishing articles in this journal and the journal is open access.

IoT Based Intelligent Vehicle Parking Solution System

[a]Shikhar Sharma [*], [b]Vedant Chauhan, [c]Manju Khari

[a,b,c] AIACT&R, Geeta Colony, New Delhi

[a]sshikhar105@gmail.com, https://orcid.org/0000-0001-9520-9779

[b]vchauhan.chauhan2@gmail.com, https://orcid.org/0000-0002-1737-4370

[c]manjukhari@aiactr.ac.in, https://orcid.org/0000-0001-5395-5335

Abstract

With rapid increase in population in urban cities, availability of parking space is real issue. This parking issue lead to traffic and encroachment of roads for parking. With implementation of smart cities is real time development, smart parking is integral part of this development. Intelligent parking system describe in this paper solve the parking issue and fits in the smart city development, this system is based on cloud-based parking system where user is able to get location of parking spot with helps sensors network and cloud computing. The user is updated with real time data of available parking spot near their destination, and they can choose the spot according to their convenience. The main components of the system are sensor layer, hardware layer, cloud layer and application layer. The sensor layer is controlled by Arduino board or other system on chip which manages the data collected by sensors, this data is sent to cloud through hardware layer cloud layer manages the data accordingly and data is sent to users' application on the reception of request through application. This interconnection of all the layers is main aspect of IoT (Internet of Things). This system will help user to get the spot in hassle free and quick way.

Keywords

Internet of Things

Cloud Computing

Computing Centre

IoT

1. Introduction

The Internet of Things (IOT) is envisioned as a network that allows any device or any system to be connected anytime and anyplace. IOT can make the communication between the machine and machine or between the machine and humans much easier than by using the conventional ways. IOT extend the use of Internet by using it for the networking between different devices and also helping in monitoring the systems and tracking the devices (Things). There are two main words in IOT "Internet" and "Things". Internet means the vast network of   servers which are connected around the globe, computers, tablets and mobiles which use protocols followed internationally and connecting systems. Internet enables to share the data between different devices and systems. Things in literal sense means an object, or an idea or some action, but in IOT

---

[*] Corresponding author

Shikhar Sharma

Email: sshikhar105@gmail.com

"things" mean the devices which are connected together and form the net of devices which is called the internet of these devices.

IOT in general consist of, track the devices in remote areas, get data to the operating systems from the different devices and can organize and analyse the data. IOT gives the vision in which the devices and systems become smart and they start sensing and analysing their surroundings and start interacting with each other through different mediums. This smart and interactive nature of IOT can be used in many applications in real life which can make our life easier and help us steal some time for ourselves.

The most common idea that comes in mind when is thought about IOT and smart devices is "Smart Cities". And the biggest problems in our normal cities are traffic and insufficient vehicle parking. To make a smart city different from the normal city these problems have to be tackled smartly. So, in these situations IOT comes into play, using IOT, a intelligent parking system can be developed which will solve this parking issue. New applications in IoT are being built by developers due to recent advancement in embedded system which have low cost and are low powered.

One such concept that is proposed using IOT is intelligent parking system which is a reality now. The intelligent parking system concept that has been proposed includes the use of hardware(sensors), cloud storage, embedded chips and mobile application. All the above-mentioned components of the system are connected to each other forming the web or net which is used for sending, receiving and processing of data which is shared between different components. The systems helps the user in finding the free parking spot available, which is closest to the destination of the user, this helps in saving the time of the user which he/she will spend on searching for free parking spot , also this will help in reducing the traffic on the road as no unnecessary vehicle will be on the road wandering for finding parking space , this will reduce vehicles on the road and hence will help in keeping the check on traffic.

2. Background and Description

2.1. Background of intelligent parking solution system

In the present urban parking is a serious problem in populated country like India, parking is major problem for most urban streets and many cities lots by covering streets and roads' sides of busy routes. This creates traffic and wastes lots of time for people searching for parking spots.

For the motorist, the convenience of being able to find a place to park vehicle and then be on one's way is of great importance.

This invention is based on concept of using intelligent vehicle parking systems based on cloud technologies in can be used as major application in the smart cities. Because of whole business-oriented characteristics of services like this, system will be an intrinsic section for smart cities which will inclusive in general operative models of IoT.

Internet of Things (IoT) Strategic Research Roadmap which was proposed by the Cluster of European Research Projects (CERP), named CERP-IoT [1] in September 2009 was recommended by European Union (EU) for encouraging, sharing and publicising the research work in the field of IoT , especially application of sensors in field of IoT, like Intelligent Transport Systems (ITS) [2], smart health sensors wearables. Application like intelligent homes, smart cities, etc are also includes in sensors technology. Wireless sensors [3] will be used in this technology for infrastructure-to-vehicle communication, tracking and tracing services can be enabled by real time location detection system, which will remarkably advance the ITS applications. Each individual driving to a specific area could get a perfect parking spot with help of smart vehicle parking systems which find, allot, and assign the parking spot to user. These systems establish an important part of the ITS. The user application provides service to user to navigate to parking lot. These systems provides whole business-oriented services, these characteristics of system may serve as a substructure and a common business framework for a general operational platform of IoT.

2.2. Description of Intelligent Parking Solution System

This invention is based on cloud based intelligent parking system, where the IoT is the integral part of this system. Architecture of this system is inherited from adjoining of architecture of IoT and cloud computing i.e. it will contain both hardware application and cloud based mobile application. This system consists of three layers i.e. hardware layer, cloud layer and application layer. The hardware layer includes

sensors and a hardware processing unit which is Arduino UNO for this proposed system. This layer will work as infrastructure and will provide all its information and current statuses to database which is a part of cloud layer. The cloud layer is main processing unit of the system, it consists of computing centre which provide best available parking spots location to the user on basis of user location request sent through application. The application layer consists of user interface which takes user information and current location of user and send it cloud layer and a navigation system to navigate user to available parking spot.
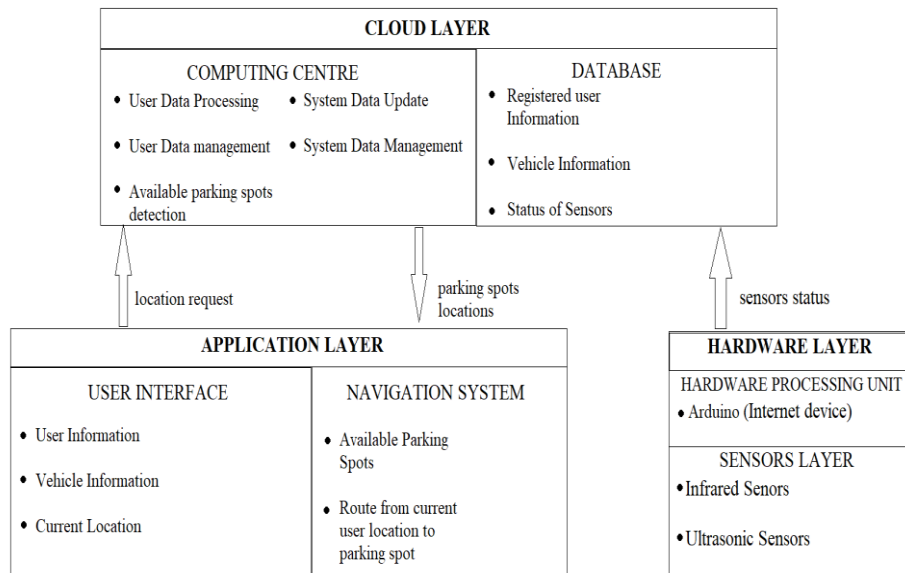
Fig.1: Architecture of Intelligent Parking Solution System

3. Detail Description of Components

3.1. Hardware Layer

3.1.1. Parking sensors

As seen from above fig1. The hardware layer consists of sensors layer. For parking system, there were options of many different types of sensors like ultrasonic sensors, infrared, passive infrared sensors [4] etc. These sensors detect the vehicles' presence and inform the system, whether the specific parking spot is vacant or occupied. In this case the infrared sensors were used because they are more accurate and efficient and can form heat image of the vehicle that is parked at the spot.

The infrared sensors are connected to the Arduino UNO boards. The sensors are externally powered by 5volt power supply through a battery, the sensors can also be powered through the Arduino boards but that would not be very efficient so it is more preferable to use external 5 volt supply(battery) to power the sensors. In this system Infrared sensors are being used. These sensors are used in detecting the presence of an object in front of it. This detection is done by either emitting or detecting the radiations by the body which is to be detected. Next, detection the focused radiation is done by infrared sensors. Then amplification of the output from the sensors is required which is done by pre-amplifiers coupled with circuitry. It is because output is very small.

3.1.2. Hardware processing unit

The processing unit comprises of the Arduino UNO board which is a processor on the chip. The work of the Arduino board is to collect the information from the sensors, process it and then send it to the cloud. It works as the intermediate between the sensors and the cloud. The data is collected by the Arduino through the input and output pins which are directly connected to the sensors, the data sent to the Arduino is then uploaded to the cloud through the internet device. Once the data is uploaded to the cloud storage then it is available to the servers and can be accessed by the app.

The Arduino Uno is a microcontroller [5] board. It is based on the ATmega328P microcontroller. It has a 16 MHz quartz crystal and a USB connection. A power jack for power supply and an ICSP header. To support the microcontroller, there are every basic element present on Arduino Uno which is required. Connection to computer can simply done through a USB cable. Also, it can be externally powered through a dc or ac source.

3.2. Cloud Layer

Cloud layer consist of two main component one is database and other is computing centre [6] as shown in fig1.

3.2.1. Database unit

In cloud, there is data of available vehicle parking lots, vehicle parking area. It also includes data of vehicles' location, users' location and their profiles, etc. It provides data storage and computation of resources of vehicle parking system. The database stores most recent data usually to assist real-time problems. On other hand the history data gets accumulated(warehousing) [7] in the cloud database. Various Map and Reduce algorithms are used for computing. One of which is recommendation algorithm for providing the suitable vehicle parking lots to users. To build an effective and scalable system, for quick and reliable decision-making process cloud system is used.

The users request locations of parking lots (on their mobile devices) from the closest location via mobile application through internet connection. The organisation of cloud-based system is in such a manner that the request is directed to the Computing Centre. Computing Centre is decision centre, the decision for the most appropriate, quickest and nearest way of parking lot to each user is made. Which is according to his/her current individual location and vehicle.

3.2.2. Computing Centre

The Computing Centre is mainly concerned with the service content and its creation. The deployment and operation of the service is also main part of Computing Centre. While control and execution of the service is also done by Computing Centre. In addition to this, some support functions are present that is initially required when each request is created, for example location management, vehicles' profile management, parking lot list etc. A depository of all profiles of both users and their vehicle is present. If there are any changes which are made by the individual user in his/her profile, then these changes or any other change in vehicle's profile are directed to the Computing Centre through user's mobile, where the source is updated. There are copies of all recently used profiles. Also, there are copies of profile updated by users. When a user is within the radius of a Computing Centre and access mobile application, the application which is installed in the user's device, and the Computing Centre mutually connect with each other. This process is facilitated through the request, feedback module within the Computing Centre. The Mobile application sends a request to the Computing Centre for user's Authorization. This request also includes a description and updates of user profile and user 's vehicle profile.

The Computing Centre directs this request to the database. This request along with the profile updates are processed. If the user is successfully authorized and there is parking spot available according to profile within the Computing Centre coverage area, a new record is generated for the user and that particular slot is occupied till the user vehicle leaves the slot.

3.3. Application Layer

On mobile device of user, the Mobile application provides the service utilization by the user. The mobile application allows the user unhindered access to the service. Through Mobile application user make a request for parking lot location while within the nearby range of Computing Centre. Application layer consist of user interface [8] for easy and simple interaction of cloud layer and user. The user interface will provide a detail form to the first-time user to fill up all his details to create an Id and get registered. This user Id can be used for getting the service of this system.

There will an option to set a default vehicle which is used by the user regularly, user can change the default vehicle. Once the user is registered, there will be an option to enter the location where user needs

parking. This location will be sent to Computing Centre as a location request, the computing centre will provide available parking spot in that parking location. User can select a particular spot and route to that particular spot will be provided from user current location.

User application would also require GPS navigation system [9] for direction of available parking spot.

4. Detail Working of Intelligent Parking Solution System

In the below section the idea of intelligent parking has been explained in more detailed and classified manner. In order to ensure the pliable and adaptive service provision which are present in each of the three layers of the system, it must interact so that user is satisfied in the best available way, no matter what type of request the user encounter. The following description describes the whole interaction that takes place between the different layers of architecture. The user get access to the information about the parking spots near them through mobile phones.

the content of the service must be more user friendly and gadget friendly so that it supports maximum devices. In the initial request, the user's mobile application sends the location and the type of the user car. Then this information is then sent to the Computing Centre which reads the user's vehicle description and location from the depository and according to this, find and provide the best available parking spot. Before using the application, every user has to register with vehicle details so that the system can provide them with best possible parking spot according to the vehicle. In this case, during the initial request the Mobile application sends a full description of the user's vehicle information. This request is forwarded to the Computing Centre which outline a sample interaction between components involved in the intelligent parking [10-11] system provision. As the user's vehicle enters the covered area of a Computing Centre. The Application which is installed in the user mobile and the Computing Centre mutually interact. The application sends a request to the Computing Centre for user's authorization. In this initial request, the user's mobile application sends the user data to the database and make the user profile and updates it whenever new data is provided. The Computing Centre registers the user in its database and updates the profile, this is done before computing user's request which come along user profile. However, profile is updated first time for the new user and if the user uses of the application regularly, the Mobile application could spontaneously predict the users request, i.e. once the application is opened, the Mobile application automatically send requests for parking location for the user's vehicle. The application directs the user request to the Computing Centre. Networks of sensors in the vehicle parking area constantly update the Computing Centre in real time so as to provide the users with real time availability of the parking area available. However, the system can be updating at larger interval during the evening period and during weekends as the rush for parking spots in this time or days is relatively low. Cloud consisted an ordered list of available parking spaces. These are according to their nearest distance to the user's desired destination. The Computing Centre then finds the current location of the user based on the location of the device. The Computing Centre then provides the best directions from user's current location to each of the available spaces.

Once the user receives the ordered list of parking spaces, s/he chooses a desired parking space. Directions to that space is given via application. Once a space is chosen, the Mobile application check the details of the space and displays accurate directions. An audio description of the route with visual directions on map will be best for this service as it will give least distraction to the user while driving. Occupancy of spot sensed by the network of sensors, any other users will not be given details of that space. When the user leaves the parking space, the sensor network sense this and send the confirmation to the Computing Centre. The Management Module within the Computing Centre accounts for the duration of the user's stay. Once the user/ service profiles have been updated, the service is terminated.

