ISSN (Print):2321–404X, (Online):2321–4384

INTERNATIONAL JOURNAL OF SOFT COMPUTING & ARTFICIAL INTELLIGENCE

VOLUME NO. 11
ISSUE NO. 2
MAY - AUGUST 2023

ISSN (Print):2321–404X, (Online):2321–4384

International Journal of Soft Computing & Artificial Intelligence

Aim & Scope

IJSCAI provides a new forum for dissemination of knowledge on both theoretical and applied research on Artificial Intelligence and Soft Computing with an ultimate aim to bridge the gap between these two non-coherent disciplines of knowledge. The Journal offers survey and review articles from experts in the field, promoting insight and understanding of the state of art, and latest trends in the field. The content includes original research and innovation ideas, applications from all over the world.

All published papers are also available freely with online full-text content. This forum accelerates interaction between the above two bodies of knowledge, and fosters unified development in next generation computational models for machine intelligence. Special Issues devoted to topics of current interest will be published occasionally.

ISSN (Print):2321–404X, (Online):2321–4384

Editor In Chief

Publication and Distribution In-Charge

International Editorial Board

ISSN (Print):2321–404X, (Online):2321–4384

International Journal of Soft Computing & Artificial Intelligence

(Volume No. 11, Issue No. 2, May - August 2023)

Contents

ISSN (Print):2321–404X, (Online):2321–4384

Human Pose Estimation Based Robust Gait Generation and Estimation in Humanoid Robot

[1] Amrit Das, [2] Ankita Sahu, [3] Rohan Giri
[1]Student, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India
[2]Student, Kalinga Institute of Industrial Technology, Patia, Bhubaneswar, India
[3]Student, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India
E-mail: [1]amritdas1998@gmail.com, [2]ankitasahu0018@gmail.com, [3]girirohan36@gmail.com

ABSTRACT

This paper proposes a methodology and observations to achieve adaptive and autonomous gait generation and correction mechanisms in humanoid robots. The methodology used in this paper has been implemented in Robotis Darwin Humanoid Robots for dynamic gait generation in uneven terrain and inclined plane. The paper follows a two-thread process where the first thread deals with generating the gait pattern and after implementing the generated gait, the second thread has a feedback mechanism for gait correction. 2D Pose Estimation is used on Human Gait videos to extract gait cycle pattern which is normalized and applied to the Robotic actuators. An IMU sensor present in the Centre of the Robot provides the feedback data for the gait correction. The cumulative score of the angles generated by the Gait Generation Unit and the Feedback Unit are fed to the actuators to perform a robust walk cycle on the Humanoid Robot.

Keywords - Robotics, Legged Locomotion, Robot Dynamics, Robust Control

I. INTRODUCTION

Humanoid Robotics has received considerable attention in varying areas of research due to their malleability for various applications. Researches on Humanoid Robotics are not just limited to their application but also their locomotion, vision, gripping mechanisms etc. Our work focuses on Bipedal Locomotion and it aims to provide a Robust Control for generation and correction of Humanoid Gait Pattern. Locomotion in Humanoid Robot is one of its essential features as it is extremely necessary for the robot to move around freely in any surface including uneven and unstable. But in dynamic surface conditions, it becomes difficult to maintain stability and balance as the Center of Gravity (CoG) can easily be displaced from the hull resulting in complete collapse of the Robot.

In a practical scenario, it is easier for any Human to dynamically walk in unstable surfaces as they can arrest falling by conserving moment or/and by varying their stride and movement of hand. In order to achieve this flexibility, we intend to attain robust control for dynamic balancing of Humanoid robots. The primary objective of our work is to achieve adaptive robust control for gait generation and optimization. Our challenge in the real time scenario is to protect the robot from collapsing. For this purpose, an efficient and effective control system unit has been designed for error estimation and angle corrections in the gait. It basically comprises of two units- GGU (Gait Generation Unit) and Feedback

Unit. The GGU generates gait pattern by observing humans and converting Human Gait pattern into Robot actuation by normalizing the limb length. Normalization is required and is essential so as to actuate upon the angle limitation of the servo motors which is discussed in below section and generates angles in order to determine the motion. After generating an acceptable gait for the Robot unit, the Feedback Unit comes into play; here an IMU sensor is connected at the centre of robot to determine its body force, angular rate and orientation. In order to determine these parameters, the IMU sensor uses a combination of gyroscopes, accelerometers. The feedback from the sensor determines whether correction in GGU generated angles is required based no threshold detection. Based on this feedback, a Limb Manager Unit produces angle corrections for the actuators to stabilize the bot.

Our proposed system thereby provides a mechanism that not only generates the gait but also corrects the gait pattern to provide adaptive and Robust Gait Control for the Humanoid Robot. The system has been tested on Robotis Darwin, the proposed system; results and observation are described below in the paper in the following manner. Section 2 of the paper deals with previous relevant works done on Gait Generation and Estimation followed by the System Setups in Section 3 and Model Overview in Section 4. The results of our experiments are described in Section 5, the Conclusion and Future Scopes are mentioned in Section 5.

II. RELEVANT WORKS

Seungsuk et. al [1] proposed an approach of developing a balanced gait pattern resembling a human gait pattern. They have resolved it in points of torque in the sagittal plane together with ZMP (Zero Moment Point) in the frontal plane. They have generated a gait pattern similar to the human ones using the genetic algorithm. [2] Suggested the way of getting the data from the observation images andpresenting a way of gait pattern generation using neural networks. In [1], [3] have outlined the simultaneous trajectories of ZMP and COG (Centre Of Gravity) through planning and which results in a stable and smooth gait pattern. We have devised a methodology in which this can subdue the method of finding the gait pattern more accurately and systematically. The pattern generated using the feedback mechanism is used to complete the gait cycle and rectifying the error. [4] has used a Lower Limb Exoskeleton Robot which records the gait sequences which are characterized to derive the gait features. The patterns that are recorded are very similar to the real trajectories and these are then enforced on the Exoskeleton robot. Increasing the training sets in [4] helps in increasing the accuracy of the gait pattern. Guo et. al [5] proposed a method on height consideration of the COM(Centre Of Mass) for biped humanoid robots. This COM can fix the swinging leg's affect getting stable and efficient gait, depleting the model defect. It is challenging to do 2D pose estimation from image datasets than estimating from the depth maps due to the vagueness of perspective and self occlusion [6].The learned neurons present in the multi-stage system with Convolutional Neural Networks are tuned [6] to the confined body parts.

The pose estimation is observed as a regression task which further trains the model to find the pose and feature space from the image datasets. In a pose estimation method from graph-based models, the connection between the different body parts is configured that matches the configuration are by a distribution function. [7,8] We try to merge the works done by [9, 10] where [9] shows a method to laterally balance a robot in uneven surfaces and [10] shows a tele-kinetic method which we have modified in order to act it as a GGU about which it is further discussed in section 4. [11] has discussed about human identification from gait, and we used their work to reverse implement, human walk into gait and [12] helped in determining 3D poses from images. Over all there has been various researches in the field of 2D Pose estimation which has been our basis to generate gait, but in order to correct and make the walk pattern adaptive, we have designed the feedback system, as discussed in Section 4.

III. THE HUMANOID ROBOT

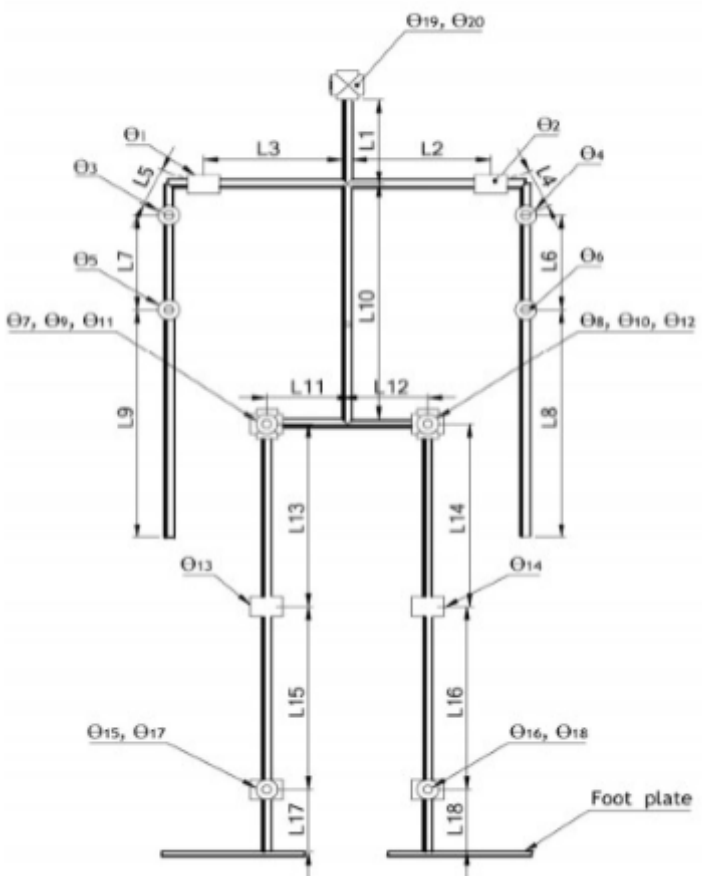

Figure 1: Robotis Darwin Humanoid Robot in a stick diagram

The Robot uses Dynamixel Smart Servos (MX28T), which actuate based on angle input and can provide current-position feedback to the microcontroller controlling the Robot. The real-time orientation and motion feedback are provided by an IMU Sensor located in the centre of the Robot. The microcontroller used by us is Hardkernel's Odroid XU4. Other technical specification regarding the Robot can be found in Table 1.

Parameters	Specification
Height, Weight	45cm, 2.5Kg
Degree of Freedom	20
Actuators	MX28T Servos
Sensor	IMU Sensor
Microcontroller	Odroid XU4
Operating System	Linux
Power Source	Battery (3S 2200mAh Li-Po)

Table 1

For our experiment, we have used a 20 DOF, Humanoid Robot. Figure 1, represents the stick diagram of the Humanoid Robot, actual image is shown in Figure 2.

Figure 2: Robotis Darwin humanoid Robot

IV. SYSTEM REPRESENTATION AND OVERVIEW

Our model enables the Humanoid Robot discussed, in Section 3, to generate a stable walk pattern for the Humanoid Robot and to attain robust control by adapting to unknown environment. Our Robot is a fixed linked robot, that is, the hands are not responsible for conserving momentum during the walk pattern and remain fixed to the hip of the robot while the Robot moves. The momentum is conserved due to the weight of the robot and stride length. The system model can be represented as Figure 2.