5. Application Results of Intelligent Parking Solution System

The above-mentioned invention was implemented in college campus area of AIACT&R Geeta Colony, Delhi. Sensors with Arduino board were installed in the parking spots available in the campus, available parking spots will be allocated for students and visitors in the campus area.

Staffs' parking area was excluded from the implementation and these spots will not be allocated to anybody.

Students and visitors with vehicle when used this system for parking, were able to find the spot in less time. This system worked in the same way as mentioned in description user through his/her mobile application sends a request to server with all details of vehicle.

The cloud server search within its database for available spots. Database of such spots was created through interactions with sensors installed in parking area. On detection of no vehicle in front of sensors sends the following information to the server. These spots will be marked as free spaces. The servers on the request of user send these spots list to the user application.

Now user can select the best spot for the vehicle and the location and route of that spot from the user's vehicle will be shown on the user application.

Unauthorised parking at spots available for authorised user can be avoided by further implementations of new techniques and idea in this system. User registered via mobile application to this system doing unauthorised parking could be requested via application with help of few enhancements in this system.

PROS

1) Reduce human effort in finding parking spot.

2) Reduce traffic on road.

3) Help developing smart city model.

CONS

1) Any nonuser cannot be stopped to occupy the space having our hardware unit.

2) The parking spot for users cannot be booked, as it is open and free parking for all.

6. Conclusion

At last conclusion on this invention is that; this Smart parking system will be a better solution for increasing parking issues of public in urban cities where parking is a major problem for general public going to a particular area. Implementation of this system on large scale could avoid wastage of time and traffic due to vehicle roaming for parking spot in a busy area.

This system can be integral part in implementation of smart cites, where public can park their vehicle efficiently without affecting movement of other free moving vehicles.

References

[1]. Vermesan, O.; Friess, P.; Guillemin, P.; Gusmeroli, S.; Sundmaeker, H.; Bassi, A.; Jubert, I.S.; Mazura, M.; Harrison, M.; Eisenhauer, M. Internet of things strategic research roadmap. *Int. Things-Global Technol. Soc. Trends* **2011**, *1*, 9–52.

[2]. Rudin-Brown, C.M. 'Intelligent' in-vehicle intelligent transport systems: Limiting behavioural adaptation through adaptive design. *IET Intell. Transp. Syst.* **2010**, *4*, 252–261.

[3]. Chi, Q., Yan, H., Zhang, C., Pang, Z., & Da Xu, L. (2014). A reconfigurable smart sensor interface for industrial WSN in IoT environment. *IEEE transactions on industrial informatics*, *10*(2), 1417-1425.

[4]. Mustapha, B., Zayegh, A., & Begg, R. K. (2013, December). Ultrasonic and infrared sensors performance in a wireless obstacle detection system. In *Artificial Intelligence, Modelling and Simulation (AIMS), 2013 1st International Conference on* (pp. 487-492). IEEE.

[5]. Badamasi, Y. A. (2014, September). The working principle of an Arduino. In *Electronics, computer and computation (icecco), 2014 11th international conference on* (pp. 1-4). IEEE.

[6]. Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., ... & Zaharia, M. (2010). A view of cloud computing. *Communications of the ACM*, *53*(4), 50-58.

[7]. Inmon, W. H., & Hackathorn, R. D. (1994). *Using the data warehouse* (Vol. 2). New York: Wiley.

[8]. Di Lecce, V., & Amato, A. (2011). Route planning and user interface for an advanced intelligent transport system. *IET intelligent transport systems*, *5*(3), 149-158.

[9]. Baus, J., Krüger, A., & Wahlster, W. (2002, January). A resource-adaptive mobile navigation system. In *Proceedings of the 7th international conference on Intelligent user interfaces* (pp. 15-22). ACM.

[10]. Pham, T. N., Tsai, M. F., Nguyen, D. B., Dow, C. R., & Deng, D. J. (2015). A cloud-based smart-parking system based on Internet-of-Things technologies. *IEEE Access*, *3*, 1581-1591.

[11]. Geng, Y., & Cassandras, C. G. (2011, October). A new "smart parking" system based on optimal resource allocation and reservations. In *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on* (pp. 979-984). IEEE.

Author's Biography

Shikhar Sharma is a student of Ambedkar Institute of Advanced Communication Technologies and Research, Under Govt. Of NCT Delhi affiliated with Guru Gobind Singh Indraprastha University, Delhi, India. He is pursuing B.Tech. in ECE. Currently he is working in field of IoT and its application. He has done training in LASTEC Lab of Defence Research and Development Organisation. He has experience of working under learned and experienced mentors.

Vedant Chauhan is a student of Ambedkar Institute of Advanced Communication Technologies and Research, Under Govt. Of NCT Delhi affiliated with Guru Gobind Singh Indraprastha University, Delhi, India. He is pursuing Btech in ECE. Currently he is working in field of IoT and its application. One of such application on which he is working is "Intelligent Parking System". he has experience of working under learned and experienced mentors.

Manju Khari an Assistant Professor in Ambedkar Institute of Advanced Communication Technology and Research, Under Govt. Of NCT Delhi affiliated with Guru Gobind Singh Indraprastha University, Delhi, India. She is also the Professor- In-charge of the IT Services of the Institute and has experience of more than twelve years in Network Planning & Management. She holds a Ph.D. in Computer Science & Engineering from National Institute Of Technology Patna and She received her master's degree in Information Security from Ambedkar Institute of Advanced Communication Technology and Research, formally this institute is known as Ambedkar Institute Of Technology affiliated with Guru Gobind Singh Indraprastha University, Delhi, India. Her research interests are software testing, software quality, software metrics, information security, optimization and nature-inspired algorithm. She has 70 published papers in refereed National/International Journals & Conferences (viz. IEEE, ACM, Springer, Inderscience, and Elsevier), 06 book chapters in a springer. She is also co-author of two books published by NCERT of XI and XII.

She has also delivered an expert talk, guest lecturers in International Conference and a member of technical program committee in an International Conference in

Hyderabad, she is also the life member of various International / National research societies viz IEEE, SDIWC, IAENG (Hong Kong), etc. Besides this, she associated with many International research organizations as Editor of Springer, Wiley and Elsevier books and Guest editor of International Journal of Advanced Intelligence paradigms, reviewer for International Journal of Forensic Engineering, Inder Science, and editorial board member of International Journal of Software Engineering and Knowledge Engineering, World Scientific.

How to Cite

Sharma, S., Chauhan, V., & Khari ,M. (2018). IoT Based Intelligent Vehicle Parking Solution System. *International Journal of Machine Learning and Networked Collaborative Engineering*, 2(04) pp 143-150.

https://doi.org/10.30991/IJMLNCE.2018v02i04.001

Domestic Mechanization System with IoT and Robotics

[a]Shakik Mahmud[*], [b]Engr. Mohammad Farhan Ferdous, [c]LamiyaTasnim,
[d]Dip Chakraborty

[b]Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan.

[a,c,d]Under-Graduate student , Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh

[a]smahmud172174@bscse.uiu.ac.bd, https://orcid.org/0000-0001-8589-7766
[b]jbratrc2018@gmail.com, https://orcid.org/0000-0001-8076-7428
[c]ltasnim172044@bscse.uiu.ac.bd, https://orcid.org/0000-0001-7323-3989
[d]dipchakraborty71@gmail.com, https://orcid.org/0000-0002-9604-0810

Abstract

In this paper, we discuss home automation IoT based and we show three projects about IoT. IOT or internet of things is an up and coming innovation which is an arrangement of interrelated computing devices, mechanical and computerized machines, articles, creatures or individuals that are given one kind of identifiers and the capacity to exchange information over a system without expecting human-to-human or human-to-PC association. It's an achievement thought it will change our whole world. New Horizons will begin in our lives by this. The premise of this research was to diminish the anguish of human. IOT based home mechanization can make the life an excessive amount of less demanding. In this research, we will attempt to interface the normal devices which are utilized as a part of a home. This paper is giving the design part of point of smart window, smart almirah, and smart bookshelf. One of the three projects is the smart windows system. It can be controlled in accordance with the weather conditions of the owner's house, and the house temperature, the proper balance of gas in the air. Using this system, the user will get comfortable weather at home, and will be able to predict any external danger to the environment. And it can be done by a mobile phone or an internet-enabled device.

Keywords

Internet of Things (IOT), Microcontroller, Smart Home, Smart Window, Smart almirah, Smart Self, Future of home automation

1. Introduction

At the present time, the internet of things (IOT) has been widely appreciated and responded. It is IOT which is building with a theme where two or more devices can exchange their data without human help. These devices are developed to collect and send data. In IoT here is two words one 'internet' another 'things'. The Internet is a system which is connected to a computer network. This network serves billions of users in the whole world. This network consists of millions of public, private, business, academic, business and government services with wireless and optical networking technologies.

[*] Corresponding author

Shakik Mahmud

Email: smahmud172174@bscse.uiu.ac.bd

In today world without some backward region, the total number of internet users is 3.54 billion, it's a huge amount. Things are real objects both living and non-living. It's not limited within electrical material or devices. So the phenomenon of things is that things are real objects in this physical or material world. A big number of researches can carry out about the IoT in the world. Around 25 billion distinguished detectable objects are expected to be a part of computing network in 2020. It's an Impressive amount. IoT has a great potential for flexibility and promises a great future.

Home automation or smart home system, which is being built with a theme of IoT. No IoT no home automation so when we develop a home with IoT, it's called a smart home. An automated home can reduce human effort like cleaning, cooking, etc. And it also saves your time and reduces your responsibility for the home.

As an example Suppose, by pressing just a button from your phone you can set an alarm in your table clock, which is just beside the bed. In the morning when the alarm clock rings this time, as usual, you will wake up and stop the alarm clock and the IoT system starts now, the clock will send a signal to the coffee maker which is in your kitchen while finishing the coffee, the coffee maker send a signal to your sandwich maker. You can observe the whole system in a website. You can skip a step or two if you want and that depends on you. Too much researching can involve more devices in a network.

2. IoT in Home Automation

Because of the large advantages smart home system gaining popularity day by day. In IOT system two or more devices interrelated with each other with a network and they also exchange their data without human or computer interaction. This method is also applicable to home automation. In home automation, some technology can be used, and this equipment provides a smart home system. RF-based system such as IEEE 802.11(wi-fi) it's too much popular wireless network at home; it has some features like location determination system, high data rate transmission, etc. Now we discuss some hardware components. These are given below:-

2.1. Bluetooth

Bluetooth is a mechanical determination for WPAN (Wireless Personal Area Network) that empowers the transmission of information and voice between various gadgets through a radiofrequency connect in the 2.4 GHz ISM band. The principal destinations to be accomplished with this standard are to encourage correspondence between cell phones, to evacuate links and connectors amongst them and to offer the likelihood to make little remote systems encouraging the synchronization of information between close to home gear. All Bluetooth gadgets have a novel address of 48 bits and a gadget name that permits the distinguishing proof of each other. That is a refreshed gear and furthermore eases. It's can work in a short range and a lithium curl cell battery can supplies its vitality for one year.

Fig 1: Bluetooth Module

(IEEE 802.15.4) or ZigBee: The ZigBee WSN includes XBee-S2 modules worked by Digit are arranged as end gadgets and convey remotely to an organizer as a work topology. ZigBee is the particular of an arrangement of abnormal state remote correspondence conventions for use with low-control advanced telecom in light of the standard of WPAN (Wireless Personal Area Network). Its objective is applications that require secure correspondences with low information rate and boost the life of batteries. Both are short-extend remote gear. ZigBee is the propelled adaptation of IEEE802.15.4. There are few home computerization frameworks that utilization ZigBee or Bluetooth for the remote association. With help of Wi-Fi.

2.2. LTE

Progressed for the fast portable system it has higher idea put and lower break.

Raspberry Pi: It's a solitary band PC. CPU 1.4 GHz 64/32-bit quad-center (refresh version).It is excessively agreeable for working framework and systems administration.

2.3. Arduino

Here we examine two kinds of Arduino

2.3.1. ArduinoATmega 2560

The Arduino Mega 2560 is a microcontroller board in view of the ATmega2560. It has 54 computerized input/yield pins (of which 14 can be utilized as PWM yields), 16 simple sources of info, 4 UARTs (equipment serial ports), a 16 MHz gem oscillator, a USB association, a power jack, an ICSP header and so forth. It contains everything expected to help the microcontroller; essentially associate it to a PC with a USB link or power it with an AC-to-DC connector or battery to begin[3]

2.3.2. Arduino Uno

Arduino UNO is known as the 'stock' Arduino. It is effectively versatile to all programming gadgets. All other Arduino sheets are diverse variants of the UNO board. It has all highlights resemble Arduino MEGA with less stick outs [3].
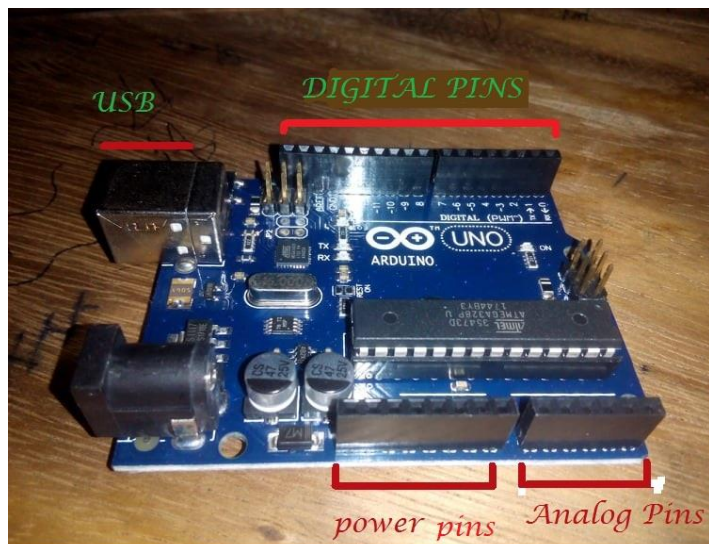

Fig 2: Arduino Uno

2.3.3. ESP8266

The ESP8266 is a minimal effort microchip with full TCP/IP, It has microcontroller capacity moreover. 32-bit microcontroller type, 160 MHz CPU. [10] The module can without much of a stretch associated with the Wi-Fi arrange. That is the significant purpose behind utilizing it in home computerization.

Relay Module: A hand-off board is utilized to control different apparatuses. Ordinarily, it's utilized with or without microcontrollers. Each 5V hand-off requirements 20mA driving current. It has LEDs for the sign of yield status.