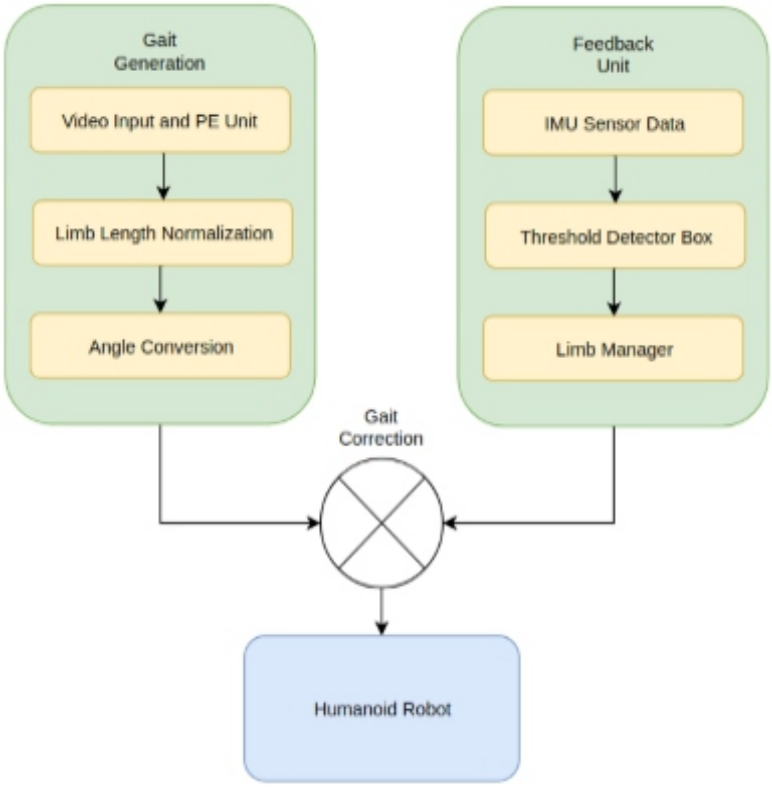

Figure 3: Model Representation

As per Figure 2, our system incorporates two major setups; the first one is external setup and is used to generate a gait pattern which resembles that of Human. This setup provides a sample gait pattern for the Robot to start moving forward. In order to generate this gait data, our system follows the following steps.

1) Pose Estimation – Our major aim of the experiment is to generate the Gait pattern. In the initial stages, this was done either by brute force and hard coding or by generating mathematical model based on the structure of the robot. Either case it is very difficult to repeat the process. We in our work have used Neural Network model to predict the pose of a Human in Real Time. This mechanism helps in generating gait cycle that resembles a Human.

2) Limb Length Normalization – The PE Unit generates an optimized walk cycle for us but this cannot be directly implemented on the Robot because of various constraints like weight of the Robot, height of the Robot, ration between their limbs etc. Due to these factors, we need to normalize the values generated from the Human Pose Estimation into suitable value that can be implemented for the Robot Pose. This Normalization is done by the Limb Length Normalization Unit that takes input of the ratio of the various limb lengths of the robots and normalizes it with the ratio generated by the Human Pose.

Joint	Angle	Joint	Angle
R-Hip	-46	L-Hip	55.96
R-Knee	-71.48	L-Knee	91.99
R-Ankle	-36.45	L-Ankle	46
R-Foot	-6.45	L-Foot	9.95

Table 2

Table 2 shows a sample of the angle generated by the Gait Generation Unit. The generated angles as per Table 2 are normalized because the smart servos have a working range of 00 to 3000 which is mapped in between -150 to 150, 0 being the mid-point. These angles are generated from the output received as per Figure 4, using equations (3) as below. These are then mapped to contain in between -150 to 150 as per our mapping convention.

3) Angle Conversion – Once we have the desired Pose, the estimated pose is converted into angles by using conventional vector algebra equations.

$$\overrightarrow{AB} = (A_x - B_x)\hat{\imath} + (A_y - B_y)\hat{\jmath} + (A_z - B_z)\hat{k} \quad (1)$$

$$\overrightarrow{BC} = (B_x - C_x)\hat{\imath} + (B_y - C_y)\hat{\jmath} + (B_z - C_z)\hat{k} \quad (2)$$

$$\theta = \cos^{-1}\left(\frac{\overrightarrow{AB}.\overrightarrow{BC}}{(|AB||BC)}\right) \quad (3)$$

Equations (1) and (2) are use to convert the link lengths into vectors. Equation (3) is vector algebra equation to calculate angle between two lines. These equations are used to calculate the angle at hip joints, knee joint, ankle joint and foot joint on either side.

These angles are again normalized to match the orientation of the motors and are stored as arrays with indexes that can be referred form Table 2.

Joint	Index	Joint	Index
R-Hip	1	L-Hip	2
R-Knee	3	L-Knee	4
R-Ankle	5	L-Ankle	6
R-Foot	7	L-Foot	8

Table 3

Once the GGU (Gait Generation Unit), completes the generation of a walk cycle by learning from Human walk pattern, we move onto our next phase of gait correction. The angles generated are implemented on the actuators and the robot is allowed to move freely. But due to physical constraints, the Robot walk pattern could still get the Robot collapsed in order to prevent this, we incorporate the Feedback Unit. The various stages for receiving and implementation of the feedback are as follows.

1) Data from Sensor - We are using an IMU sensor to get a feedback of the current orientation and motion of the Robot. IMU Sensors have inbuilt gyro sensors and accelerometer that can give an approximation of roll, yaw, pitch motions that is filtered by us to determine the current orientation of the Humanoid Robot.

2) Threshold Detector - From a hardcode brute force trial it was noted that correction was only required in the 3rd and 4th phase of the generated walk pattern where the Robot will tend to fall forward due to misplacement of its CoG outside the hull. Hence whenever any such event is predicted by the data from the IMU Sensor, an interrupt flag is raised, which turns on an estimator unit that estimates the orientation of the fall of the Robot, and based on the feedback value from the sensor, generates a stride length that would arrest the fall of the Robot preventing the collapse of the motion. This stride length is generated by Limb Manager.

3) Limb Manager - Based on the threshold value/parameter that is generated by the IMU, the Limb Manager generates angle values for the Hip, Knee and Ankle motors. The best method to calculate these angle values according to us would be by using Reinforcement Learning method, but due to the lack of resources we have rather used a hit and trial method to find a proportional value for angles based on the sensor data. This is not very accurate but it works well for inclined surfaces and slightly uneven surfaces.

The angles generated by the Gait Generator can be summed up with the feedback angles generated by the feedback unit to improve the walk cycle.

V. OBSERVATION OF EXPERIMENTAL RESULTS

Section 4 discusses about the two methods to generate and to correct the angles generated. The cumulative summation of both the methods produces a correct posture that can be mimicked by the robot to generate an adaptive Gait Pattern. The result of the pose estimation is shown in Figure 3.

The estimated pose cannot be directly implemented on the robot; hence after pose generation as discussed in Section 4, the values are normalized. These normalized angles are corrected using the

Feedback Unit and the summation of angle is taken for record and implemented for actuation and can be seen in Figure 4.

Figure 4: Pose Estimation Results

Figure 5: Final Angles implementation on Robot

This paper summarizes the design of biped robots for gait generation and control as well as adaptive gait correction to ensure the maintenance of the robot's locomotion. The proposed model clearly depicts the methodology to prevent the robot from collapsing on uneven surfaces.

The generated gaits receive feedbacks from the IMU sensors which help to determine the locomotion of the robot. In addition to that, a correction unit has been implemented which supervises the robot's balance on uneven and rugged surfaces. This enhances the durability and robustness of the robot, exposing it to a wider range of applications.

The application of the feedback can be better understood based on Figure 6 which represents the ground truth, robot stabilization on plane surfaces and robot stabilization on unstable or dynamic surfaces. Figure 6 (c) represents how the feedback system can help stabilize the robot on unstable surfaces. Whenever the robot tries to cross the positive or negative threshold values, the robot has a tendency to fall, in order to arrest the fall here, the feedback system will generate angle against the angles produces as produced by gait unit. Sample can be seen as per Table 2.

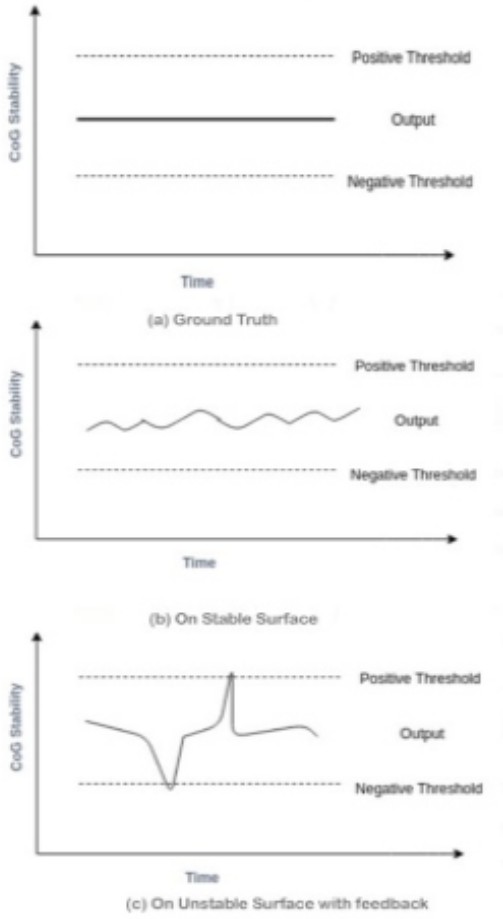

Figure 6: Feedback System Implementation

VI. CONCLUSION AND FUTURE SCOPE

The propose method aims to contribute in the advancement in the legged locomotion field of Humanoid Robotics. Our proposed systems has used a brute force method, which helps the Robot to adapt to its surface conditions and maintain balance in inclined surfaces and slight uneven surfaces.

The design of biped robots are mainly focused on investigation of gait and control systems for locomotion on stable and even surfaces. The proposed model is not only based upon gait generation and control, but also on the adaptive gait correction techniques. However, it takes the hit and trial approach to find out the correct position which isn't feasible. In order to get a more accurate result, reinforcement learning approach [13] can be taken into consideration. This increases the probability of getting the correct result after every process of learning. Even if it is durable and withstands inclined and uneven surfaces to some extent, it can be made more efficient with its locomotion on highly uneven terrain to maintain its balance.

REFERENCE

[1] Seungsuk Ha, Youngjoon Han, and Hernsoo Hahn, "Generation of Biped Robot based on Captured Human Motion Image," International Conference on Smart Manufacturing Application, 2008.