2.3.4. RFID

Radio recurrence distinguishing proof (RFID) work programmed recognizable proof and information catch utilizing radio waves, a tag and a peruser. The tag can store more information between conventional scanner tags. Electronic Product Code (EPC), a worldwide RFID-based thing distinguishing proof framework created by the Auto-ID Center. 3 kinds of labels are utilized RFID. Inactive RFID labels depend on radio recurrence vitality exchanged from the peruser to the tag to control the tag; they are not battery-controlled. The usage is found in international IDs, supply chains, electronic tolls, thing level following and so forth. A functioning RFID labels have it's own battery supply and can correspondence with a peruser. These labels can contain outside sensors to screen, weight temperature, synthetics, area, and different conditions. RFID labels are utilized in assembling, research facilities, clinic and IT administration. Semi-uninvolved RFID labels utilize batteries to control the microchip while conveying by drawing power from the peruser. Dynamic and semi-latent RFID labels cost more than aloof labels. Remote sensor systems (WSN) Wireless sensor systems (WSN) comprise of spatially appropriated self-sufficient sensor-prepared gadgets to screen physical or natural conditions and can participate with RFID frameworks to all the more likely track the status of things, for example, their area, temperature, and developments. RFID equipment is currently especially in dynamic advancement around the globe.

3. Some project about Home Automation

In this paper, we discuss home automated projects, these are:

i) Smart window

ii) Smart Almirah

iii) Smart bookshelf

3.1. Smart window
3.1.1. Introduction and Components

The window is one of the principle parts of a house. Would you be able to envision a room without the window? Now and again answer is yes however is it legitimate? The brilliant window can make your home a keen house. It can decrease your physical exertion and duty. Think when the rain comes, The window will consequently near risky your home from water. Or on the other hand if the residue level is high, at that point, the window will naturally be shut. Or then again the outside temperature will cool or hot, at that point a similar framework will happen and the window will be shut. Presently we talk about the segment:-

i) Arduino/Raspberry Pi

ii) Wi-Fi or Bluetooth Module

iii) Sensors:

    (a) LM 35 Temperature Sensor

    (b) Optical Dust sensor

    (c) Water level sensor

    (d) Gas Sensor MQ-2 (SEN 00091)

i) 4 Channel Relay board

ii) GSM Shield SIM900A

iii) Servo Motor

iv) LCD Display with header (16x2)

v)   Bread Board

vi) Adapter

3.1.2. Implementation

We will divide all components in two categories.

1. Sensors

2. Module (without sensor)

The simple hardware diagram of a smart window is given below:
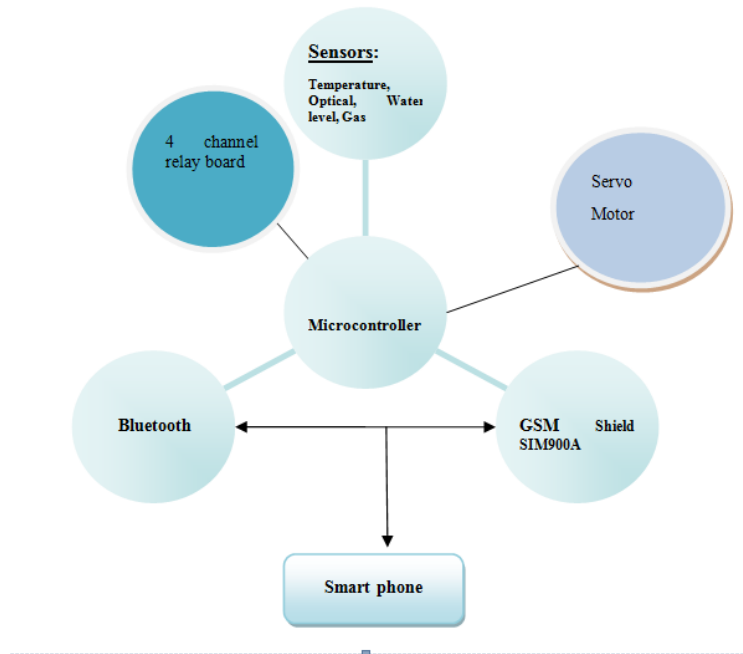
Fig 3: The basic hardware diagram of a smart window

In this venture, at first, we should set up the parts with the arduino board, in this means we typically fall into an issue which is stick mode. So we give the stick mode as much as we usually like.

3.1.3. For Dust Sensor

Here is the stick rundown of a residue sensor :

Stick 1: V-LED-3.3V (150 Ohms in the middle)

Stick 2: LED-GND-GND Pin

Stick 3: LED-Digital Pin

Stick 4: S-GND-GND Pin

Stick 5: VO – Analog Pin

Stick 6: VCC – 3.3V Pin (Direct)

3.1.4. For Bluetooth Module association with Arduino

i) Ardunio RX (PIN 0) to Bluetooth TX

ii) Ardunio TX (PIN 1)to Bluetooth RX

iii) Arduino 5V to Bluetooth VCC

iv) Arduino GND to Bluetooth GND

In the wake of associating all modules, we ought to prepare the product to work. For the product usage, in this venture we utilized Arduino, so we need to compose a code in Arduino IDE. It's called Sketch. For legitimate work, we need to partition the code into three areas.

i) Sensor part

ii) Modules part

iii) Application part(Android base)

We need to set appropriate rationale here, it chips away at rationale. The more grounded your rationale is, the more grounded it works.

3.2.   Smart Almirah

3.2.1.   Introduction and Components

Individuals store their valuable things and archives in almirah yet typically it's simple. In our task, the almirah will be computerized or brilliant. Cell phone and the almirah will recognize its proprietor by fingerprints. It will demonstrate how much space is free or how much space is full, and it will keep up the activity (if the crate was full you didn't put there any things). Intriguing right? In this way, how about we begin the task

Parts:

i) Arduino Mega

ii) Sonar sensor

iii) Servo Motor

iv) Wi-Fi Module

v) 4 channel Relay

vi) Fingerprint Recognition Module

vii) LED show (16*2)

viii) Piezo Buzzer

ix) Bread Board

x) Adapter

By and large, we think about every one of the parts appeared above aside from the Fingerprint Recognition Module. It's a module which is utilized to recognize and confirm individual with biometrics, mage rendering, estimation, highlight finding, and seeking. This module can without much of a stretch be associated with a microcontroller. This gadget works in 100mA with top 150mA. Baud rate is N*9600 bps. It has the ability to store 256 example prints.

3.2.2.   Diagram

Each part will be associated with a microcontroller and in the wake of composing code the gadget will work. Sonar sensor will be set in each cabinet for space-related data.
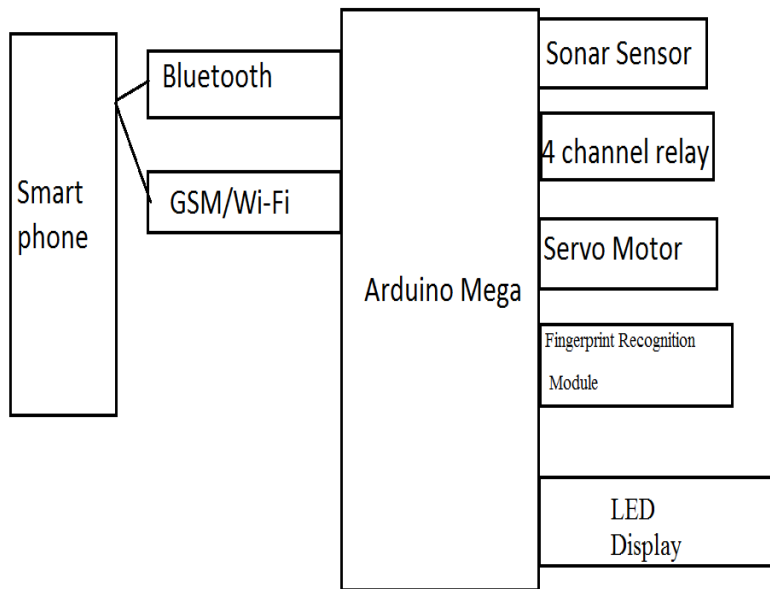
Fig 4: Smart Almirah

3.3. Smart Bookshelf
3.3.1. Introduction and Components

The idea of smart bookshelf is, the self can distinguish a book and toss it on the floor then a bin will get this then the container will convey it and put it on the table. It can spare our opportunity and can assuage us from finding the issue. It's an excessive amount of exhausting work however savvy self can decrease the issue.

Here are two sections one is an automated hand and another is a bushel which bases web-based after framework.

3.3.2. For Robotic hand

i) Arduino

ii) Servo Motor

iii) RFID

iv) Bluetooth Module

v) Wi-fi-Module

vi) Robotic Hand Structure

vii) Jumper wire

viii) Battery

Here RFID is utilized to distinguish books and Bluetooth utilized for client and wifi for interconnecting with the container.

3.3.3. For container

i) Arduino

ii) A container

iii) Wheel

iv) Motor

v)Motor module

vi) IR and Sonar sensor

vii) Jumper wire

viii) Battery.

ix) Wi-Fi and Bluetooth Module

Fundamentally, the framework made with a line devotee robot framework. Be that as it may, it can speak with Robotic hand. It will just work when a book tumbles down on it. For development, it will just take after the line following framework.

Robotic hand diagram

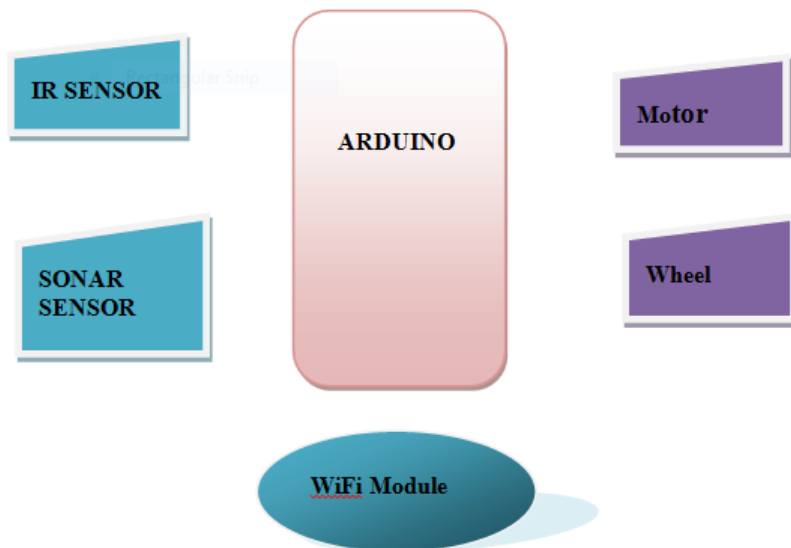Fig 5: Robotic Hand Diagram

For Basket

Fig 6: Basket Diagram

4. Future of home automation

Home computerization is developing step by step. Legitimate research and undertaking can make it more created. In 2020 it will be more famous and individuals will depend more on it since it diminishes human exertion and makes the existence straight. For its attention, top-level ventures should come to help, and the home computerized gadget ought to be made for business reason yet at a sensible cost. Be that as it may, there is a central issue about security and protection issues and there is another inquiry, is it a human-accommodating framework?

Presently we examine some trying future issue for IoT

4.1. Information Management issue

Data administration is an administration which incorporates all parts of information like arranging, taking care of, investigation, and store. IoT gadgets and modules accumulate loads of information for their work. For administration procedure, the server farm needs to store and process. The present circumstance of the server farm is disappointing. The framework structure of a server farm can't store individual and undertaking information. Not very many undertakings are fruitful in Data administration field. They can gather and store their IoT information in their system. Essentially, they just need information for task and framework reinforcement. For future days, server farms require more proficiency and enhance their procedure. Else, it will be an obstruction for IoT future.

4.2. Information Mining

The web of things will deliver a lot of information. On the off chance that we go to a general store and watch the production network which embraces RFID innovation. The configuration of crude RFID information is EPC, area, time, EPC(unique identifier) check the area of the peruse and perusing time. A RFID needs 18 bytes for another crude record. In this way, consider the store case once more. There are least 700000-600000 RFID labels. In the event that we figure each and every second least 11-12 GB RFID information will be created. What's more, the sum will achieve 450-500 TB for each day. In this way, now daily's information mining is excessively imperative. It's important to create compelling techniques for overseeing, dissecting and mining these information.

The information reliable with customary discrete information as well as the information created from sensors, area, development, temperature, and mugginess and so forth so on. Along these lines, the utilization of information mining devices turns into a need. For information mining, there must update the model of information mining. Multi-layer information mining can diminish the issue.

4.3. Protection Challenge

Accept the case with wise home computerization or other IoT based structure. This IoT based devices can convey a colossal proportion of data on the customer. These are an improvement, RFID and distinctive sensors data all of which can begin basic security concern. Security is must, without insurance, it will be valueless. IoT is the method for upgrading the overall public living quality. Clusters of web customers agreed that the upsides of splendid contraptions surpassed any security concerns. Thusly, we can express that the inevitable destiny of IoT will depend upon the affirmation of customer's security.

4.4. Security challenge

The quantity of developing associated gadgets is brought into IoT systems, the potential security danger raises. The latest the IoT enhances the efficiency and the organizations improve the nature of individuals' lives, the IoT will likewise build the potential assault surfaces for programmers. An ongoing report by Hewlett Packard (2014) uncovered that 70% of the most ordinarily utilized IoT gadgets contain genuine vulnerabilities. An absence of transport encryption, unreliable Web interfaces, lacking programming assurance, and deficient approval IoT gadgets are in under vulnerabilities. Likely, all things considered, every gadget seized 25 openings or dangers of trading off the home system. These gadgets don't utilize information encryption strategies by and large. Some of IoT applications support delicate infra-structures and imaginative administrations, for example, the keen matrix. Other IoT applications will progressively produce gigantic measures of individual information about the family unit, wellbeing, and money related status that undertakings will have the capacity to use for their organizations. The absence of security and protection will

make protection from the reception of the IoT by firms and people. Security difficulties might be by made appropriate strides like uncommon preparing for designers.

4.5. Chaos challenge

The disclosure of IoT advances like chips, sensors, remote and RFID and so forth advances is in a hyper-quickened development cycle that is speedier than the commonplace purchaser item in the advancement cycle. These are as yet contending with measures, deficient security, protection concern, complex interchanges framework and multiplying quantities of inadequately tried gadgets. On the off chance that the plan isn't done painstakingly, multi-reason gadgets can transform our lives into confusion. In a detached world, a little mistake does not cut tumble down a framework; be that as it may, in a hyper-associated world, a solitary blunder in one a player in a plan can cause disorder all through. The control frameworks and shrewd home applications made with interconnected sensors and specialized gadgets and controllers. In the event that a sensor of a home applications and control framework breakdowns, the controller may get a false flag, which may demonstrate deadly to the patient. It isn't hard to figure savvy home packs, for example, indoor regulators and private power meters separating or being assaulted by programmers, making startling wellbeing issues. The Internet transfer speed can get immersed with information movement of multiplying gadgets, making framework wide execution issues. A solitary gadget may have an unimportant issue, yet for the framework, all in all, the chain responses of other associated gadgets can wind up awful. To forestall mayhem in the hyper-associated IoT world, organizations need to bend over backward to decrease the unpredictability of associated frameworks, improve the security and institutionalization of uses, and certification the wellbeing and protection of clients whenever, anyplace, on any gadget.

Fortunately, there has been as of now many researches finished on IoT security concern. These accomplishments must be additionally extended as opposed to concentrating towards looking for the new conceivable security arrangements. Presently the security question is a future test of IoT and home computerization likewise in light of the fact that they associated with the cloud framework. So the marvel is lawful structures, legitimate controls and strategies must be formulated to guarantee stable advancement of the security advances.

Conclusion

The name of a revolution that is currently running is the IOT, which will have more branches in the future. Because of increasing the use of digital devices, strong reliance on human devices, time and effort to save human life, this revolution is accelerating. If the security of the cloud storage system and the controller system is improved, the IOT will work a lot easier. Smart windows will help to keep the environment in our home, which will work to remove the cell's bacteria. It also plays an important role in the release of toxic gases in the air. This is very important in the industry of the future. So we can see if we use the IOT in the right way, it will give us some great gifts, and in future it will be possible to lead the world of technology. Need more research with IOT, thinking that how it can be utilized for more welfare, so that business organizations and big industries will come forward, to invest so that it can be dreamed of something bigger in the future.

References

[1]. Nunberg, G. (2012) The Advent of the Internet: 12th April, Courses. Link: http://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1482944

[2]. https://statista.com/statistics/273018/number-of-internet-users-worldwide/

[3]. https://www.arduino.cc/en/Main/arduinoBoardMega

[4]. http://www.gartner.com/newsroom/id/2905717/

[5]. X. Zhao, "The strategy of smart home control system design based on wireless network," in Computer Engineering and Technology (ICCET),2010 2nd International Conference on, vol. 4, 2010, pp. V4-37

[6]. https://en.m.wikipedia.org/wiki/ESP8266#

[7]. C.Liu,Y.Zhang,J.Zeng,L.Peng,R.Chen,ResearchonDynamical Security Risk Assessment for the Internet of Things inspired by immunology, in Eighth International Conference on Natural Computation (ICNC), 2012

[8]. Rafiullah Khan, SarmadUllah Khan, R. Zaheer, S. Khan, Future Internet: The Internet of Things Architecture, Possible Applications and Key Challenges, in 10th International Conference on Frontiers of Information Technology (FIT 2012), 2012

[9]. B. Edson, "Creating the Internet of Your Things", Microsoft Corporation

[10]. E. J. Bruno, J. Kornwitz, P. Lombardi, and M. Samuelsson, "The Intelligent Internet of Things with Axeda and Oracle Java Embedded," no. June, 2014

[11]. Neng Wang and Lu Tan. Future Internet: The Internet of Things, International Conference on Advanced Computer Theory and Engineering (ICACTE). Pages 376- 380, 2010.

[12]. AmeetChauhan, SangeetaAgarwal "SMS based remote control System in International Journal of computer Science and Management, 2011.

[13]. P. S. Pandey, D.S. Chauhan, R. Singh, "The Real Time Hardware Design and simulation of moving message Display System Integrated with PLCC Modem" Innovative Systems Design and Engineering, ISSN 22221727 (Paper) ISSN 2222-2871 (Online), Vol. 3, No. 10, (2012).

[14]. TRUSTe. (2014). TRUSTe Internet of Things privacy index——US edition. Retrieved from http://www.truste.com/resources/ privacy-research/us-internet-of-things-index-2014

[15]. Sundmaeker, H., Guillemin, P., Friess, P., &Woelffle´, S. (2010). Vision and challenges for realising the Internet of Things. Accessible at http://www.researchgate.net/publication/ 228664767_Vision_and_challenges_for_realising_the_ Internet_of_Things

[16]. Hewlett Packard. (2014, July 29). HP study reveals 70 percent of Internet of Things devices vulnerable to attack. Retrieved from http://www8.hp.com/us/en/hp-news/press-release. html?id=1744676#.VOTykPnF-ok

[17]. Gartner. (2014, March 19). Gartner says the Internet of Things will transform the data center. Retrieved from http://www. gartner.com/newsroom/id/2684616

Author's Biography

Shakik Mahmud is an undergraduate student of Computer Science and Engineering department in United International University (UIU), Dhaka, Bangladesh. Recently he worked as a Deputy Technical Director at Japan Bangladesh Robotics and Advanced Technology Research Center (JBRATRC). He also passed his higher secondary exam from Cantonment Public School and College, Saidpur, Nilphamari.

Mohammad Farhan Ferdous is an Engineer. He completed his Diploma engineering in Mechanical branch from Dhaka Polytechnic Institute, Bangladesh. He has also completed B.Sc. in Mechatronics Engineering from World University of Bangladesh. He has completed his Masters in Information Science from Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan. He is the founder and Chairman of Japan-Bangladesh Robotics and Advanced Research Center (JBRATRC).

Dip Chakraborty is a graduation student of Computer Science in United International University. He completed HSC from Daffodil International College and passed SSC from Rayer Bazar High School.

Lamiya Tasnim is a B.Sc. engineering student department of Computer Science and Engineering in United International University (UIU), Dhaka, Bangladesh. She also passed her secondary and higher secondary exam from Cantonment Public School and College, Saidpur, Nilphamari.

How to Cite

DNA Based Storage: Introduction, Characteristics, Applications and Challenges

[a]Deepak Kumar Sharma[*], [b]Shiv Kumar, [c]Amit Kumar

[a,b,c]Division of Information Technology, Netaji Subhas University of Technology(formerly NSIT), New Delhi, India

[a]dk.sharma1982@yahoo.com, https://orcid.org/0000-0001-6117-3464

[b]shivk.it.16@nsit.net.in, https://orcid.org/0000-0003-4792-9207

[c]amitk.it.16@nsit.net.in, https://orcid.org/0000-0003-1380-9055

Abstract

Over the years, as humans have made progress, data has come to the forefront and has become one of the principal elements of life. No matter the field, all aspects of life are now dependent on data in one way or the other. Be it hospitals or financial institutions; sports teams or researchers, all operate on some form of data during their functioning. This ever-increasing dependency on data further leads to the need for its storage. The capability of present storage mechanisms is not able to keep up with the exponentially increasing demand. This along with other factors such as high setup costs, high maintenance charges, security and accessibility are pushing towards an alternative avenue of storage. DNA or the code of life is very similar to the binary based data systems that we operate on, hence is being looked at, as the alternative to conventional methods. This field has seen massive amounts of developments in the recent past and is finding a strong footing. Its theoretical capability to store all the data ever created in a finger sized device is one of the many factors, which makes it such an interesting field to study and know about. This paper describes how this domain of storage system(s) basically functions, the work done in this field in the past, its advantages and limitations along with the challenges that this domain needs to overcome to become practically viable bringing a paradigm shift in computing.

1. Introduction

In this era of information, where data is rightfully treated as one of the principal elements of life, one thing that cannot be ignored is the explosion in the rate of its creation and exchange. Fields like big data, data mining, human computer interaction, genomics, social interaction etc. are becoming increasingly popular around the world and are yielding enormous amounts of data and information every second. Enterprises are using the ability and insight provided by domains like Artificial Intelligence and Machine Learning to gain

---

[*] Corresponding author

Deepak Kumar Sharma

Email: adk.sharma1982@yahoo.com

significant amounts of knowledge, about the created data in order to produce better products and services, creating even more data in the process. From autonomous cars to smart home devices, our world is changing, and fast, transforming our understanding of everything around us and data is at the center of it all.

The realisation that over 90 percent of the worlds data ever created has come up in the last 2 years of civilisation[1] alone, is overwhelming. Humans are estimated to be creating around 20 quintillion bits of data every day, which is astounding [1]. Also, the fact that the rate at which data is being produced is speculated to increase exponentially in the years ahead is quite hard to digest [1]. Based on current capabilities and technologies, the storage systems are only capable of storing 18% of the world data and this amount will hit rock bottom at around 3% by 2030 and reduce to 0.5% by 2040. [2]

Apart from the worries of analysing, understanding and finding patterns in such big, unstructured and complex amounts of data, there lies one, seemingly trivial but important issue of its storage. Present approach of data storage which ranges from portable hard drives and flash drives to cloud-based data centers just won't be able to keep up with the increasing demand for storage space. Plus, not only do these methods require real estate and infrastructure which is indeed very limited, but they also add to the costs that come in the form of initial setup, electricity, manufacturing and maintenance. Another issue is that of security; even with multiple layers of security, if the data is on the internet, it can be hacked and retrieved. The severity of such a threat, cannot be taken lightly. With it also comes the issue of accessibility, i.e. if the data size is very large, then it can't be carried in flash drives and external hard drives, and the user must be connected to the internet to access the data present in some remote cloud-based service and, hence if the user is offline then carrying big data becomes a big problem. These solutions also run the danger of being corrupted when brought in touch with certain external environments, which brings in contention their ability to act as archival systems for very long periods of time.

Silicon along with other elements which are non-biodegradable, are also a concern to increasing pollution in the environment. Hence, it can be said, that it is time that we look for other alternative solutions to the ones presently employed for data storage. Otherwise, we are headed towards a crisis for which we are heavily unarmed.

One alternative that will be discussed in this paper is that of deoxyribonucleic acid (DNA) based storage [3], i.e. nature's own storage medium. They are of interest in this area due to some of their properties which include but are not limited to being highly-dense, sturdy and long lasting.

Just one gram of DNA is said to store about 455 exabytes of data [4], solving the real-estate and infrastructure side of problems and since that much amount of data can be localised to that little space, it allows users to carry their own data centers along with them wherever they go, providing them with better accessibility as well as returning physical ownership of their data. This also solves the problem of lack of privacy in conventional storage mediums.

Another aspect to be considered here is, the power required while working with DNA. It is very little as compared to conventional data storage solutions which require huge amounts of energy to maintain. [5]

It has also been seen that scientists have been able to extract information from DNA, known to be thousands of years old [6,7], hence making this form of storage very robust and immune to corruption. Generating copies of DNA is also possible using Polymerase Chain Reaction techniques, hence exchange of information does not get hindered if the transition takes place to this medium of storage. Also, since DNA is bio-degradable, it poses less of a threat to the environment as compared to its conventional counter-parts, which increases the life of the planet humans inhabit.

In the present paper, the structure of DNA and the concept of memory hierarchy are explained in brief in the second section which is used in section three to understand and review the work already done in this field. This is used in section four to understand the issues and challenges that DNA based storage faces in order to become the standard of storage around the world. We conclude by understanding the current situation and the future possibilities in the field.

2. Background

DNA or deoxyribonucleic acid [3] is an extremely long chain of molecules that contains all the information necessary for the functioning of any living cell. It is the basic unit of genome which is an

organisms complete set of genetic instructions, and hence considered the fundamental unit of life. Its structure (as shown in figure 1.) is said to be in the form of two strands made of molecules called nucleotides, intertwining to form a structure referred to as the double helix structure.

The nucleotides which make up these strands are themselves made up of some fundamental blocks, that are:

"one of four possible bases that include, Adenine (A), Cytosine (C) , Guanine (G) and Thymine (T) , a phosphate group and the five carbon sugar."

Each base has a corresponding complimentary base and it forms a pair with only that complimentary base on the opposite strand, through hydrogen covalent bonds and together these complimentary pairs are called base pairs (thymine compliments adenine using two H bonds and cytosine compliments guanine using three H bonds). The double helix structure is very stable due to its regular shape and stacking of layers. The different arrangement of these base pairs gives different properties to all living beings. From eye colour to better agility, from regeneration properties to genetic diseases, everything is governed by the arrangement of these pairs. Different combinations of these base pairs hence, although very simple in concept leads to millions of possibilities which is similar to how the binary codes employed in the computational systems of today, use Boolean logic to construct systems of such high complexity. The bases sequence can be seen analogous to a base 4 system i.e. 4 possible fundamental units to form complex combinations. This similarity with binary system is one of the prime reasons for the consideration of DNA as an alternative to current methods of data storage, and hence is also called the *code of life.*
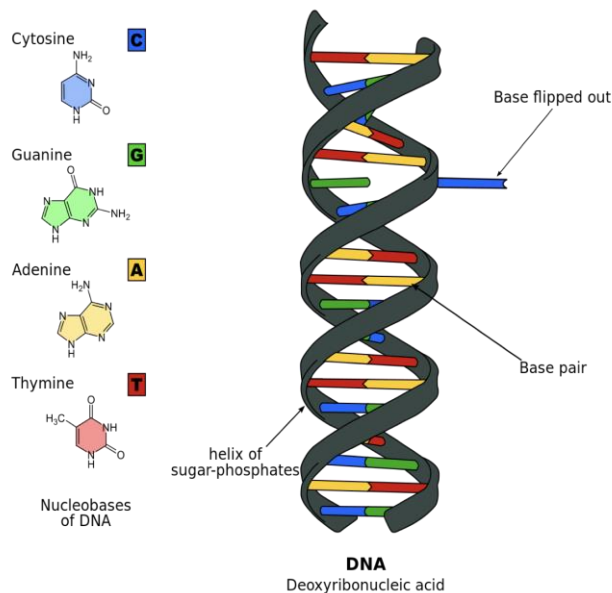

Fig. 1. Molecular structure of DNA [3]

Since we have established a direct relation between the binary form of data we are familiar with and the nucleotides based DNA, we need to understand how we can convert the binary form of data to nucleotide form of data.

There are many ways to establish this, but the basic idea is to assign codes to DNA nucleotides, encoding it in a way. For example, 00 be coded as A, 01 be coded as C, 10 coded as T and 11 coded as G. Now, for example, to store an image file, we encode it into a string of 0s and 1s. Let's say the first 8 bits are as follows 00111001. On breaking them into pairs, in sequence and then encoding as mentioned above, we get A-G-T-C. That's how we need to join the nucleotides to form a single strand of DNA to represent equivalent binary data.

So we can see the nucleotides equivalent to a 4 base system and accordingly encode data as is done by combining 0s and 1s to represent complex forms of data. The important thing post the theoretical conversion

of this binary form of data into corresponding base(s) compatible data is to be able to read and write it accurately. That is where processes such as DNA synthesis and DNA sequencing come in.

DNA synthesis [8] is the natural or artificial process of synthesising strands of DNA. This method is used to create synthetic strands of DNA, and used to write the base(s) form of data obtained from binary data. First the single strands are assembled with the combination of the code to be encoded in the form of A,C,G and T, the sequence of which is obtained from the process as explained above; to form oligonucleotides (short DNA molecules) and later pairing them up correctly, using techniques such as oligonucleotide synthesis and polymerase chain reaction to form DNA strands. Further explanation of which is available here. (provide reference).