[2] Genci Capi, Yasuo Nasu, Leonard Barolli, Kazuhits Mitobe, "Real-time gait generation for autonomous humanoid robots: A case study for walking," Robotics and Autonomous Systems, vol. 42, pp. 107-116, Issue 2, Feb. 2003.

[3] Kensuke Harada, Shuuji Kajita, Kenji Kaneko, and Hirohisa Hirukawa, "An Analytical method for real-time Gait planning for Humanoid Robots", International Journal of Humanoid Robotics, vol. 3, no. 1,.pp. 1–19, 2006.

[4] Xinyu Wu, Du-Xin Liu, Ming Liu, Chunjie Chen, Huiwen Guo, "Individualized Gait Pattern Generation for Sharing Lower Limb Exoskeleton Robot," IEEE Transactions on Automation Science and Engineering, vol. 15, Oct. 2018.

[5] Guo Fayong, Zhao Jianghai, Mei Tao, Luo Minzhou, Song Xiaobo, "A modified gait generator for humanoid robots based on height compensation of center of mass, IEEE International Conference on Robotics and Biomimetics, 2014.

[6] Sijin Li Zhi-Qiang Liu Antoni B. Chan, "Heterogeneous Multi-task Learning for Human Pose Estimation with Deep Convolutional Neural Network".

[7] Felzenszwalb, P.F., Huttenlocher, and D.P. Pictorial structures for object recognition. IJCV, pp. 55–79, 2005.

[8] S. Johnson and M. Everingham, "Learning effective human pose estimation from inaccurate annotation," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Jul. 2011.

[9] Aditya Sripada, Harish Asokan, Abhishek Warrier, Arpit Kapoor, Harshit Gaur, Rahil Patel, Sridhar R, "Teleoperation of a Humanoid Robot with Motion Imitation and Legged Locomotion", 3rd International Conference on Advanced Robotics and Mechatronics (ICARM), July 2018.

[10] Aditya Sripada, Abhishek Warrier, Arpit Kapoor, Harshit Gaur, B Hemalatha, "Dynamic Lateral Balance of Humanoid Robots on Unstable Surfaces," International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Dec 2017.

[11] Murat Ekinci, Murat Aykut, "Human Identification Using Gait," IEEE 14th Conference Signal Processing and Communications Applications, 2006.

[12] Mun Wai Lee, I. Cohen, "A model-based approach for estimating human 3D poses in static images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, June 2006.

[13] Jan Peters, Sethu Vijayakumar, Stefan Schaal, "Reinforcement Learning for Humanoid Robotics," Third IEEE-RAS International Conference on Humanoid Robots, Sept. 2003.

ISSN (Print):2321–404X, (Online):2321–4384

A Review of Machine Learning for DNA Sequence Classification

[1] H. M. I. Janakee Herath, [2] K. A. Dilini T. Kulawansa
[1,2]Faculty of Information Technology Department of Computational Mathematics, University of Moratuwa, Moratuwa, Sri Lanka
E-mail: [1]janakeeherath@gmail.com, [2]dilini_kulawansa@hotmail.com

ABSTRACT

Machine learning is a data processing technique and it helps to make decisions, predicitions, classification and recognitions by using traning data. It is a subsection of Artificial Intelligence and do not follow any theory to do their tasks. Machine learning is used for different applications which are Agriculture, Computer vision, Gaming Industry, Linguistics etc. DNA Sequence Classification is one of its affected are in today. In here nucleic acid sequence and its order is determined in order to identifed the mutations, diagnose different diseases, recognise normal gene and abnormal gene etc. Machine Learning models of Convolutional Neural Network, Deepl Learing, Genetic Algorithm are used to this DNA Sequence Classification. This paper presents how the techniques of machine learning use to classify the DNA Sequence and issues of this process, applications of Machine Leraning and future implements to be done for enhancing this process.

Keywords - Machine Learning, Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Deep Learning, DNA Sequence Classification, Bioinformatics, DNA sequencing.

I. INTRODUCTION

Machine learning is one of subdivision of Artificial Intelligence and it enables machines to do real world tasks intelligently, while learning from data. Its algorithms used to analyze the data and they act as black box [27]. That means, when some data input to the black box and it outputs a number, likewise it produces output. Machine learning also divided into subcategories which are ANN, CNN, Deep Learning, Genetic Algorithm etc. When considering the neural network, it can be used for identifying the subset of data in a sequence and it have ability to identify the particular data separately. After reading the example of toy, it can be clearly understood [27]. In today machine learning is used for many applications which are fundamental researches in cheminformatics, bioinformatics, cosmology to quantitative social science, physics [21], agriculture, computer vision, gaming industry, linguistics etc.

DNA sequence expresses many information about species which are behaviors, appearance, their parent's information etc. Those information helps to identify a species separately form other species. Therefore, identifying the DNA Sequence order and their classification is significant need in today. Nowadays the Genetic expression is used frequently in medical sector. It applies to recognize the cancers, down syndrome, mutations, tumors etc. [7]. The goal of DNA sequencing is identifying the

order of nucleotides of a given DNA section. When talking about its classification, mainly it divided into two parts as ―Functional DNA‖ and the other one is ―Rubbish DNA‖ [2]. Functional DNA isdivided into another two parts called ―Literal DNA‖ and ―Indifferent DNA‖ [2]. Rubbish DNA also divided into another wo parts as ―Junk DNA‖ and ―Garbage DNA‖ [2]. During the evolution affiliation may happen among above DNA segments. There are three methods can be used to categorize the DNA sequence which are distance-based method, feature-based method and the other one is model-based method [9]. Understanding sequence of DNA is very important as well as accuracy and the speed of identifying is important. DNA sample have million of genes. Therefore, it is a difficult task to classify this. Most of the time in these days sequencing technologies are used for achieving that task. Machine learning is the newly introduced technology for solve that problem. ANN, CNN, Deep Learning, Genetic Algorithms are the recent machine learning techniques which are used to overcome from this issue.The position information of the DNA and their classification can be recognized by using the deep learning model with hot vectors [11]. The sequence of DNA inputs to these models and get the outputs. Likewise, different aspects of machine learning can be used to determine DNA sequence and their classification.

This paper is provided a description about how machine learning techniques use to classify the DNA sequence, some challenges of this process, some applications of machine learning and investigates some researchers in context of DNA sequence classification with the help of machine learning.

The remaining sections of this review paper is arranged as follows. Overview of machine learning to the field of DNA sequence classification is under second section. In the third section major researches in Use of machine learning for DNA sequence classification will describe and the fourth section investigates about the current issues in Use machine learning for DNA sequence classification. In the fifth section discovers about application of machine learning. The discussion is under sixth section and at last we describe about the future direction related to this topic.

II. OVERVIEW OF MACHINE LEARNING AND DNA SEQUENCE CLASSIFICATION

Considering about modern systems, involvement of Artificial Intelligence has become turning point for that systems. Studying about AI and its subdivision is a need during these days. Considering about their applications, DNA sequence classification gets one of a major role in today. Hence, discussing about this topic is an important fact.

A. Machine Learning

The technology machine learning is a turning point for modern systems. Therefore, discovering about it is also significant. Before a deep studying about machine learning, knowing its origin is important. Machine learning is founded after the AI.The birth of AI was in 1956. AI is mainly divided into two subdivisions which are known as ―Machine Learning‖ and the ―Artificial Cognitive Systems‖.

An artificial cognitive system is based on symbolic base problem solving, reasoning, thinking etc. and this involves theories. Expert Systems, Robotics, Natural Language processing, Fuzzy Logic and Multi Agent Systems are come under this division.

Machine Learning is about training by large amount example even without following any theory. Machine learning was founded in 1959 by Arthur Samuel. ANN, Deep Learning, CNN and Genetic Algorithms are come under machine learning.

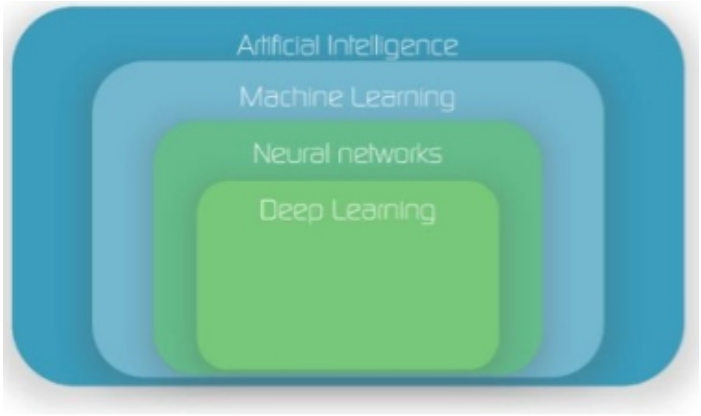

Figure 1: Difference between AI and the machine learning.

According to their approaches types of machine learning algorithm are different to each other. In different situations the input and out types are distinct not only that the type of the problem also different. So that there are main three algorithms which are supervised learning algorithms, unsupervised learning algorithms and reinforcement learning algorithms.

In Machine learning the sample data called as ―training data‖ [27] and they are used to make decisions and predictions which are relevant to some problem or task.

Considering about machine learning applications, there are lots of areas it covered. Few of them are fundamental researches in cheminformatics, bioinformatics, cosmology to quantitative social science, physics [21], agriculture, computer vision, gaming industry, linguistics, Economics etc.

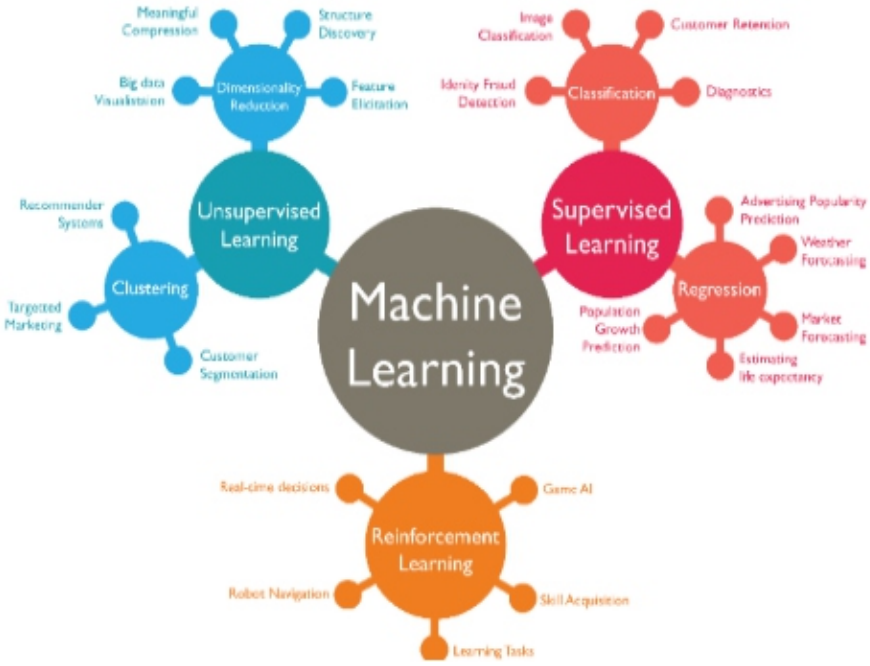

Figure 2: Types of machine learning

B. DNA sequence classification

The sequence data of DNA is a main object of bioinformatics study. When considering the term of DNA sequencing, it means the process of finding the order of nucleotides in a given nucleic acid sequence. The term classification meansclassify the nucleic acid or its combinations which called as gene into separate sections.