DNA sequencing [9] is the process of extracting the sequence of nucleotides from the strands of DNA, and is of prime importance in the process of retrieving the encoded data. There are many methods employed to sequence long strands of DNA with high efficiency and precision. More information on the methods used currently and ones in development can be found here.(provide reference)

Apart from the basic knowledge of DNA, another aspect to look into is of memory hierarchy which is an arrangement of different storage modules based on response time. Based on their response time and costs associated with them, different storage modules serve different purposes in a computer system. Processor registers are on top of this hierarchy yielding fastest response time followed by cache and RAM. The slowest response time are of tape based memories which are used to back up data.

How DNA based storage, if needed, can complement the current system of memory modules is another important topic of discussion based on its access time and costing.

3. Review of Previous Work

The use of DNA in computational systems to solve problems like Hamiltonian graph, was first proposed by L.M. Adleman [5]. Clelland, Risca and Bancroft devised an approach of storing information in DNA, by encoding it into the four bases as stated before [10]. The inspiration for this approach was derived from the second world war where a steganography technique based around a microdot containing a downscaled picture of a letter, was used to communicate secret data. The researchers applied this approach on a DNA scale and hid data in a DNA strand. They synthesised the encoded strand sandwiched between polymerase chain reaction primer sequences and also used an encryption key. The encryption key as showcased in figure 2, was the mapping between the combination of three base(s) and the corresponding numerals and alphabets. This encoded strand was then surrounded by more than hundred times size worth of human DNA molecules meant to conceal the data. This provided the encoded message much required privacy and security. Only, a person with access to the PCR primer sequences and the encryption key would be able to decipher the code.

An interesting and important observation from this study was that no matter the human DNA, if the primer sequences and encryption key were known then the data could be extracted. This research also established the strength of DNA storage as a much more secure and private storage medium when compared to its conventional counterparts in the field of steganography.

Bancroft et. al[11] followed a similar mechanism. They classified DNA into two types for this process, i.e. information DNAs (iDNA) which contained the information and the poly primer key (PPK) (which is key to extracting information stored in iDNAs). Each iDNA strand contained some basic elements, i.e. unique information segment, small common spacer indicating the start of the stored information, unique sequencing primer and common flanking forward and reverse amplification primers. They came up with an experiment which included amplification using polymerase chain reaction to extract data. They encoded and accurately extracted the opening lines of Charles Dickens A Tale of Two Cities. Although, this study did not take into account the response of this storage media in external conditions since it was studied under strict laboratory conditions, this experiment proved the possibility of development of dedicated DNA based storage systems, which was a huge step in paving the way for future work to be done in this field.

The next major contribution[12] to this field was made by Wong et al , based on the problems in the structure of DNA strands that could break at both ends, which could lead to loss of information. They suggested that to prevent DNA from harsh conditions, a dependable medium to store the encoded strands and synthesised gene sequences were required. They employed a vector containing encoded data which is able to

grow and accumulate to ensure longevity of encoded data, for this purpose. They used agents with high rapid regeneration rate and tolerance towards radiation as well as vacuum, such as Escherichia coli and Deinococcus radiodurans. This research was important as it showcased the protection mechanism needed to be employed for the protection of encoded data in this storage mechanism from harsh external conditions. This further proved the competence and capability of DNA based storage devices to be used for archival purposes given its density and longevity.

Up until 2012, the largest project to encode data in DNA was of 7920 bits. Then in [4], George M. Church et al. published a landmark paper highlighting their work on converting an html coded draft of a book that included 53,426 words, 11 JPG images, and one JavaScript program into a 5.27-megabit bitstream into encoded DNA. They were able to recover all data blocks with a total of 10 bit errors out of a total 5.27 million which was remarkably unprecedented. Their method had five advantages over the techniques employed in the past. The most important ones are highlighted here; they used one bit to encode each base instead of the formerly used two which allowed them to encode data in many ways avoiding sequences that might be difficult to read or write. They had split the bit stream into address blocks eliminating the need for long DNA strands that were difficult to assemble at that scale. The advances in DNA synthesis and sequencing technology (which had been a major hindrance for research in this area in the past years) had allowed them to encode and decode data in large amounts in almost hundred thousand times less the cost as compared to first generation encodings.

Nick Goldman and his team at EMBL-European Bioinformatics Institute[13] created a way to encode

"739 kilobytes of hard-disk storage, synthesized DNA, se- quenced it and reconstructed the original files with 100% accuracy. Their data comprised of Shakespeares Sonnets in ASCII format, Watson and Cricks 1954 Classic Paper, a medium resolution colour photograph, a snippet of the audio from Martin Luther King's famous speech in MP3 format and a Huffman code (again in ASCII text scheme) for an aggregate data size of 757,051 bytes."

Their analysis had showcased that DNA based storage systems could be scaled to handle global storage requirements and could become a realistic alternative for digital archiving, (which places the DNA based storage medium at the bottom of the storage hierarchy). This research came out as a major breakthrough for DNA based storage devices. It indicated that if the costs related to DNA synthesis were to reduce at the pace that they were, their scheme could become cost effective within a decade.

Later in [14], researchers at Columbia University by the name of Dina Zielinski and Yaniv Erlich along with New York Genome Centre published a new technique called DNA fountain, using which they were able to store a complete Operating Systems and related files, that had a data storage density of 215 petabytes per gram of DNA which was approximately 85% of the theoretical limit of the Shanon capacity, orders of magnitude higher than previous studies and attempts. Figure 3 summarises the process, employed in this research. The issue with this technique was that it was very costly (around $3500 per MB to synthesise) and not feasible for large scale use.

A problem in information retrieval had been the need to sequence the entire pool of DNA data, even if a small subset was required which was inefficient and was a major contributing factor to the cost associated with DNA based storage. In [15], Microsoft in association with University of Washington demonstrated storage and retrieval of an unprecedented 200 MB of data across 35 distinct files. As depicted in figure 4, they used a large collection of primers to enable individual recovery of data. This research tackled the problem of sequential access of data in DNA based storage systems by showcasing random access and proved the viability of large scale DNA based storage systems for storage and retrieval.

4. Issues and Challenges of DNA based storage

Considering the work done up until now, it is very probable that DNA based storage systems will be the dominant technology around the world one day. But there are some issues and challenges that need to be handled in order for that to happen. The biggest challenge for DNA based storage remains the cost associated with the synthesis and sequencing of information on scale. It is imperative for this cost to be made much more affordable in order for this technology to be adopted by industrial ventures in the near future. With the boom in data creation, we need such affordability to come as soon as possible. Large strides have been taken in the last few years and if the growth continues at the same rate, then there is no reason why this technology cant become viable. Another challenge is that of the total time required in the process of storing data in DNA

based systems, which is many magnitudes more, as compared to its conventional counterparts. Even if it is placed in conjunction with the current memory hierarchy and placed at the bottom of the pyramid, still the write and read time needs to improve many folds for this technology to be both scalable and practical. The machines used today in the process of storing data in DNA are throwing very less error rates but even these error rates are significant when we consider the complications related to the process of storing data and therefore the machines need to be made much more efficient and accurate in order to achieve better rates of successful conversions.

5. Conclusion and Future Scope

We are past the days where DNA based storage was considered a part of people's imagination. DNA based storage devices are here to stay and like any other revolution in the past, it also has to face some big challenges in order to be adopted by the world. The benefits this technology provides such as stability, storage density, energy efficiency and robustness are too significant to ignore, and it is for sure that this technology will be used for archival purposes. In conjunction with present tape technologies. Expecting major leaps in development of this technology won't be asking too much looking at its figures in the past where it has given over a million-fold improvement in recent years. Against the 1.5 times a year progress in electronic technology, this technology has grown 10 times every year since we started reading and writing. A major contribution to this progress goes to the improvement in cost efficiency. We have recorded five- and twelve-times reduction in the cost of synthesis and sequencing respectively. The migration to DNA based storage would also lead to growth in the DNA computing field. This would further speed up the process of reducing synthesis and sequencing costs.

References

[1]. Data Never Sleeps 5.0. "https://www.domo.com/learn/data-never-sleeps-5?aid=ogsm072517 1&sf100871281=1" (Accessed on September 27 2018)

[2]. Next Chapter. How DNA Storage will transform data centre design be- yond recognition. https://data-economy.com/next-chapter-dna-storage-will-transform-data-centre-design-beyond-recognition/ (Accessed on September 18 2018)

[3]. What is DNA? https://ghr.nlm.nih.gov/primer/basics/dna (Accessed on September 18 2018)

[4]. George M. Church, Yuan Gao, Sriram Kosuri, Next-Generation Digital Information Storage in DNA (2012, September) (10.1126/science.1226355)

[5]. L. M. Adleman, Molecular computation of solutions to combinatorial problems. Science 266, 1021 (1994). (10.1126/science.7973651)

[6]. J. Bonnet et al., Chain and conformation stability of solid-state DNA: Implications for room temperature storage. Nucleic Acids Res. 38, 1531 (2010). (10.1093/nar/gkp1060)

[7]. S. Pbo et al., Genetic analyses from ancient DNA. Annu. Rev. Genet. 38, 645 (2004). (10.1146/annurev.genet.37.110801.143214)

[8]. DNA Synthesis. http://www2.csudh.edu/nsturm/CHEMXL153/ DNASynthesis.htm/ (Accessed on September 21 2018)

[9]. DNA Sequencing. https://www.genome.gov/10001177/dna-sequencing-fact-sheet/ (Accessed on September 23 2018)

[10]. Taylor Clelland, Catherine & Risca, Viviana & Bancroft, Carter. (1999). Hiding messages in DNA microdots. Nature. 399. 533-4. (10.1038/21092)

[11]. C.Bancroft, T.Bowler, B.Bloom & C.T.Clelland, Long term storage of information in DNA, Science, vol. 293, no. 5536, pp. 1763-1765, 2001. (10.1126/science.293.5536.1763c)

[12]. Chung Wong, Pak & Wong, Kwong-Kwok & Foote, Harlan. (2003). Organic Data Memory Using the DNA Approach. Commun. ACM. 46. 95-98. (10.1145/602421.602426.)

[13]. N.Goldman, P.Bertone, S.Chen, C.Dessimoz, E.M.LeProust, B.Sipos, E. Birney, Towards practical, high-capacity, low-maintenance information storage in synthesized DNA, Nature 494:7780 (10.1038/nature11875)

[14]. Yaniv Erlich, Dina Zielinski, DNA Fountain enables a robust and efficient storage architecture, Science, 355:950-954 (10.1126/science.aaj2038)

[15]. L. Organick, et al., Random access in large-scale DNA data storage Nature Biotechnology, Nature Biotechnology, 36:242-248 (10.1038/nbt.4079)

Author's Biography

Deepak Kumar Sharma received the B.Tech. in Computer Science Engineering from G. G. S. Indraprastaha University, M. E. in Computer Technology Applications, and Ph.D in Computer Engineering from University of Delhi, India. He is presently working as Assistant Professor in the Division of Information Technology, Netaji Subhas University of Technology (formerly Netaji Subhas Institute of Technology), New Delhi, India. His research interests include opportunistic net- works, wireless ad hoc networks, sensor networks and network security.

Shiv Kumar is an undergraduate student at Netaji Subhas University of Technology (formerly Netaji Subhas Institute of Technology), New Delhi, India. His research interests include big data, data science and machine learning.

Amit Kumar is an undergraduate student at Netaji Subhas University of Technology (formerly Netaji Subhas Institute of Technology), New Delhi, India. His research interests include big data, data science and machine learning.

How to Cite

A Machine Learning Approach for Speech Detection in Modern Wireless Communication Environment

[a]Shibanee Dash[*], [b]Mihir Narayan Mohanty

[a]R.V.R & J.C College of Engineering(Autonomous), Andhra Pradesh

shibaneedash.3@gmail.com , https://orcid.org/0000-0002-3828-1325

[b] ITER, Siksha 'O' Anusandhan(Deemed to be University),Bhubaneswar

mihir.n.mohanty@gmail.com , https://orcid.org/0000-0003-1252-949X

Abstract

Modern wireless communication has gained a improved position as compared to previous time. Similarly, speech communication is the major focus area of research in respective applications. Many developments are done in this field. In this work, we have chosen the OFDM modulation based communication system, as it has importance in both licensed and unlicensed wireless communication platform. The voice signal is passed though the proposed model to obtain at the receiver end. Due to different circumstances, the signal may be corrupted partially at the user end. Authors try to achieve a better signal for reception using a neural network model of RBFN. The parameters are chosen for the RBFN model, as energy, ZCR, ACF, and fundamental frequency of the speech signal. In one part these parameters have eligibility to eliminate noise partially, where as in other part the RBFN model with these parameters proves its efficacy for both noisy speech signals with noisy channel as Gaussian channel. The efficiency of OFDM model is verified in terms of symbol error rate and the transmitted speech signal is evaluated in term of SNR that shows the reduction of noise. For visual inspection, a sample of signal, noisy signal and received signal is also shown. The experiment is performed with 5dB, 10dB, 15dB noise levels. The result proves the performance of RBFN model as the filter. The performance is measured as the listener's voice in each condition. The results show that, at the time of the voice in noise environment, proposed technique improves the intelligibility on speech quality.

1. Introduction

In 70's the Wireless and cellular concept was developed and become more popular than expected at the time. Since then the focus in this area increased till now. The clarity at the receiver end is highly essential. Day-by-day the communication techniques are developed along with the type of modulation. Researchers focus on the capacity as well as the reception capability [1]. The technology including OFDM,

---

[*] Corresponding author

Shibanee Dash

Email: shibaneedash.3@gmail.com

MIMO and coding types are to be observed for higher bandwidth utilization with better response in less time. Congestion in licensed spectrum increases in gradual manner for which unlicensed spectrum utilization has been occurred that is developed by researchers [2-3].

Many different factors are included in wireless communication, such as environmental factors, channel condition, types of speakers, and the way we transmit in the channel. Research of speech technology requires changing the way of transmission, and reception in specific type of communication [4].

The Speech is a common mode of Communication between of human being. Speech signal is an important data for communication network. The detection and recognition accurately is most important criteria. To maintain wireless communication successfully speech of human being should listen and speak clearly. As well the matching among network should be maintained perfectly. In such cases the synthetic speech is also verified in many problems [5-7] .

Speech has potential of being important mode of interaction with computer for this evaluation. For errorless communication and noiseless signal reception, initially speech signal is processed for noise suppression and enhancement. It helps in choosing the technique in digital age point of view [8-9].

It is a challenge for the next technological development to make the natural speech reception through HCI at the user end. Speech processing is exciting areas of research in signal processing and one type of pattern recognition problem. The choice and use of features should relevant for the purpose of detection. Again it must well manage at the time of training and testing on use of machine learning techniques.