In [2] considering the selected-effect function genes are classified into main two categories which are ―Functional DNA‖ and Rubbish DNA‖ [2]. Functional DNA has the unselected-effect function and Rubbish DNA does not have that function. Functional DNA isdivided into ―Literal DNA‖ and ―Indifferent DNA‖ [2]. InLiteral DNA, nucleotides order is under selection and only the presence or absence of the sequence of Indifferent DNA is under selection. When considering the Rubbish DNA, it is also divided into ―Junk DNA‖ and Garbage DNA‖ [2]. Junk DNA contributes from the organism fitness and due to that it evolves under selective neutrality [2]. Garbage DNA reduces fitness its carriers. Above categories of DNA can be transcribed and translated or transcribed but do not translated or not transcribed [2]. In evolution, the affiliation of a DNA section to an exact functional category can be changed. For an example Functional DNA can be changed as Junk DNA and Junk DNA may be changed as Garbage DNA and so on [2].

Analyze and interpretation of DNA data are two main difficult tasks in bioinformatics. Classification and also prediction methods are the major techniques to address these tasks. In [9] there are three main categories of DNA sequence classifications methods are included. They are distance-based method, feature-based methods and mode-based method.

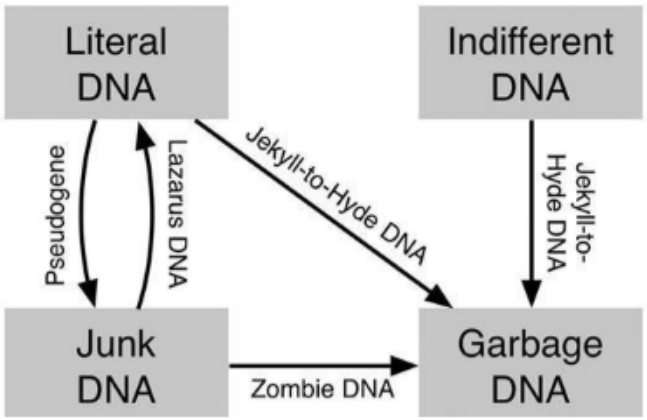

Figure 3: DNA categories

When considering DNA sequence classification, the genetic data helps to identify species from another species separately. In today in medical sector is used this technology for diagnose diseases, recognize tumors, identity the mutation, select the treatment to overcome some illness and etc. [7]. These are the importance of this process.

III. MAJOR RESEARCHES IN USE OF MACHINE LEARNING FOR DNA SEQUENCE CLASSIFICATION

After examining the value of this area, people tried to find some newest things to implement this. They wanted to do this process faster and with accuracy. Here some significant and interesting contribution of other researcher have done so far relevant to this study area.

Liang you Chen and Lois Boggess conducted a research about gene classification by utilizing the genome signature. They used the technology of neural networks and its four methods of back-propagation, functions of radial-basis, committee machines and self-organizing maps for the purpose of studying this matter[23]. After the experiment, they got the best result from committee machine and it had 16.88% of average error.

When considering the research done by Dr. P. Kiran Sree, Dr. P.S.V. Srinivasa Rao, S.S.S.N. Usha Devi N introduced a new concept for gene prediction and a new unsupervised classifier which is used machine learning model of deep learning [5]. It used hybrid cellular automata with the environment of deep

learning to overcome some issues in bioinformatics. They able to achieve 98.7% in .8 nano second accuracy with this classifier. According to Vrinda V. Nair, Karthika Vijayan, Deepa P. Gopinath and Achuthsankar S. Nair done a research to classifying the unknown genomic fragments and it is based on ANN and also used Chaos game representation [4]. Eight subsets from taxonomical distribution which are belong to the Eukaryotic organism were gathered to evaluate this proposed method. Distinct ANN configurations were tested, and also significant accuracy was obtained.

Using Neural Network and expectation maximization algorithm Qicheng Ma, Jason T. L. Wang, Dennis Shasha and Cathy H. Wu done a research to classify the DNA sequence.[12]. They presented new concept for DNA sequence classification and it was focused on identifying E. Coli promoters in their DNA. In here determined given DNA sequence whether it is E. Coli promoter or not. Above problem is called as binary classification problem.

Jun Miyake, Yuhei Kaneshita, Takashi Hirano, Satoshi Asatani, Seiichi Tagawa, Hirohiko Niioka investigated a new approach to classify the DNA sequence by using deep learning [14]. In here this classification is done as a graphical classification and used the DNA of human leukocyte antigen alleles. Data is collected from the Database of Immuno Polymorphism and the data are compressed to two-dimensional representation.

When considering the research done by Karthika Vijayan, Deepa P. Gopinath, Achuthsankar S. Nair, Vrinda V. Nair investigated a new method for organism classification which is developed on combination of FCGR and ANN [4]. They addressed problem of organism classification into distict categories by using genomic sequence.They used few distinct species for this. They got 86.8% of accuracy for PNN based calssification. By using machine learning and data mining Jonathan Auerbach, Damian Gola, Elizabeth Held, Emily R. Holzinger, Marc- Andre Legault, Rui Sun, Nathan Tintle, Hsin-Chou Yang, Inke R. König were adressed the problem in complex genomic data classification [17]. They felt analsing problems and also combination of distinct data strudtures. Another research regarding to classification of metagenomic and genomic sequence by using discriminative k-mers. This was investigated by Steve Wanamaker, Timothy J Close, Stefano Lonardi, Rachid Ounit and proposed a new accurate and efficient tool to classify the objects which is based on the reduced set of k-mers [6]. 10% of metagenomic sequence were choose from every of the data set in 10 compenent genome. When considering the research is done by Zhaoli which is naming Maching Learning in Bioinformatics give an overview of the involvement of machine learining to the field of Bioinformatics [19].In here addressed some problems in bioinformatics and identified problem in data pre-processing and also analyzing those data. Another problem identify is a analyzing the biological images.

Kun Wang, Huixiao Li, Yang Jia, Xiaoqin Wu, Yaning Du, Wei You was conducted a research regarding to DNA sequence classification which is based on the dinucleotide compostions. They were introduced a new method for classifying the DNA sequence in bacteria [8]. They were got 84.3 % accuracy in the classification and used only 4 kinds of bacteria.

Brian Hudson, David Whitley, Martyn Ford, Phil Picton, Hassan Kazemian, Antony Browne were conducted research about neural networks [15]. They provide an overview about existing algorithm regrading to neural network. They dicussed about decision tree and their limitations and data mining technique relevant to this topic area.

IV. CURRENT ISSUES IN USE OF MACHINE LEARNING FOR DNA SEQUENCE CLASSIFICATION

Some problems can be seen under this topic. Few of them are discussed in here.

A. Classification of Biomedical Data and Prediction.

When considering the issues, there is an issue in the classification of biomedical data and the prediction of those data. It is about find a unique assessment metric for above classification and for the prediction [10]. This is a difficult task can be seen in this study area.

B. Storage of DNA Sequencing Data.

There is another issue with storage of large amount of DNA sequencing data. These data should be storage in a proper way and if it is stored like that, then it would be easy to access those data. As well as it should be reliable to access [20].If the data is meaningless, there isnotany available data. So that, it is very important fact to be concerned.

C. Binary Nonlinear Classification Issue.

When examine the issues, there is an issue with the DNA sequence classification. It can be recognized as the binary nonlinear classification issue [24]. Binary classification is also called as binomial classification which means the classification of given elements set into two distinct groupsbased on the classification rule.

D. DNA Sequence Length Normalization.

There can be see another issue in DNA sequence length normalization. Normally genomes are different in their length and length normalization usually results in adding which is called as up-sampling or losing which is known as down-sampling, some of information [22]. This may be made problems when dealing with the subset of genomes and lager genome sequence.

E. Binning Step

When considering the issues, binning step is the one of main difficult task that happen in analyzing the metagenomic data. It happens in the process of assigning the sequenced read to a taxonomic clade [16]. Binning method must be faster and also accurate when dealing with large amount of metagenomic datasets. This is also a challenge.

F. Problems in Support Vector Machine Learning. There are some problems in Support Vector MachineLearning. One of them is convex optimization issues [1] which is arise in the research conducted by Qingshan Jiang, Dan Wei, Qingda Zhou. Another one is detecting problem [3] and it is in the research conducted by Shailendra Singh, Trilok Chand Aseri, Neelam Goel. In here they used Support Vector Machine Learning along with the markov model

G. Optimization Problem

Optimization problem is arisen in the analyzing process of genome. As a solution to this problem Genetic Algorithm was introduced by John Holland in 1975 [26]. It is a computing search technique. GA used evolution principles found in nature to solve the problem of finding an optimal solution.

V. APPLICATIONS OF MACHINE LEARNING

Development of Machine Learning has been influenced by many other areas. For an example Cheminformatics, cosmology, physics, analysis of gravitational wave, detection of black hole, material design [21], Linguistics, recognition of images [10] search engines can be considered. Here few applications of Machine Learning are studied.

A. Cheminformatics.

This is the study about use of computer and also informational techniques used to a range of issues in chemistry. In here machine learning is used for database mining process, molecular mining process, virtual screening etc. There were so many researches were conducted for finding new approaches of this.

B. Virtual Personal Assistants.

When considering the example of virtual personal assistants systems Siri, Google and Alexa are few of them. After asking some questions it finds solutional and answering for the question. For answering, it finds information, recalls the related queries otherwise send a command to collect information to the other resources. Machine learning is a significant part of these virtual personal assistant systems. It collects and refine the relevant information which is based on the user's previous engagement with them. After that, above data is used for final outcomes which are tailored to user desires.

C. Search Engine.

Google, Yahoo and other search engines utilize machine learning to implement the finding outcomes for the users. Each and every time user executes a search and algorithm in the back end keep attention at how user respond to relevant outcomes. This process also uses the machine learning for implementations.

D. Email Spam.

Considering the email spam, there are some spam filtering approaches which can used by the email client. To prove that this spam filters are always updated, and they are usually powered by the modes of machine learning. Not only that decision tree induction and multi-layer perception are powered by the machine learning.

E. SLTR/SINE PROBLEM

In human genome there are lots of copies of SINEs and LTRs. Considering the example of SINE Alu, it has 300 000 copies, one for each 6K of the genome. Due to the fact that, entire ERVs are referred to as LTRs, then refer their LTR portion as an sLTR. We called this issue as sLTR/SINE problem [13].

VI. CONCLUSION

With the introduce of machine learning, lot of system become more familiar to user and easy of use since different applications were created. Then machine learning become more popular with people. Through this study, we discover its usage to the field of DNA sequence classification. During this process there are so many challenges, issues, drawbacks were created and few of them are discussed here. Different people done different researches to avoid these barriers and tried to find new concepts and methods for implement this process with more efficiency and in a very accurate way and here we have studied about 27 research papers.

VII. FUTURE DIRECTIONS OF USE OF MACHINE LEARNING FOR DNA SEQUENCE CLASSIFICATION

Different researches were conducted by different people base on the topic of DNA sequence classification with the help of machine learning. There are so many researches still doing their researches for further implementation of this area.

Some researchers have an idea for building the classifier of state-of-the-art compositional in the near future [16]. Katya Rodriguez, Roberto A. Vazquez, Beatriz A. Garro planned to study about deep analysis of parameters that helps to train ANN and the accuracy increment of the the propsed methodology [7].Another direction is that introduce a hybrid system which is integrated with non-linear

cellular automata and fuzzy sets [5].Vu Anh Tran and his other researchers aim to applythe method that they introduced, using with the CNN for other sequence classification [11]. There is an another interesting direction is wished to develop in near future by Abida Sanjana Shemonti and his team. It is about to develop a new model to recognize a species from its any part of DNA sequence [25]. Shailendra Singh and other members expect to study new area of predicating the low fequency non – canonical splice sites [3]. Not only that, Tingting Zhu and his team wish to discover some new resistance mechanism and improve the performance of existing method [18].

ACKNOWLEDGMENT

Doing a research in the area of Machine learning and DNA Sequence Classification is somewhat a difficult task. So, I am thankful to my parents for supporting me to do this work. And also, the people who help me to do this research. Not only that I am thankful to people who have done a great service by dedicating their lives to research about this area.

REFERENCE

[1] Qingshan Jiang, Dan Wei, Qingda Zhou, "A New Method for Classification in DNA Sequence," in The 6th International Conference on Computer Science & Education, 2011.

[2] Yichen Zheng, Ricardo B. R. Azevedo, Dan Graur, "An Evolutionary Classification of Genomic Function," vol. 7, no. 3, p. 4, 2015.

[3] Shailendra Singh, Trilok Chand Aseri, Neelam Goel, "An improved method for splice site prediction in DNA sequences using support vector machines," in 3rd International Conference on Recent Trends in Computing, 2015.

[4] Karthika Vijayan, Deepa P. Gopinath, Achuthsankar S. Nair, Vrinda V. Nair, "ANN based Classification of Unknown Genome Fragments using Chaos Game Representation," in Second International Conference on Machine Learning and Computing, 2010.

[5] Dr P. S. V. Srinivasa Rao, S.S.S.N Usha Devi N, Dr P. Kiran Sree, "CDLGP: A Novel Unsupervised Classifier using Deep Learning for Gene Prediction," in IEEE International Conference on Power, Control, Signals and Instrumentation Engineering, 2017.

[6] Steve Wanamaker, Timothy J Close, Stefano Lonardi, Rachid Ounit, "CLARK: Fast and accurate classification of metagenomic and genomic sequences using discriminative k- mers," vol. 16, p. 13, 2015.

[7] Katya Rodriguez, Roberto A. Vazquez, Beatriz A. Garro, "Classification of DNA microarrays using artificial neural networks and ABC algorithm.," vol. 38, p. 13, 2015.

[8] Kun Wang, Huixiao Li, Yang Jia, Xiaoqin Wu, Yaning Du, Wei You, "Classification of DNA Sequences Basing on the Dinucleotide Compositions," in 2 nd International Symposium on Computational Intelligence and Design, 2009.

[9] Ngoc Giang Nguyen, Favorisen Rosyking Lumbanraja, Mohammad Reza Faisal, Bahriddin Abapihi, Bedy Purnama, Mera Kartika Delimayanti, Mamoru Kubo, Kenji Satou, Dau Phan, "Combined Use of k-Mer Numerical Features and Position-Specific Categorical Features in Fixed-Length DNA Sequence Classification," vol. 10, no. 8, p. 12, 2017.

[10] Feng Liu, Hai Tan, Deshou Song, Wenjie Shu, Weizhong Li, Yiming Zhou, Xiaochen Bo, Zhi Xie, Chensi Cao, "Deep Learning and Its Applications in Biomedicine," vol. 16, no. 1, p. 16, 2018.

[11] Vu Anh Tran, Duc Luu Ngo, Dau Phan, Favorisen Rosyking Lumbanraja, Mohammad Reza Faisal, Bahriddin Abapihi, Mamoru Kubo, Kenji Satou, Ngoc Giang Nguyen, "DNA Sequence Classification by Convolutional Neural Network," vol. 9, p. 7, 2016.

[12] Jason T. L. Wang, Dennis Shasha, Cathy H. Wu, Qicheng Ma, "DNA Sequence Classification via an Expectation Maximization Algorithm and Neural Networks: A Case Study," in IEEE Transactions on systems, man, and Cybernetics—part C: applications and reviews,, 2001.

[13] Suprakash Datta, Wendy Ashlock, "Evolved Features for DNA Sequence Classification and Their Fitness Landscapes," in IEEE Transaction on Evolutionary Computation, 2013.

[14] Yuhei Kaneshita, Satoshi Asatani, Seiichi Tagawa, Hirohiko Niioka, Takashi Hirano, Jun Miyake, "Graphical classification of DNA sequences of HLA alleles by deep learning," vol. 31, no. 2, p. 4, 2018.

[15] Brian Hudson, David Whitley, Martyn Ford, Phil Picton, Hassan Kazemian, Antony Browne, "Knowledge Extraction from Neural Networks," in IEEE International Conference, 2003.

[16] Pierre Mahe, Maud Tournoud, Jean-Baptiste Veyrieras, Jean- Philippe Vert, Kevin Vervier, "Large-scale Machine Learning for Metagenomics Sequence Classification," vol. 32, no. 7, p. 2015, 14.

[17] Jonathan Auerbach, Damian Gola, Elizabeth Held, Emily R. Holzinger, Marc-Andre Legault, Rui Sun, Nathan Tintle, Hsin-Chou Yang, Inke R. König, "Machine learning and data mining in complex genomic data- a review on the lessons learned in Genetic Analysis Workshop 19," p. 8, 2016.

[18] Katherine E. Niehaus, Timothy M. Walker, Zamin Iqbal, A. Sarah Walker, Daniel J. Wilson, Tim E. A. Peto, Derrick W. Crook, E. Grace Smith, Tingting Zhu, David A. Clifton, Yang Yang, "Machine learning for classifying tuberculosis drug-resistance from DNA sequencing data," vol. 34, no. 10, p. 6, 2018.

[19] Zhaoli, "Machine Learning in Bioinformatics," in International Conference on Computer Science and Network Technology, 2011.

[20] Ghanshyam I. Prajapati, Pooja Dixit, "Machine Learning in Bioinformatics: A Novel Approach for DNA Sequencing," in Fifth International Conference on Advanced Computing & Communication Technologies, 2015.

[21] Xiaopeng Li, S. Das Sarma, Dong-Ling Deng, "Machine learning topological states.," p. 12, 2017.

[22] Kathleen A. Hill, Lila Kari, Gurjit S. Randhawa, "ML- DSP: Machine Learning with Digital Signal Processing for ultrafast, accurate, and scalable genome classification at all taxonomic levels.," p. 23, 2018.

[23] Lois Boggess, Liangyou Chen, "Neural Networks for Genome Signature Analysis," in 9th International Conference on Neural Information Processing (ICONIP'OZ).

[24] De- Shuang Huang, Hong- Qiang Wang, Xing-Ming Zhao, Xin Huang, "Representation of DNA Sequences with Multiple Resolutions and BP Neural Network Based Classification," in IEEE International Conference, 2004.

[25] Abida Sanjana Shemonti, Atif Hasan Rahman, Tasnim Kabir, "Species Identification using Partial DNA Sequence: A Machine Learning Approach," in IEEE 18th International Conference on Bioinformatics and Bioengineering, 2018.

[26] Wajdi Bellil, Chokri Ben amar, Abdesselem Dakhli, "Wavelet Neural Networks for DNA Sequence Classification Using the Genetic Algorithms and the Least Trimmed Square," vol. 96, p. 10, 2016.

[27] Timothy Cohen, Bryan Ostdiek, Spencer Chang, "What is the Machine Learning?," vol. 97, no. 5, p. 6, 2018.

ISSN (Print):2321–404X, (Online):2321–4384

Performance Analysis of Lte- A Systems with Femtocell Overlays using Different Modulation Techniques Under Different Multipath Fading

[1] Haseeb Ahmad Qureshi, [2] David John, [3] Simant Prakoonwit
[1]Department of Computing and Informatics, Bournemouth University, United Kingdom
[2,3]Department of Creative Technology, Bournemouth University, United Kingdom
Email: [1]hqureshi@bournemouth.ac.uk, [2]djohn@bournemouth.ac.uk,
[3]sprakoonwit@bournemouth.ac.uk

ABSTRACT

The developments in mobile communication systems from 1G to 5G have increased demands on the network due to the increased number of devices and increasing volume of data. Therefore, networks need to be more efficient to deliver the expected increase in volume. Femtocells are identified as a crucial way to the delivery of the increased demands for heterogeneous networks in which macrocells work in combination with femtocells to provide coverage to offices, homes and enterprise. A simulation framework is used to implement the LTE-A system of macrocells with femtocell overlays. A comprehensive and easy to use Graphical User Interface has been set up with the desired two-tier network topology. It estimates the throughput and path loss of all femto and macro users for all operating frequencies of an LTE-A system using different modulation schemes. A series of tests are carried out using the simulator for a range of scenarios. The LTE-A system performance is analysed in terms of throughput and path loss in an urban environment taking into account the attenuation and penetration losses because of multipath fading. It is concluded that for any operating frequency of an LTE-A system with femtocell overlays, the throughput of a femto user is increased two times for 16QAM modulation scheme as compared to QPSK/4QAM. The throughput is increased to three times for 64QAM as compared to QPSK. It is also deduced that path loss is independent of the modulation scheme and increases with the increase of distance from its base station.