Different Technologies are used for faster communication in both licensed and unlicensed spectrum utilization. These are

(i) OFDM is a FDM scheme where digital multicarrier modulation method is utilized. The data is subdivided into parallel data streams sharing every sub-carrier. Further every sub-carrier is modulated with a modulation method like QAM, and variant of PSK.

(ii) MIMO technique is used with multiple antennas along with transmitters and receiver to improve performance. It offers increase in data and receivers throughput and link range without additional bandwidth.

(iii) Turbo Code-It is a category of high performance error correction codes which was developed in 1993.The coding technique is one of the significant technologies in communication. It is helpful for maximum information transmission without error/ relatively small error.

The respective generations of wireless communication have many advantages like higher bandwidth, Better response time. It works at 2.6 GHz frequency implies that better coverage even though with same tower. The gradual developed generations provide higher flexibility as compared to already existing technologies [10].New technology with less cost and better usage needs to simplify hardware with effective design so that the versatility can be maintained with the same handset with better reception capability for different generations.

The paper is organized as follows. Section 1 introduces the work. Section 2 provides the methodology proposed in this work. It explains the principle of hearing the speech communication model. Through the model the speech is communicated. For detection purpose Radial Basis Function Network (RBFN) is used and explained with its parameters as the problem formulation. Section 3 explains the result and section 4 concludes the work.

2. Methodology

One important mechanism for received signal is source separation that has the capability to remove time-frequency regions where the speech signal is less distorted [11]. To increase the success of communication, adaptation is required at different levels, such as subject, place and vocabulary. As a result the Lombard effect [5] can be analyzed. The parameters for Lombard speech like intensity, vowel duration, speaking rate, energy distribution, spectral tilt, formant frequency are observed by researchers.

OFDM System

OFDM technique is used due to better transmission capacity and high bandwidth efficiency in wireless communication for both licensed and non-licensed spectrum. Such system is based on spreading technique with low rate carriers. The spacing between the orthogonal components is generated using the Fast

Fourier Transform technique [12]. The data is converted to parallel stream and grouped. Further, it is modulated using either Quadrature Amplitude Modulation (QAM), or Quadrature Phase Shift Keying (QPSK), or Binary Phase Shift Keying (BPSK). Finally, required spectrum is then converted back to its time domain signal using an Inverse Fast Fourier Transform (IFFT). At the receiver end it is converted from parallel to serial for transmission of data. With this technology the system is designed by considering the Gaussian noise channel [13].

The basic model of OFDM system is presented in Fig. 1. Input signal as considered for transmitted symbols through the wireless Gaussian channel.
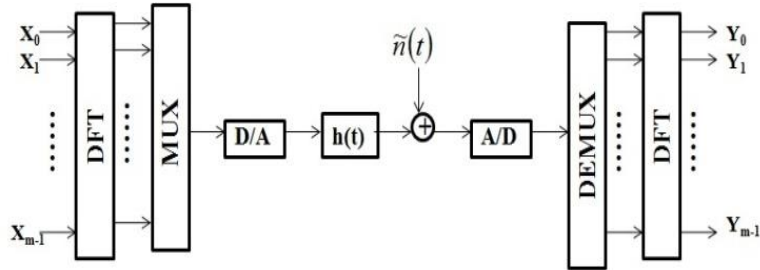

Fig. 1. Basic OFDM system

The impulse response of the channel can be expressed as,

$$h(t) = \sum_m a_m \delta(t - \tau_m T_s) \tag{1}$$

Where α_m represents the amplitudes. It is formed by a N-point DFT_N and is expressed as,

$$y = DFT_N(IDFT_N(x) \otimes \frac{h}{\sqrt{n}} + n) \tag{2}$$

For N-independent channel the expression will be, $y_i = h_i X_i + n_i, i = 0........N-1$

Where $h = (h_0, h_1........h_{n-1})^T$ that can be considered as attenuation of the channel and $n = (n_0, n_1........n_{n-1})^T$ is a noise vector. The system can be formulated as,

$$y = XFg + n \tag{3}$$

Where X is the input data and can be expressed in terms of twiddle factor as,

$$F = \begin{pmatrix} W_N^{00} & \cdots & W_N^{0(N-1)} \\ \vdots & \ddots & \vdots \\ W_N^{(N-1)0} & \cdots & W_N^{(N-1)(N-1)} \end{pmatrix} \tag{4}$$

The twiddle factor is defined as,

$$W_N^{nk} = 1/\sqrt{N} e^{-j2\pi nk/N} \tag{5}$$

The MMSE estimate of h becomes,

$$\hat{h}_{MMSE} = R_{hy} R_{yy}^{-1} y \tag{6}$$

Where,

$$R_{hy} = E\{hy^H\} = R_{hh} F^H X^H$$

$$R_{yy} = E\{yy^H\} = XFR_{hh}FHX^H + \sigma_n^2 I_N$$

signifies the cross covariance matrix and the auto covariance. Again R_{hh} is the auto covariance matrix of h and σ_n^2 denotes the noise variance. Assuming these quantities to be known, the MMSE estimates (h_{MMSE}) will be,

$$\hat{h}_{MMSE} = F\hat{h}_{MMSE} = FQ_{MMSE}F^H X^H y \tag{7}$$

Where Q_{MMSE},

$$Q_{MMSE} = R_{hh}[(F^H X^H XF)^{-1}\sigma_n^2 + R_{hh}]^{-1}(F^H X^H XF)^{-1} \tag{8}$$

The LS estimator for channel impulse response h is analyzed as follows,

Similarly the least square channel estimator can be formulated as,

$$\hat{h}_{LS} = FQ_{LS}F^H X^H y \tag{9}$$

where,

$$Q_{LS} = (F^H X^H XF)^{-1} \tag{10}$$

considering the two equations we have,

$$\hat{h}_{LS} = X^{-1}y \tag{11}$$

From equation (7) and (11) it is shown the LS estimate has a high mean square error as compared to MMSE estimation technique.

Parameters for Speech Detection

Energy

It is defined as the squared signal. In speech signal case it is analyzed frame wise. Hence the short time energy is to be evaluated considering different windowed signal [14]. The energy of the speech signal reflects the amplitude variations. Short-time energy can define as:

$$Energy = \sum_{i=-\infty}^{\infty}[s(i)w(m-w)]^2 \tag{12}$$

where , $s(i)$ represents the signal, $w(m)$ represents the window and En represents the energy.

Zero-Crossing Rate

The rate of change of signal from positive to negative is defined as the Zero crossing Rate (ZCR). It is a measure of number of times in particular time interval/frame. As a result the amplitude of the speech signals passes through a value of zero. The zero crossing rate of a signal can be found by using

$$ZCR = \sum_{-\infty}^{\infty}|sgn[s(i)]sgn(i-1)|w(n-w) \tag{13}$$

where, $sgn[s(i)] = 1$ if $s(i) \geq 0 = $ -1 if $s(i) < 0$

and $w(n) = \dfrac{1}{2N}$, $0 < n < N-1$ (N is the length of signal) $= 0$, otherwise.

The model for speech production suggests that the energy of voiced speech is concentrated about 3

kHz as the spectrum fall of glottal wave and for unvoiced speech, the energy is found at higher frequencies. Since high frequencies imply high zero crossing rates, and energy. There is a strong correlation between zero-crossing rate and energy distribution. Therefore, another parameter is considered as autocorrelation co-efficient to keep relevancy at received signal.

Autocorrelation

For clean speech and separation of noise autocorrelation coefficient has a major role alike to energy and ZCR. It works not only for noise elimination, but also for smoothening the signal. It is a type of cross correlation and convolution. The relation is expressed as,

$$R_{xx}(j) = \sum_n x_n \bar{x}_{n-j} \tag{14}$$

It is of finite energy of Signal. Similarly for the measurement of frequency as low or high fundamental frequency is an important parameter of speech and is considered.

Fundamental frequency.

As the human voice varies over a range of frequencies, the fundamental frequencies cannot be considered as a specific value. Though it is an essential component of speech and speaker recognition, it has a similar application in voice communication and is taken as an attribute.

Radial basis function Network (RBFN) Model for Detection

Both RBFN is used and tested for detection accuracy. Different possible hybridization of features has been attempted. Radial basis function Network (RBFN) consists of an input layer, a hidden layer and a linear output layer. In this case, the Gaussian kernel as activation function is used and the distance is evaluated [15]. The hidden layer depends on a non-linear RBF activation function [16-17]. The output of the network is found as the distance between the input vector and the vector of the centre of the Gaussian function and can be expressed as [18-20].

$$
\begin{bmatrix} y1 \\ y2 \\ \vdots \\ y_j \end{bmatrix} =
\begin{bmatrix}
R\|x_1 - c_1\| & R\|x_1 - c_2\| & \cdots & R\|x_1 - c_j\| \\
R\|x_2 - c_1\| & R\|x_2 - c_2\| & \cdots & R\|x_2 - c_j\| \\
\vdots & \vdots & \vdots & \vdots \\
R\|x_j - c_1\| & R\|x_j - c_2\| & \cdots & R\|x_j - c_j\|
\end{bmatrix}
\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_j \end{bmatrix} \tag{15}
$$

where, R is the RBF, c_j is the center, $\|x - c_j\|$ is the distance between input and the center. $x_1, x_2, ...x_j$ are represented as the inputs, $y_1, y_2....y_j$ are the outputs and $w_1, w_2,...w_j$ are the weights of the network. The target output is obtained by updating the corresponding weights. The output to weight and input is given as,

$$y = \sum_{j=1}^{N} R(\|x - c_j\|) w_j \tag{16}$$

where, w_j is the weight of the jth center and N is the length of the signal. The structure of the network is shown in Figure1. The network is operated with the activation function that is Gaussian and is expressed as,

$$R(\|x - c_j\|) = \exp\left[-\frac{(x - c_j)^2}{2\sigma^2}\right] \tag{17}$$

where, σ is the width of the center. The network is trained with adaptive learning method and is described in following subsection as proposed method.
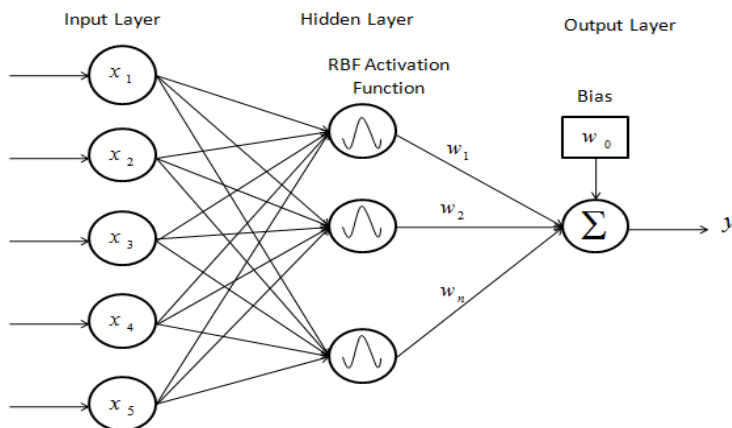

Fig 2. Structure of RBF Network

3. Result

The work consists of two parts as modulation technique and noise elimination through neural network model. The results are obtained from both the techniques and depicted in this section. The bit error rate is obtained to validate the OFDM system and is shown in Fig. 3. To strength it the MSE is found and is shown in Fig. 4.

Once the system found suitable, the chosen parameters are given to the RBFN model. One of the sample of speech is shown in the Fig.5 for visual aid. The corresponding outputs for original signal, noisy signal and obtained result with noiseless signal are shown in Fig. 6 and Fig. 7. From this result it is clear that the voice signal is well communicated and can be suitable for next generation wireless network.
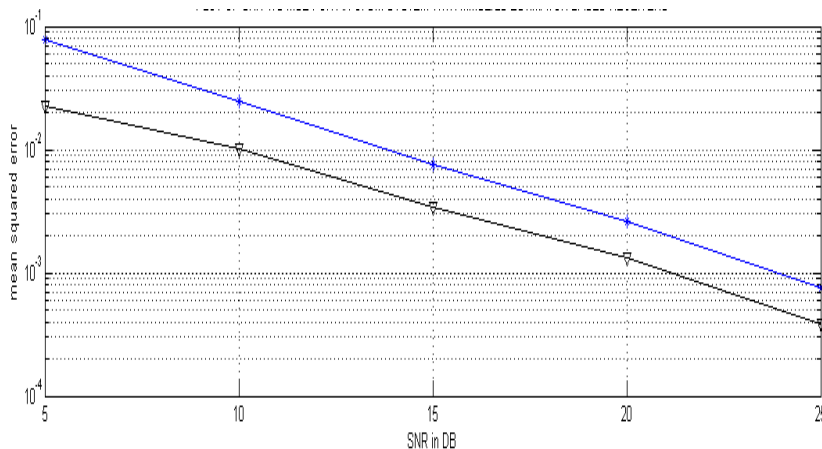