Keywords - Quadrature Amplitude Modulation, Femtocells, Main Base Station, Long Term Evolution – Advanced and Multipath Fading.

I. INTRODUCTION

There is a significant change in the mobile wireless technology in terms of reliable connectivity and fast data transmission with the advent of 4G network in mobile communication system. It has been estimated that in the next 15-20 years the volume of the mobile data traffic will increase to 1000 times than the present volume[1]. A propitious energy and cost efficient way to cope such an anticipated increase in demand of voice and data is the idea of dense deployment of small cells i.e. Femtocells. Femtocells are low power, less cost, short range and minimal effort cellular Base Stations (BSs) which are installed by consumers [2]. The advantages of deployment of femtocells are high data rates and efficient spectrum use [3], energy saving [4], money saving [5, 6], plug and play [7], less congestion to a Main Base Station (MBS) [5] and easy handoff [6]. They provide fast response data services and clear voice quality.

Femtocells increase the network capacity, for example, the materials used in the construction of large buildings resist wireless network signal access, however femtocells can be placed within these locations (inside these buildings) to provide excellent coverage and signal reception. The substantial growth in the wireless telecommunication networks results in a heterogeneous network with largemacrocells in combination with small cells working together to provide the network coverage [1].

It allows service providers to extend coverage indoors and achieve high capacity gain. It also helps to offload the data and reduce traffic on expensive macrocell networks[8].

Quadrature Amplitude Modulation (QAM) can refer to both analog and digital modulation scheme used for communication. QAM transmits 2 digital bit streams (or 2 analog message signals) by modulating the amplitudes of the carrier waves. It is carried out by using either Phase Shift Keying (PSK) or Amplitude Shift Keying (ASK) in digital modulation schemes or by Amplitude Modulation (AM) in analog modulation schemes. Quadrature Amplitude Modulation is primarily used for digital data modulation. The analog carrier signal in QAM is first modulated by a discrete signal. Therefore, the digital signal is converted to analog signal during modulation and the analog signal is converted back to digital signal during demodulation. QAM can double the bandwidth. This makes it more effective to be used in digital communication systems[9]. The types of QAM include 2. 4, 8, 16, 32, 64, 128, 256 and 1024 QAM. The other forms include circular, rectangular etc. QAM modulation provides more secure data, extra capacity, better compatibility with digital data services, increased bit packing ratio and higher quality of communication [10]. Indeed, the higher order data QAM carries more data but this would also cause reliability issues. If there is any distortion or insertion, higher order QAM becomes less effective and reliable because it delivers data with higher Bit Error Rate (BER) as compared to lower order QAM [9]. Considering the reliability and effectiveness of lower order QAM, QPSK, 16QAM and 64QAM are chosen for performance analysis and simulation of Long Term Evolution - Advanced (LTE-A) systems[12-14] having femtocell overlays under different multipath fading. Table 1 presents the recapitulation of the relation between no. of bits, symbol rate and bits per symbol of the different formats of QAM.

Multiplier	Modulation Scheme	Number of Bits	Bits per Symbol	Symbol Rate	Symbol Example
1x	On-Off Keying	2^1	1	1 x Bit Rate (BR)	0
2x	4QAM/QPSK/4PSK	2^2	2	1/2 x Bit Rate (BR)	01
4x	16QAM	2^3	4	1/4 x Bit Rate (BR)	0101
5x	32QAM	2^4	5	1/5 x Bit Rate (BR)	01010
6x	64QAM	2^5	6	1/6 x Bit Rate (BR)	010101
7x	128QAM	2^6	7	1/7 x Bit Rate (BR)	0101010
8x	256QAM	2^7	8	1/8 x Bit Rate (BR)	01010101
9x	512QAM	2^8	9	1/9 x Bit Rate (BR)	010101010
10x	1024QAM	2^9	10	1/10 x Bit Rate (BR)	0101010101

Table 1: Characteristics of different formats of QAM[9, 11]

From Table 1 it can be deduced that 64QAM provides 50% more bandwidth than 16QAM because it can transfer 6 bits/symbol whereas 16QAM can only transfer 4 bits/symbol. Similarly, 64QAM can provide 3 times or 300% more bandwidth than QPSK while 16QAM can provide 2 times or 200% more bandwidth than QPSK.

Femtocells can work together with macrocells in LTE-A systems using the same or different frequencies. A user friendly simulation framework[8]is used to analyse the performance of LTE-A systems having femtocell overlays using different modulation techniques (QPSK, 16QAM and 64QAM) under different multipath fading by taking walls into account. An environment is defined to test a scenario with a range of operating frequencies for this paper in which femtocells are working in conjunction with macrocells to create a two-tier mobile network. The simulation framework calculates the Throughput and Path Loss (PL) of all femto and macro users deployed in the scenario under different modulation schemes.

The remaining sections of the paper are organised as follows. The Section 2 explains the mathematical modelling of the design and its implementation in the simulation framework. Section 3 elaborates the simulator model and configuration followed by Section 4 which demonstrates the results of simulation and analyse the network performance. Finally, Section 5 completes the paper with a conclusion and an identification of future work.

II. NUMERICAL MODELING OF THE SIMULATION FRAMEWORK

In a two-tier system, the inappropriate deployment of femtocells may cause interference resulting in a decreased network capacity. Since the main focus is to analyse the performance of an LTE-A system with both macro and femtocells working together under different modulation schemes, the scenario is kept simple and constant to mitigate the effect of interference. The following section explains and analyses how the throughput in the LTE-A system integrating femtocells and macrocells is calculated and estimates the Path Loss (PL) of all femto and macro users connected to their respective femto or macro base stations with respect to their distance. It also takes into consideration the propagation models and losses due to attenuation and multipath fading.

Path Loss Model

The path loss between a MBS and macro user and also between a Femto Base Station (FBS) and femto user, in the same apartment (no walls between them) needs to be calculated.

The path loss of a macro user in an outdoor urban area in dB is calculated as [15]:

$$PL\ dB = 37.6 \log_{10} D + 15.3 \quad (1)$$

while in comparison the path loss of an indoor macro user in dB is calculated as:

$$PL\ dB = 37.6 \log_{10} D + 15.3 + LPOW \quad (2)$$

In the above equation, D denotes the separation between the transmitter and receiver (in meters) and LPOW represents the penetration loss of outdoor walls. The path loss of a FBS to femto user in dB is calculated as [15]:

$$PL\ dB = 20 \log_{10} D + 38.46 + 0.7d_{2D,indoor} + 18.3n^{((n+2)/(n+1)-0.46)} + LPiw \quad q \quad (3)$$

whereas n denotes the no. of penetrated floors, LPiw represents the penetration loss of the wall splitting the apartments and qrepresents the no. of walls dividing the apartments between FBS and femto user.

$0.7d_{2D,indoor}$ accounts for the penetration loss because of the walls inside an apartment calculated in meters.

In a case when an outdoor femto user makes connects with an indoor femto base station the outdoor wall loss has to be considered. The path loss in dB will then be measured according to [15]:

$$PL\ dB = \max 37.6 \log_{10} D + 15.3\ ,\ 20 \log_{10} D + 38.46 + 0.7d_{2D,indoor} + L_{Piw} \quad q + L_{POW} + 18.3n^{((n+2)/(n+1)-0.46)} \quad (4)$$

The Channel Gain (CG)yields different values for indoor and outdoor scenarios and is significantly affected by path loss. It is calculated as:

$$CG = 10^{-PL/10} \quad (5)$$

Throughput Computation

The total throughput of a serving macrocell is calculated according to the equation (6):

$$TM = \beta_{m,n} C_{m,n} \quad (6)$$

where $\beta_{m,n}$ denotes the subcarrier assignment. The subcarrier n is allocated to the macro user m if $\beta_{m,n} = 1$. Else $\beta_{m,n} = 0$. Each subcarrier is assigned to a macro user at a time for every time slot in a macrocell, because of the attributes of Orthogonal Frequency Division Multiple Access (OFDMA) used in LTE-A technology. This results in the following equation:

$$\sum_{m=1}^{Nm} \beta_{m,n} = 1 \quad (7)$$

N_m in the above equation represents the no. of macro users in a macrocell.

III. SIMULATOR MODEL AND CONFIGURATION

The simulationis performed according to the mathematical modelling explained in Section 2. The desired topology is executed in the simulation framework provided in [8]. The number of femto base stations, the number of macro users, the number of femto users and the number of walls in the x and y direction are fed as an input in the GUI. The purpose of using GUI is to allow flexibility of choosing the desired environment and running the simulation many times by only changing the QAM scheme and keeping the LTE-A environment constant to analyse the performance. Maximum throughput estimation and path loss of the femto and macro users are calculated for the defined test scenario consisting ofa macrocell environment integrated with femtocell overlay.

The inputs of the simulation framework are the properties of macrocell environment such as the dimensions of the interfering buildings, positions of the femto base stations, and positions of the femto users along with their distance to their respective femto base stations. It also includes the type of modulation scheme and the transmission powers of macro and femto base stations. According to the defined femto and macro base station transmission powers, numerical model presented in Section 2 and taking the White Gaussian Noise(WGN) into account, the simulation framework can calculate the throughput of all femto and macro users at any position given in the LTE-A network under examination. It also calculates the path loss of the macro and femto users. The GUI allows the user to interact with the system. The framework requires user inputs including the total number of macro and femto users, the number of femto base stations users are connected to and the precise location coordinates (x, y) of the femto base stations in a defined macrocell coverage region. The framework also needs the road width and the number of buildings in x and y axis so that it can generate the building map. As LTE-A operates on multiple configuration modes, type of modulation scheme and operating frequency parameters are necessary. The distance between users and their respective antennas are calculated according to the specified location coordinates. It results in the estimation of path loss. The location of base stations, users' location and the hexagonal layout of the map are the sufficient parameters to compute interfering walls and consequently calculate the path loss. Since the deployment of femto base stations are mostly restricted to urban area because of the need and population density in urban areas, this specific topology only considers the models of an urban area while calculating throughout and path loss.

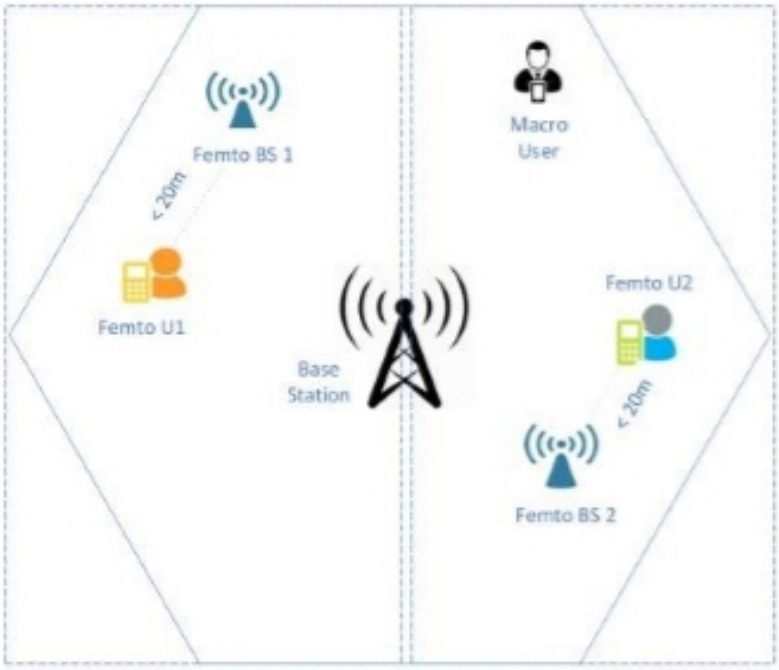

Fig. 1. Visual Representation of the test scenario

The visual representation of the scenario being tested is shown in Fig. 1. Since the base station antennas follow hexagonal coverage patterns to reduce black spots, the coverage area of the macrocell also displays hexagonal structure. A macrocell antenna is shown in the middle of the map. It is encircled by all deployed femto base stations and the femto and macro users attached to their base stations in their respective positions. Based on user defined positions, the walls of the buildings in x and y directions are also displayed on the topology's map. Fig. 1 depicts the visual representation of the scenario being tested.

IV. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

The proposed scenario as shown in Fig. 1 has been repeatedly run and simulated against different frequencies and modulation schemes. The metrics are calculated according to the equations described in Section 2 and outputted in real time. The values of the parameters used in the simulation framework are based on [16]and are stated in Table 2.

Parameter	Value
Femtocell Radius (m)	20
Macrocell Radius (m)	250
Carrier Frequency	2 GHz
Femto Base Station (FBS) Power (dBm)	20
Macro Base Station (MBS) Power (dBm)	46
Indoor Walls Loss (L_{Piw})	5 dB
Outdoor Walls Loss (L_{POW})	20 dB
White Noise Power Density	-174 dBm/Hz
Subcarrier Spacing	15kHz
Bandwidth (MHz)	1.4 3 5 10 15 20
Modulation Scheme	QPSK/4QAM/4PSK 16QAM 64QAM

Table 2: Parameters of Simulation Framework

To start with, the end user selects the number of femto base stations located within the macrocell coverage area, the no. of macro users, and the no. of femto users connected to every FBS. Then the user selects the channel bandwidth (20, 15, 10, 5, 3 or 1.4 MHz according to the existing LTE-A standards) and modulation scheme (64QAM, 16QAM or QPSK/4QAM/4PSK). Then the user selects the number of buildings along x axis and y axis. The user then defines the street width on the map because the urban environment is chosen in this simulation framework. The manual deployment of femto base stations and femto and macro users takes place by clicking on the points of user choice in the area of macrocell. As the last macro user is placed on the layout, the deployment is considered to be completed. Fig. 1 represents the visual representation of the placement of femto base stations, femto users attached to their respective femto base stations (the distance should be less than 20m as the femtocell base station radius is taken as 20m in this scenario), macro users and base station antenna. The framework also allows the end user to attach an additional macro or femto user to the already positioned femto base stations. For this particular scenario, an additional user is attached to Femto Base Station 1. Table 3 provides an overview of the user inputs used for the simulation framework.

User Input	Value
Num. of Femtocells	2
Num. of Macro Users (MUs)	1
Num. of Femto Users (FUs)	2
Num. of Buildings (x)	2
Num. of Buildings (y)	1
Road Width (m)	6
Additional User connected to	Femto BS 1

Table 3: User Inputs for Simulation Framework

The described scenario is repeated for all the operating frequencies of LTE-A systems (20, 15, 10, 5, 3 and 1.4 MHz) against each modulation scheme (64QAM, 16QAM and QPSK/4QAM/4PSK). The throughput and path loss of all femto users are calculated and the distance of femto users from their respective base stations is take into account. Table 4 presents the results of the simulation based on the user defined topology.

The throughput and path loss are calculated for each femto user according to the equations described above in Section 2. Since an additional user is also connected to Femto BS 1, therefore Femto BS 1 has to serve two users in its proximity whereas Femto BS 2 has to serve a single femto user. Therefore, calculated throughput of Femto U1 is always half than the throughput of Femto U2 for each scenario because available bandwidth to Femto U1 is shared with an additional user also served by the Femto BS 1. From the Table 4 it is also evident that 64QAM provides 3 times more bandwidth as compared to

QPSK or 4QAM against same frequency for each scenario. Table 4 also shows that 16 QAM provides twice the bandwidth as compared to QPSK/4QAM. 64QAM provides 1.5 times more bandwidth as compared to 16QAM as seen from the Table 4.

Frequency (MHz)	Modulation Scheme	Femto User 1 (Femto U1)			Femto User 2 (Femto U2)		
		Throughput (Mbps)	Path Loss (dB)	Distance (m)	Throughput (Mbps)	Path Loss (dB)	Distance (m)
20	64QAM	54	62.82	16.52	108	63.29	17.43
20	16QAM	36	63.07	17	72	63.05	16.97
20	4QAM/QPSK	18	62.61	16.12	36	61.98	15
15	64QAM	40.5	62.79	16.46	81	62.54	16
15	16QAM	27	62.42	15.78	54	63.1	17.05
15	4QAM/QPSK	13.5	62.26	15.49	27	62.1	15.20
10	64QAM	27	62.87	16.61	54	63.55	17.96
10	16QAM	18	62.54	16	36	63.29	17.43
10	4QAM/QPSK	9	62.12	15.23	18	63.56	17.99
5	64QAM	13.5	62.74	16.37	27	62.87	16.61
5	16QAM	9	62.33	15.62	18	63.11	17.08
5	4QAM/QPSK	4.5	63.11	17.08	9	63.49	17.85
3	64QAM	8.1	62.59	16.09	16.2	61.96	14.97
3	16QAM	5.4	62.59	16.09	10.8	58.31	9.83
3	4QAM/QPSK	2.7	57.58	9.04	5.4	57.08	8.53
1.4	64QAM	3.24	52.59	5.09	6.48	52.42	4.99
1.4	16QAM	2.16	55.69	7.27	4.32	52.42	4.99
1.4	4QAM/QPSK	1.08	41.45	1.41	2.16	50.75	4.11

The comparison of the throughput of Femto Users 1 and 2 is shown in Fig. 2 below. Femto U1 is denoted by a blue line whereas Femto U2 is denoted by a red line to clearly show the comparison of obtained throughputs of both the users against each modulation scheme for all the available frequencies of LTE-A systems with femtocell overlays. The path loss increases as the distance of femto users from their femto base stations increase. The path loss does not have a linear relationship with distance because of the complexity of the equation to compute path loss considering the indoor and outdoor penetration losses. It also takes into account the attenuation caused by multipath fading in an urban environment. From Table 4 it is confirmed that the path loss is independent of the modulation scheme. Fig. 3 illustrates the comparison of path loss (in dB) of femto users against their distances from the femto base stations for Femto U1 and Femto U2.

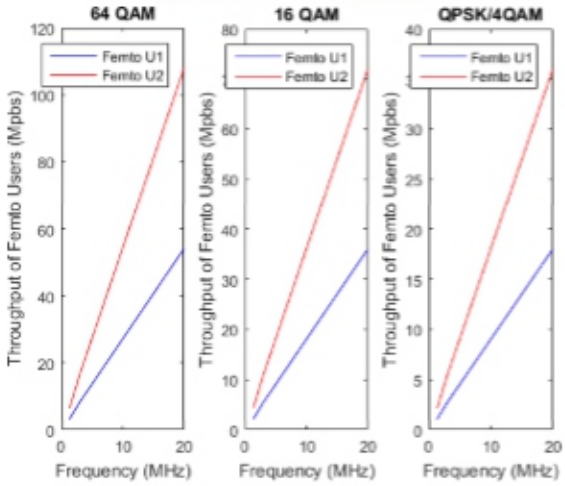

Fig.1. Throughput Comparison of Femto U1 and Femto U2 for 64QAM, 16QAM and QPSK/4QAM

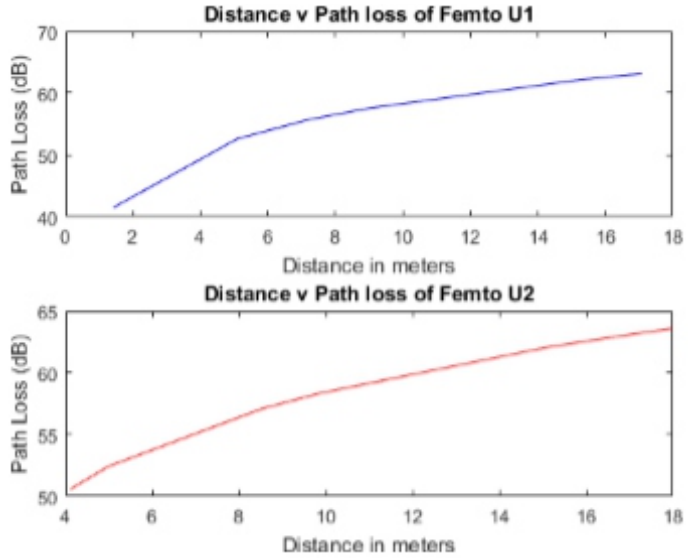

Fig.2. Path Loss Comparison of Femto U1 and Femto U2

V. CONCLUSION AND FUTURE WORK

The GUI of the simulation framework allows researchers to set up their desired two-tier network topology. It allows the calculation of throughput and path loss for all operating frequencies of an LTE-A system using different modulation schemes. A series of tests were carried out using the simulator to calculate maximum throughput and path loss for a range of scenarios. Same user topology was applied for each scenario. The bandwidth and modulation scheme was varied for each scenario to analyse the effect they have on the throughput of the femto users. The experiments were run several times and the simulation results were evaluated in terms of throughput and path loss. Hence it is concluded that for any bandwidth of an LTE-A system with femtocell overlays, the throughput of a femto user is increased two times for 16QAM modulation scheme as compared to QPSK/4QAM. The throughput increased to three times for 64QAM as compared to QPSK/4QAM. It is also concluded that path loss is independent of the modulation scheme. Path loss increases with the increase of distance of femto or macro user from its base station because of attenuation and multi path fading in an urban environment.