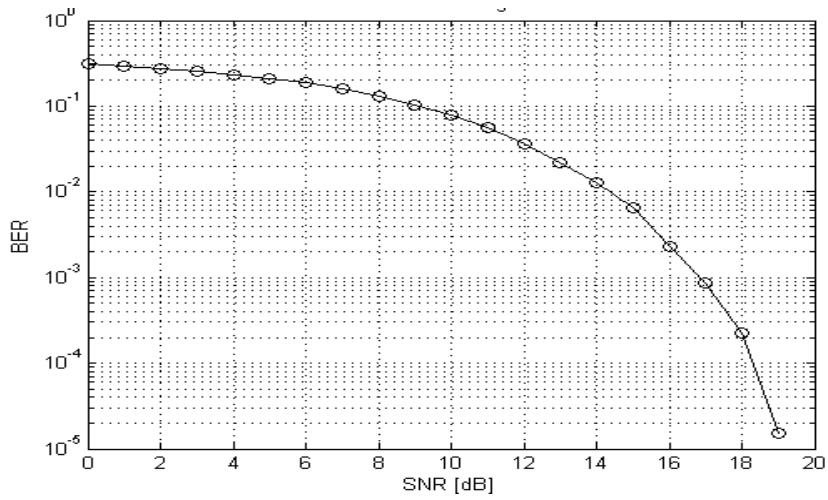
Fig 3. Plot of SNR vs. MSE for OFDM system
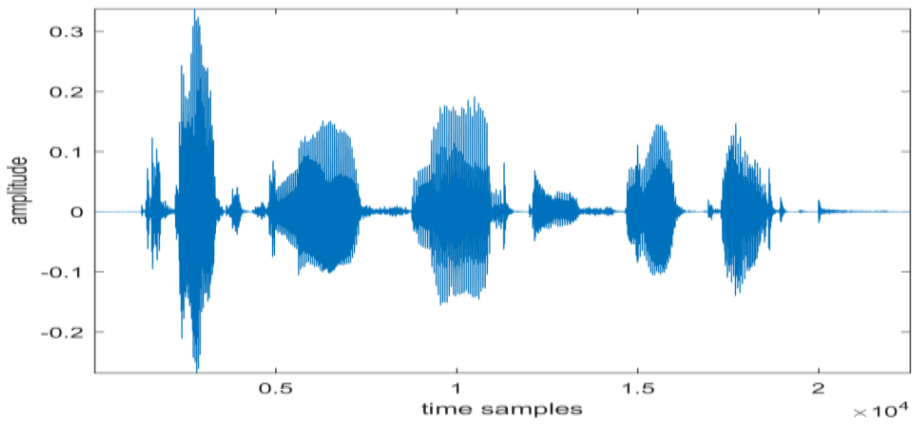
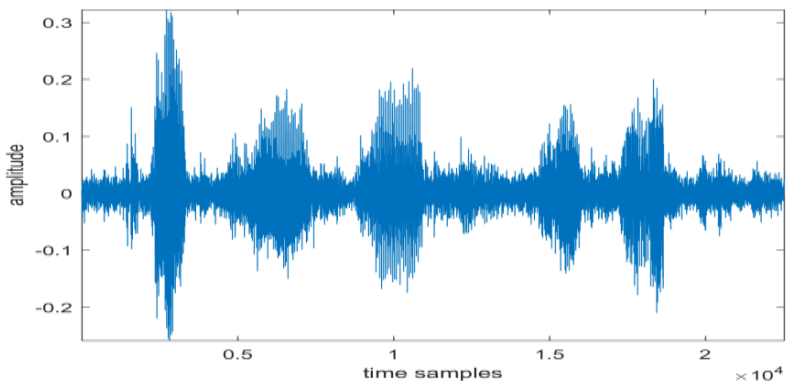Fig 4. BER vs. SNR of OFDM system

Fig. 5.Original Speech signal

Fig. 6.Noisy Speech signal with SNR of 5dB
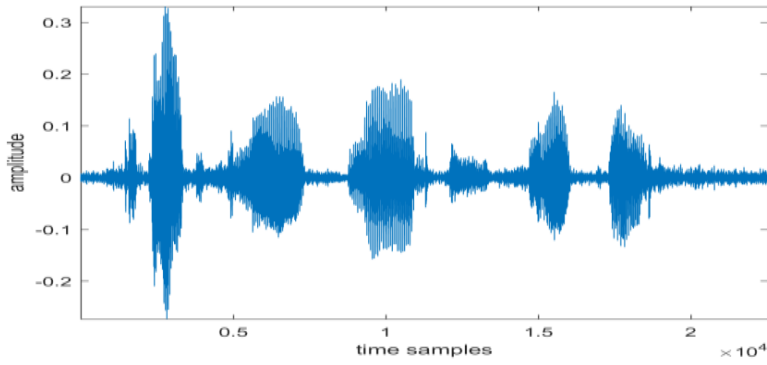
Fig.7: The Enhanced Signal at the Receiver End

Table 1: Accuracy performance of RBFN with different combinational features

Features	Average Accuracy	MSE
	80.44%	0.706
$Z + ACF$	80.53%	0.683
$E_{sT} + F0$	75.03%	0.813
$F0 + Z$	75.40%	0.808
$E_{sT} + ACF$	80.55%	0.680
$F0 + ACF$	75.33%	0.810
$E_{sT} + Z + ACF$	85.82%	0.436
$E_{sT} + F0 + ACF$	84.16%	0.487
$ACF + Z + F0$	83.14%	0.535
$E_{sT} + Z + F0$	83.18%	0.514
$E_{sT} + F0 + Z + ACF$	**90.38%**	**0.352**
$E_{sT} = STE$, $F0 =$ Fundamental Frequency, $ACF =$ Autocorrelation Coefficient, $Z =$ Zero Crossing Rate		

4. Conclusions

From the work, it is concluded that the detection accuracy depends largely on the types and size of the features fed as input. Graphical analysis shows that Intensity or energy appears to be the best feature for detection. Recognition of emotional speech in communication can provide a new future direction.

References

[1].   Haykin, S. S. (2011). Modern wireless communications. Pearson Education India.

[2].   Mohanty, M. N., Mishra, L. P., & Mohanty, S. K. (2011). Design of MIMO space-time code for high data rate wireless communication. International Journal on Computer Science and Engineering, 3(2), 693-696.

[3].   Dash, S., & Mohanty, M. N (2018). Voice Detection for Cognitive Radio Receiver in Cooperative Spectrum Sensing Environment, AESPC (Accepted).

[4].   Tse, D., & Viswanath, P. (2005). Fundamentals of wireless communication. Cambridge university press.

[5].   Junqua, J. C. (1993). The Lombard reflex and its role on human listeners and automatic speech recognizers. The Journal of the Acoustical Society of America, 93(1), 510-524, doi.org/10.1121/1.405631.

[6].   Loizou, P. C. (2007). Speech enhancement: theory and practice. CRC press

[7].   Summers, W. V., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. The Journal of the Acoustical Society of America, 84(3), 917-928, doi.org/10.1121/1.396660.

[8].   Mowlaee, P., Stahl, J., & Kulmer, J. (2017). Iterative joint MAP single-channel speech enhancement given non-uniform phase prior. Speech Communication, 86, 85-96, doi.org/10.1016/j.specom.2016.11.008.

[9].   Rabiner, L. R., & Schafer, R. W. (1978). Digital processing of speech signals (Vol. 100, p. 17). Englewood Cliffs, NJ: Prentice-hall.

[10]. Mohanty, M. N., & Mishra, S. (2013, March). Design of MCM based wireless system using wavelet packet network & its PAPR analysis. In Circuits, Power and Computing Technologies (ICCPCT), 2013 International Conference on(pp. 821-824). IEEE, doi/10.1109/ICCPCT.2013.6528867.

[11]. Cooke, M. (2003). Glimpsing speech. Journal of Phonetics, 31:579 – 584

[12]. Mishra, D., Mishra, S., & Mohanty, M. N. (2011). Estimation of MIMO-OFDM Based Channel for High Data Rate Wireless Communication. IJCSIT) International Journal of Computer Science and Information Technologies, 2(3), 1263-1266.

[13]. Mishra, B., Mishra, S., & Mohanty, M. N. (2012). Design of Wavelet Packet Based Model for Multi Carrier Modulation. International Journal of Engineering Science and Technology, 4(04), 1572-1575.

[14]. Mowlaee, P., Stahl, J., & Kulmer, J. (2017). Iterative joint MAP single-channel speech enhancement given non-uniform phase prior. Speech Communication, 86, 85-96, doi.org/10.1016/j.specom.2016.11.008.

[15]. Haykin, S. S. (2009). Neural networks and learning machines (Vol. 3). Upper Saddle River: Pearson.

[16]. Phooi, S. K., & Ang, L. M. (2006, November). Adaptive RBF neural network training algorithm for nonlinear and nonstationary signal. In Computational Intelligence and Security, 2006 International Conference on (Vol. 1, pp. 433-436). IEEE, doi.org/10.1016/j.specom.2016.11.008.

[17]. Palo, H. K., Mohanty, M. N., & Chandra, M. (2015). Design of neural network model for emotional speech recognition. In Artificial intelligence and evolutionary algorithms in engineering systems (pp. 291-300). Springer, New Delhi, doi.org/10.1007/978-81-322-2135-7_32

[18]. Cheng, J. C., Su, T. J., Li, T. Y., & Wu, C. H. (2015, September). The Noise Reduction of Speech Signals Based on RBFN. In Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2015 International Conference on (pp. 449-452). IEEE, doi.org/10.1109/IIH-MSP.2015.111.

[19]. Mohapatra, S. K., Palo, H. K., & Mohanty, M. N. (2017). Detection of Arrhythmia using Neural Network. Annals of Computer Science and Information Systems, 14, 97-100, doi.org/10.15439/2018KM42.

[20]. Singer, E., & Lippman, R. P. (1992, March). A speech recognizer using radial basis function neural networks in an HMM framework. In Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on (Vol. 1, pp. 629-632). IEEE, doi/10.1109/ICASSP.1992.225830.

Author's Biography

Shibanee Dash is presently working as a Assistant Professor in the Department of Electronics and Communication Engineering, at R.V.R & J.C College of Engineering (Autonomous), Guntur, Andhra Pradesh, India. She has Master of Technology in Electronics and Telecommunication at Kalinga Institute of Industrial Technology (Deemed to be University), India. She has 3 years of experience in teaching and research.

Mihir Narayan Mohanty is presently working as a Professor in the Department of Electronics and Communication Engineering, Institute of Technical Education and Research. Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India. He has published over 300 papers in International/National Journals and Conferences along with approximately 25 years of teaching experience. He is the active member of many professional societies like IEEE, IET, EMC & EMI Engineers India, ISCA, ACEEE, IAEng, CSI and also Fellow of IETE and IE (I). He has received his M.Tech. Degree in Communication System Engineering from the Sambalpur University, Sambalpur, Odisha and done his Ph.D. work in Applied Signal Processing. His area of research interests includes Applied Signal and image Processing, Digital Signal/Image Processing, Biomedical Signal Processing, Microwave Communication Engineering and Speech Processing.

How to Cite

Study of Concurrency Control Techniques in Distributed DBMS

[a]Manoj K. Gupta[*], [b]Rakesh K. Arora, [c]Bhoopesh Singh Bhati

[a]Professor, Rukmini Devi Institute of Advanced Studies, Delhi, India.

manojkgupta5@gmail.com, http://orcid.org/0000-0002-4481-8432

[b]Asso. Professor, Krishna Engineering College, Ghaziabad, UP, India.

rka1211@gmail.com , https://orcid.org/0000-0002-8601-2816

[c]Asstt. Professor, Ambedkar Institute of Advanced Communication Technologies and Research, Delhi, India.

bhoopesh.cse@gmail.com , https://orcid.org/0000-0001-8476-2798

Abstract

Concurrency control is one of the important task of any database management system. Without the proper concurrency control technique it is infeasible to maintain the integrity of the database system in concurrent environment. The concurrency control algorithms focus on maintaining consistency and integrity of databases through synchronized access. In centralized environment it is simple to synchronize among the various concurrent transactions. But, it becomes very complex as compared to centralized framework when the concurrency control algorithms are implemented for distributed framework because of the requirements of consistency and integrity within the multiple fragments / copies of the database, synchronization among the various distributed concurrent transactions, and isolation of the complexities of the algorithms / operations. A variety of concurrency control methods have been proposed by the many researchers so far for the centralized framework as well as for the distributed framework. The concurrency control methods are broadly classified as locking based methods, timestamp ordering based methods and optimistic methods. Two-phase locking based methods are the most popular and widely used methods used in both centralized and distributed framework of the database systems. A number of algorithms based on two-phase locking have been proposed by the many researchers for Distributed DBMS. This paper consolidates and discusses various lock based concurrency control techniques for Distributed DBMS. This paper also presents a comparative study of various two phase locking based concurrency control techniques.

Keywords

Database,

Distributed Database,

Distributed DBMS,

Concurrency Control Techniques,

Two Phase Locking,

1. Introduction

Database system is the backbone of the many organizations as is stores and manages the operational data of the organization. Most of the applications of the organizations are dependent on database systems. The database systems can be either implemented based on centralized approach or distributed approach. The

---

[*] Corresponding author

Manoj K. Gupta

Email: manojkgupta5@gmail.com

database system which implemented based on distributed approach is called as Distributed Database System. A database system (DBS) is a combination of Database Management System (DBMS) and Databases of the organization. The shift from centralized to distributed architectures over the years attributes to the demand for higher performance and availability. Distributed database system (DDBS) technology in the field of database systems is also the result of the same. This technology may be viewed as combination of database system, computer network technologies and the concept of distributed computing (or distributed processing) [3, 12, 14, 17].

"*A distributed database (DDB) involves a collection of number of logically related databases spread over a computer network*" [9]. Özsu [9] defined *distributed database management system* (DDBMS) as "*the software system concerned with the management of the distributed database and makes the distribution transparent to the users.*"

The combination of DDB and DDBMS is referred to as "distributed database system" (DDBS).

Internally, DBMSs perform several functions in order to manage and manipulate the data properly such as transaction management, concurrency control, recovery, security, etc. The concurrency control ensures the consistency and integrity of the databases in the concurrent execution environment. As we know that the DBMSs support database sharing among various transactions. The uncontrolled concurrent execution of the transactions may lead to the database systems into an inconsistent state. Therefore, there is requirement of controlling the concurrent execution of the transactions so that the consistency and integrity of the database systems can be ensured. The concurrency control is the mechanism through which DBMS can ensure the consistency and integrity of the databases even in the case of concurrent execution of the transactions without affecting the degree of concurrency.

2. Concurrency Control

In DDBS, at the same time, multiple users can access the database concurrently where each user thinks that he/she is working alone on dedicated system but this is not the case. In case of concurrent execution of transactions, the consistency of the database can ensured with the help of serializable schedules because the result obtained will be equivalent to one of the serial execution of the transactions. The use of only serial schedules limits the degree of concurrency hereby affecting performance. Therefore, concurrency control techniques need to be applied to guarantee that the schedules generated by concurrent execution of transactions are serializable [1].

Concurrency control involves the coordination among concurrent accesses to maintain consistency and integrity of database [1, 24]. The major problem in attaining this objective is to make ensure that database updates performed by one the transaction should not affect the updates and retrievals of the another transaction [9, 12, 13, 16, 25]. The problems regarding concurrency control is more difficult in a DDBMS due to the following reasons:

- Data may be accessed by the multiple users at number of distant sites,

- Database is fragmented and/or replicated across multiple sites,

- Complexity during the synchronization among concurrent transactions executed at multiple remote sites by multiple users, and

- Concurrency control techniques implemented at one location must ensure the consistency of the database at all other sites.

3. Concurrency Control Techniques

A number of concurrency control mechanisms have been proposed by the leading researchers so far. These concurrency control mechanisms broadly can be classified in the categories such as (a) Locking Based Protocols, (b) Timestamp-Ordering Based Protocols and (c) Optimistic Protocols. The first two categories are based on the pessimistic approach. The pessimistic protocols are used when there is high activity on the databases whereas the optimistic protocols are used in case of low activity on the databases. Implementation of these concurrency protocols / algorithms for the centralized environment is simpler than that of distributed environment.

The number of concurrency control techniques implemented for centralized environment can easily be extended to handle the problem in distributed databases, but there are some concurrency control techniques that are not suitable for a distributed environment [1, 16, 25]. The detailed classification of concurrency control techniques is as follows:

3.1. Pessimistic Techniques

These techniques involve the synchronization among the transactions during initial phase of their execution life cycle as these consider the high conflicts among the concurrent transactions. The pessimistic techniques are mainly suitable when there is high activity in the database system. In other words, if a large number of conflicting transactions are executed concurrently on the database systems frequently then it is preferred to use pessimistic techniques which reduce the wastage of resources if conflicts identified at later stage. In addition, if particular application involves activity that is beyond the rolling of database capability like printing, than such activity do not take place whenever conflict is generated, hereby removing the responsibility of performing undo operation. The pessimistic techniques further can be classified as:

3.1.1. Two Phase Locking (2 PL) Based techniques

Two phase locking technique is based on the locking technique which ensures the serializability of the concurrent transactions in order to maintain the consistency and integrity of the database system. In this technique, each data object of the database is associated with a shared variable called as lock which stores the state of the data object in order to control the shared access of the data object by the mutually exclusive transactions [1, 9, 19, 21, 23]. In two-phase locking, there are two phase of operations. The first phase is called as growing phase in which the transactions can acquire or upgrade the locks and in the second phase the transactions can release or degrade the locks only. To implement two phase locking the following rules need to be followed (a) Conflicting locks should not exist in two transactions. (b) Unlock operation cannot be performed before lock operation in any of the transaction. (c) Until and unless all locks are obtained, no data are affected in any of the transaction. This approach may results in deadlocks and starvation.

3.1.2. Timestamp-Ordering Based techniques

In this technique, the transactions are ordered based on some value called as timestamp. The timestamp can be either a value generated by a global incremental counter variable or current timestamp of the master clock which ensures the uniqueness of the timestamp. The transaction which is created earlier is assigned a lower timestamp as compared to the transaction which is created later. The transaction with lesser timestamp is older than the transaction with higher timestamp. Therefore, the ordering of transaction can be done based on the assigned timestamp and hence called as timestamp-based ordering technique which avoids the deadlock situation [1, 19, 20, 22, 23]. In this technique, each data item X is stored as a triplet X = {x, *write-timestamp, read-timestamp*} where x is the latest value of X, *write-timestamp* is the timestamp of the youngest transaction who has written the latest value of *X*, and *read-timestamp* is the timestamp of the youngest transaction who has read the value of *X*. There are two-approaches based on this technique (a) wait-die and (b) wound-wait. This technique has the advantage of eliminating deadlocks as the transactions need not to wait. This technique results in cascading rollbacks. Starvation can also occur if same transaction is aborted and restarted again and again.

3.1.3. Integrated

Both two-phase locking technique and timestamp-based ordering techniques are having their relative pros and cons. In order to exploit the pros of both the techniques, some DDBMS uses the integrated approach in which combination of the above two techniques are implemented at the same time. This approach proves to be advantageous when heterogeneous databases are connected together [1].

3.2. Optimistic techniques

In these techniques, synchronization of concurrent execution of transactions is delayed until their termination [19, 20, 21]. These techniques allow the transactions to perform their operations as they desire except write-phase in order to maximize the degree of concurrency. In this approach, the operations of the transactions are logically divided in to three phases (a) read phase, (b) validation phase, and (c) write phase. In the read phase, all the transactions are freely allowed to perform their operations in the local memory. Before writing the changes to the database the transaction has to pass the validation phase in which DDBMS

check whether the proposed write phase of the transaction will lead the database in to consistent state or not. If validation phase indicates the consistency of the database then the write-phase will be performed by the transaction otherwise the write-phase of the transaction is denied. Optimistic approach can be implemented either using 2PL or timestamp-based ordering techniques. No chance of cascading rollback is there because the actual write operation occurs only when the transaction initiating the write operation has committed.

The comparison between pessimistic and optimistic approach is presented in Table 1.

Table 1: Comparison between Pessimistic Approach and Optimistic Approach

Basis	Pessimistic Techniques	Optimistic Techniques
Synchronization	The synchronization of transactions occurs during initial phase of their execution life cycle.	The synchronization of concurrent execution of transactions is delayed till their termination.
Conflict	Large number of transactions will conflict with each other.	Relatively less number of transactions is in conflict with each other.
Phases of Transaction Execution	Validate, Read, Compute, Write	Read, Compute, Validate, Write
Strengths	• Simple • Suitable where transactions conflicts are more. • Lower Storage Cost	• Submitted operations never delayed. • Higher degree of concurrency
Limitations	• Lower degree of concurrency • May lead to deadlocks	• Suitable for applications where transactions conflicts are very rare. • Wastage of Resources (in case of failure of Validation phase) • Higher Storage Cost for intermediate results • Originally based only on timestamp ordering. • Concentrates on centralized DBMS.

4. Two Phase Locking Based Techniques

Locking-based techniques are widely used in centralized database systems in which logical and physical locks on data items are used for synchronization among transactions. If any transaction wants to read and/or write any granule of database, it has to first acquire the lock on that granule and after the completion of operation the locks has to be released. The locking and unlocking on data items is handled by lock manager on behalf of transactions. There are four major categories for implementing lock based algorithms:

4.1. Centralized two phase locking algorithm

It is a variation of 2 PL in which a single lock table is maintained at any one designated site for the whole distributed database. All the requests for locking and unlocking are sent to that site only [1, 26]. Only the designated site decides the grant of requisite lock to the requested transactions based on the compatability of requests. This technique is useful in the case of both replicated and fragmented distributed databases. The load on the single site will increase as all the rquests for locking and unlocking are made at single site only. If the designated site fails, then the operation of the DDBMS will fail.

4.2. Primary Copy two phase locking algorithm

This approach is useful when multiple copies of same data item are stored at different locations (replication). One of the copies at any particular site is designated as a primary copy. All the requests for locking and unlocking are sent to the designated site only with respect to the specific data items. [1, 14, 15, 18, 20]. This approach is suitable for replicated or mixed databases. Effect of the failure of one site is lesser than that of centralised two phase locking as each site is designated as primary copy for specific data items only.

4.3. Distributed two phase locking algorithm

Locking of data items are done at all sites where transaction accesses these data items. Once locking of all data items have been completed by the transaction, only than unlocking begins [24]. The trasactions are required to submit the locking and unlocking requests to all sites. The lock request will be granted to the transactions when it is allowed by all the sites. This approach is more complex then centralized two phase locking and primary copy two phase locking. In this approach, the network traffic will increase. Lot of communication overhead is involved and handling of deadlocks becomes more complex. This technique is more reliable as failure of centralized site will affect less number of transactions.

4.4. Majority Consensus 2 PL Algorithm

This technique is the variation of distributed two phase locking algorithm. Unlike, distribute two phase locking, it requires locking on majority of copies, i.e. at least (n+1) / 2 copies, instead of all copies. If the data item is to be updated, the transaction would have to send updated value to all sites where data item is stored [1, 18]. Less communication ovehead as compared to distributed two phase locking technique as less number of requests have to be made for locking and unlocking data items.This algorithm is more efficient than the distributed two phase locking algorithm.

The comparison among above two-phase locking based techniques is presented in Table 2.

Table 2: Comparison among Two Phase Locking (2 PL) Based Techniques

Basis	Centralized 2PL	Primary Copy 2 PL	Majority Consensus 2 PL	Distributed 2 PL
Control	The sites designated as primary site has the responsibility of granting locks on all data items to the transactions	For each data item , one of the site (in case of replication) is designated as the primary copy that has to be locked by the transactions	Each site has the responsibility of granting locks on its all local lock units to the transactions. A majority of copies are required to be locked instead of all copies	Each site has the responsibility of granting locks on its all local data items to the transactions
Lock Management	Centralized (only at one site)	Mixed (Centralized for each primary copy and Distributed for several primary copies)	Distributed (at each site)	Distributed Lock Management (at each site)
Number of Lock Managers	Only one (at primary site)	Equals to number of distinct lock units (for each primary copy)	Equivalent to the number of sites	Equivalent to the number of sites
Suitable for Data Distribution	Fragmented, Replicated or Mixed	Replicated or Mixed	Replicated or Mixed	Fragmented, Replicated or Mixed
Strengths	• Easy to implement	• Improved	• Number of messages	• Increased

		performance	for locking and unlocking is less	performance for local transactions
	• Less costly	• Reduces the load on central site	• Number of deadlocks is relatively less.	• Reduces the load of central site
		• Communication cost is less		
Limitations	• Less reliable as the failure of the site would result in major system failures [9]	• High Complexity	• High Complexity	• High Complexity
	• Performance degraded at high loads due increase in amount of work.	• Demands a more sophisticated directory at each site	• Large communication overhead as compared to Centralized 2PL	• Large communication overhead
	• Bottleneck as all requests for locking and unlocking are made at same site.			• May cause the distributed deadlock
				• Demands a more sophisticated directory at each site

5. Conclusion

Several categories of concurrency control methods have been proposed and implemented for Distributed DBMS. The pessimistic methods are still widely used but are cause of concern as they may lead to number of problems like lowering the degree of concurrency, deadlocks, starvation and cascading rollbacks. The optimistic methods can be used to increase the degree of concurrency but the current works on optimistic methods concentrates mainly on centralized DBMS in contrast to distributed DBMS. The pessimistic approach is mostly suitable for the distributed database systems with high activity ratio whereas optimistic approach is mostly suitable for the distributed database systems with low activity ratio.

References

[1]. Bernstein, P. A. and Goodman N., "Concurrency Control in Distributed Database Systems". ACM Computing Surveys, 13(2), June 1981 (DOI: 10.1145/356842.356846).

[2]. Bernstein, P. A. and Newcomer, E., "Principles of Transaction Processing for the Systems Professional", Morgan Kaufmann, 1997.

[3]. Ceri, S. and Pelagatti, G., "Distributed Databases: Principles & Systems", McGraw-Hill.

[4]. Desai, B. C., "An Introduction to Database Systems", Galgotia, 2000.

[5]. G. Schlageter, "Problems of Optimistic Concurrency Control in Distributed Database Systems", ACM SIGMOD Record, 13(3):62–66, April 1982 (doi>10.1145/984505.984510)

[6]. http://www.cs.helsinki.fi/u/jplindst/

[7]. http://www.vldb.org

[8]. Kumar, V., "Performance of Concurrency Control Mechanisms in Centralized Database Systems", Prentice-Hall, 1996.

[9]. Özsu, M.T. and Valduriez P., "Principles of Distributed Database Systems", Prentice-Hall, 3rd Edition, 2011

[10]. Ramakrishnan, R., and Gehrke, J., "Database Management Systems", McGraw Hill International Edition, 2nd Edition

[11]. Silberschatz, A., Korth, H.F., and Sudarshan, S., "Database System Concepts", McGraw Hill International Edition, 5th Edition

[12]. Joshi, H. and Bamnote, G.R., "Distributed Database: A Survey", International Journal Of Computer Science And Applications Vol. 6, No.2, Apr 2013

[13]. Rawashdeh, Obaidah A., Hiba A. Muhareb, and Nedhal A. Al-Sayid. "An optimistic approach in distributed database concurrency control", 2013 5th International Conference on Computer Science and Information Technology, 2013.

[14]. Gupta, S., Saroha, K. and Bhawna, "Fundamental Research of Distributed Database", Intl. Journal of Computer Science and Management Studies, Vol. 11, Issue 02, Aug 2011

[15]. "Distributed Databases", Encyclopedia of Information Systems, 2004

[16]. Zubi, Z.S., "On Distributed Database Security Aspects", IEEE Explore, 2009

[17]. Tomar, P. and Megha, "An Overview of Distributed Databases", International Journal of Information and Computation Technology, Volume 4, Number 2 (2014), pp. 207-214

[18]. Yadav, A.K. and Agarwal, A., "A Distributed Architecture for Transactions Synchronization in Distributed Database Systems", International Journal on Computer Science and Engineering, Vol. 02, No. 06, 2010, pp. 1984-1991

[19]. Harding, R. and Aken, D.V., "An Evaluation of Distributed Concurrency Control", Proceedings of the VLDB Endowment, Vol. 10, No. 5, 2017

[20]. Batra, N. and Kapil, A.K., "Concurrency Control Algorithms and its Variants: A Survey", AIP Conference Proceedings, 2010

[21]. Sinde, V. and Aware, P.A., "Concurrency Control in Distributed Database Systems", International Journal for Research in Engineering Application & Management (IJREAM), Vol-01, Issue 10, JAN 2016

[22]. Kahoro, P., Wanjiru, C. and Karumba, N., "Concurrency Control In Distributed Databases", United States International University Africa, July 22, 2014

[23]. Abuya, T.K. and Cheruiyot W.K., "Guaranteeing Global Conflict Serializability in Concurrent Distributed Database Systems using Commitment Ordering", International Journal of Computer Science and Mobile Computing, Vol.3 Issue.6, June 2014, pg. 700-707

[24]. Sah, M.K., Kumar, V. and Tiwari, A., (2014) "Security and Concurrency Control in Distributed Database System", International Journal of scientific research and management (IJSRM), Volume 2, Issue 12, Pages 1839-1845, 2014

[25]. Gupta M.K. and Arora, R.K. (2009), Concurrency Control Techniques in Distributed DBMSs: A Comparative Study, IIMT Business Review 1 (1), 27-31

[26]. Vasileva S. And Milev A. (2016), "Simulation Studies of Distributed Two-phase Locking in Distributed Database Management Systems" , Information Technologies and Control, 13(1-2), DOI: 10.1515/itc-2016-0010

Author's Biography

Dr. Manoj Kr. Gupta is presently working as Professor at Rukmini Devi Institute of Advanced Studies (Aff. to Guru Gobind Singh Indraprastha University), Delhi, India. He is also Dean Examination, Admission & Administration and Course Coordinator MBA in the Institute. He has more than 20 years of experience in teaching and administration. His interest areas are Database Systems, Data Warehousing and Data Mining. He has 4 books and 18 international / national research papers to his credit.

Dr. Rakesh Kr. Arora is currently working as Associate Professor at Krishna Engineering College, Mohan Nagar, Ghaziabad, Uttar Pradesh, India. He has more than 16 years of teaching experience in reputed institutes. His interest areas are Data Warehousing and Data Mining, Database Systems and Operating System. He has number of papers in International Journals and Conferences to his credit.

Bhoopesh Singh Bhati is Pursuing Ph.D. degree from the Guru Gobind Singh Indraprastha University, Delhi. He has obtained his M.Tech. Degree in Information Security and B. Tech. (Computer Science and Engineering) from the Guru Gobind Singh Indraprastha University, Delhi. He is presently working as Assistant Professor in the Department of Computer Science and Engineering at Ambedkar Institute of Advanced Communication Technologies & Research, Govt. of NCT, Delhi, India. He has a number of Research Papers in International Journals and Conferences. His current research area is Information Security.

How to Cite