An interesting future work will be to keep the modulation scheme and operating frequency constant and then to change the position of femtocells inside the buildings. The singularities of femtocells, characteristics of an indoor environment and mobility pattern of indoor users will pose different challenges. In such an environment, a number of femtocells can be placed in different locations inside buildings. The optimised positioning of the femtocells can be determined on the basis of Throughput, Signal to Interference Noise Ratio (SINR) and path loss.

REFERENCE

[1] I. Chih-Lin, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: A 5G perspective," IEEE Communications Magazine, vol. 52, no. 2, pp. 66-73, 2014.

[2] X. Wang, A. V. Vasilakos, M. Chen, Y. Liu, and T. T. Kwon, "A survey of green mobile networks: Opportunities and challenges," Mobile Networks and Applications, vol. 17, no. 1, pp. 4-20, 2012.

[3] W. Wang and Q. Zhang, "Local cooperation architecture for self-healing femtocell networks," Wireless Communications, IEEE, vol. 21, no. 2, pp. 42-49, 2014.

[4] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: challenges and research advances," Network, IEEE, vol. 28, no. 6, pp. 6-11, 2014.

[5] H. ElSawy, E. Hossain, and D. I. Kim, "HetNets with cognitive small cells: user offloading and distributed channel access techniques," Communications Magazine, IEEE, vol. 51, no. 6, pp. 28-36, 2013.

[6] N. Bhushan et al., "Network densification: the dominant theme for wireless evolution into 5G," Communications Magazine, IEEE, vol. 52, no. 2, pp. 82-89, 2014.

[7] X. Zhou et al., "Toward 5G: when explosive bursts meet soft cloud," Network, IEEE, vol. 28, no. 6, pp. 12-17, 2014.

[8] C. Bouras, V. Kokkinos, K. Kontodimas, and A. Papazois, "A simulation framework for LTE-A systems with femtocell overlays," in Proceedings of the 7th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks, 2012, pp. 85-90: ACM.

[9] N. Voudoukis, "Performance Analysis, Characteristics, and Simulation of Digital QAM," European Journal of Electrical Engineering and Computer Science, vol. 1, no. 1, 2017.

[10] B. Sklar and F. J. Harris, Digital communications: fundamentals and applications. Prentice-hall Englewood Cliffs, NJ, 1988.

[11] J. Houghton, "Calculating QAM Bandwidth Gains," https://www.neophotonics.com/how-much-more-band width- 64-qam/ [Accessed 16 June 2019], 2017.

[12] L. Frenzel, "An Introduction to LTE-Advanced: The Real 4G," Electronic Design, vol. 8, 2013.

[13] D. Soldani and A. Manzalini, "Horizon 2020 and beyond: On the 5G operating system for a true digital society," IEEE Vehicular Technology Magazine, vol. 10, no. 1, pp. 32-42, 2015.

[14] R. Sood and A. Garg, "Digital society from 1G to 5G: a comparative study," International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 3, no. 2, pp. 186-193, 2014.

[15] J. Ikuno, M. Wrulich, and M. Rupp, "3GPP TR 36.814 V9. 0.0-Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects," Tech. Rep., 2010.

[16] M. Simsek, T. Akbudak, B. Zhao, and A. Czylwik, "An LTE- femtocell dynamic system level simulator," in 2010 International ITG workshop on smart antennas (WSA), 2010, pp. 66-71: IEEE.

ISSN (Print):2321–404X, (Online):2321–4384

EDGE Computing and Data Mining Techniques for Smart Healthcare Using Wearble DeviceS

[1] Sushrut Hiray, [2] T. P. Vaidya
[1]Student, Sinhgad College of Engineering, Pune
[2]Assistant Professor ,Department of Computer Engineering, Sinhgad College of Engineering, Pune
E-mail: [1]hiray.sushrut6@gmail.com, [2]tpvaidya.scoe@sinhgad.edu

ABSTRACT

This paper presents use of various data mining techniques and use of wearable devices for healthcare. Devices mainly gather the real time data from patient and can store on cloud or for limited time on the device. In this paper we are mainly focusing on various data mining techniques used in decision making and also in emergency call generation. We mainly focused on use of wearable device such as 'Fit band' for data collection such as Pulse, Sleep ,pattern , temperature. Collected data can be stored on local memory for one or two days. We can apply some data analytics techniques such as association rules, regressions and statistical methods .We are mainly introducing new concept of 'Fog computing' which will give us effective computation speed and fast access to patient .This techniques will reduce the data traffic on smart city network and Emergency can be handled in more precise way.

Keywords - Wearable Devices, Association Rule, Data Analytics Techniques, Fog Computing, Smart Healthcare.

I. INTRODUCTION

The global population continues to grow at a steady pace, and more people are moving to cities every single day. Experts predict that world's urban population will double by 2050 , which means we are adding the equivalent of seven New Delhi Cities to the planet every single year. Urbanization accompanies economic development. Urban areas also contribute a higher share of the GDP. This trend of urbanization that is seen in India over the last few decades will continue for some more time. It is for this reason that we need to plan our urban areas well and cannot wait any longer to do so.

An introduction to ubiquitous computing has tremendously improved the techniques and practices in Health industry which is one of the major economic and social problems around the world and it costs tremendous health expenses and resources.

The world is advancing day by day, in the system of healthcare and using the advanced technology today, data is collected from various sources . It can be clinical, e-health prescription, test reports, records and records of the test images. The main obstacle comes in extracting the data and providing helpful knowledge and information to medical researchers and practitioners. These data are extracted and used in the scenario of real life giving profit to the common people. If Big Data Analytics is applied in the field

of healthcare than using the historical data, the diseases can be predicted especially the epidemic diseases. The quality of life is improved by adverse reaction of the drug obtained from social media and helps in avoiding preventing disease.

This opens remarkable opportunities to reduce the costs of health care as well as diagnosing the diseases in a much simpler way. Currently, healthcare domain is to offer better for people anytime and anywhere in the world in a more Profitable and patient friendly way. Accurate and exact prediction of the heart disease mainly depends on Electrocardiogram (ECG) data and patient clinical data. With a more health-conscious population, coupled with advancements in medical and digital technologies — it is no wonder that wearable sensors such as fitness trackers and smart watches are becoming common. Most often used for tracking physical activities, some wearable sensors can even track calories burnt, sleeping patterns and heart rate, keeping tabs on general health and wellness. It cannot be easily predicted by the medical practitioners as it is a difficult task which demands expertise and higher knowledge for prediction. An automated system in medical diagnosis would enhance medical efficiency and also reduce costs. We will design a system that uses various data analytic techniques which can efficiently discover the severity and the risk level of patients based on the given parameters about their health and then using wearable devices for targeting high risk patients only. Main focus of our system is enhancing more critical patient care by use of 24 by 7 monitoring of patient by Fit Band which is easily available low cost hardware.

II. RELATED WORK

Data mining techniques are used to Data mining techniques are used to analyze huge collection of data from different perspectives and deriving useful information. Various data mining techniques normally used related to Bigdata analytics we can say are Classification, Clustering, Association Rule mining, Regression. In paper[1] discussed various data mining techniques for heart disease predictions. For the evaluation of the interaction between the survival of the patient and the interaction between the physiological parameters several data preprocessing, mining and transformation approaches are used for the determination of knowledge about these interactions. The prediction of the survival of the patients is exhibited by decision making algorithms and the rules helped in making these decisions.

These decisions play a significant role in the survival of the new patients which are unseen. All the significant parameters were identified and interpreted using the process of data mining and were reported related to the medical significance. A concept was introduced by the author [2] in their research, the approach used decreases the cost and effort of the selection of patients for the clinical studies. The significant parameters are discovered and on the basis of the prediction of results patients can be selected

[2]. By the help of the artificial neural networks, using the analysis of pattern electroretinography signals the optic nerve disease can be diagnosed. The complete technology used was ANN was forward with multilayer feed along with Levenberg Marquart (LM) back propagation algorithm. There was an efficient and effective interpretation which was made using this PERG method and the classification of the final results amongst the healthy and the diseased could be performed easily [3].According to Mark Weiser, his vision is also true for the healthcare industry regarding the ubiquitous computing and smart objects. The law of Moore can also be applied to the biomedical sensors. This law is applicable to more devices above imagination. There vision was the interaction of people amongst both the cyber space and physical space. The main power of such cyber-physical system is intelligence, i.e. the smartness lying in their adaptive behavior [4-7].

Another algorithm used was a technique of machine learning named Support Vector Machine and Random Forest which were used to classify compare and study of the data sets of the cancer, liver and the heart diseases, with varying the kernel and its parameters. The results were obtained from various data sets of breast cancer, liver and the heart diseases. The results were further merged with different kernels for the selection of proper parameter. For the establishment of better learning techniques regarding the prediction of the diseases results obtained were better analyzed. Altering results were observed with the classification of SVM techniques using different functions of kernel [3].

Further two neural techniques were used along with the BPA, RBF, SVM etc and compared for better efficiency and accuracy. The BPA being the Back propagation algorithm, RBF being Radial basis Function, SVM being Support Vector Machine is one non-linear classifier. For the determination of kidney stone diagnosis, WEKA 3.6.5 tool was implemented to determine the best and better technique amongst the three algorithms used. The main aim of their work was to exhibit the best tool for the diagnosis of diseases, reduce its time of treatment and efficiency and accuracy were improved regarding the disease of kidney stone identification.

This paper, a new unsupervised classification system is adopted for heart attack prediction at the early stage using the patient's medical record. The information in the patient record are pre processed initially using data mining techniques and then the attributes are classified using a Fuzzy C means classifier. In paper [8] conceptual flow of healthcare system that stores the EHR is discussed. In this paper authors discussed three type of decision support: 1. Visualization 2.

Alerts. 3. Classification. For larger database classification is ok, but for critical patients and less data attributes alert and visualization is more required. 15 attributes are related with predicting heart attract

collecting its real time data we will monitor the status of highly severe patients. By applying association rule and generating specific rules we are suggesting some health changes and then recommending a doctor based on location, health condition and economic status. An alert will also be sent to three of the patients emergency contacts.

IV. METHODOLOGIES AND ALGORITHMS Y

We have already published one paper in ICRMR2019 Goa Conference with Apriori algorithm. In the proposed strategy after getting the severity level of the patient we are starting with second module. In this after collecting the data from fit band for 24 hours by taking average count per hour and this list will give us data which is then analyzed using FP growth algorithm.

which in turn makes the candidate list from dataset and after reducing the number of candidates at each iteration based on a certain value of confidence it gives the final candidate list. The stronger the rules the better the result. Support, confidence and association rules are defined below.

1. Support:

The support of an item set, is the proportion of transaction in the database in which the item X (In our case the characteristic) appears.

2. Confidence:

Confidence of a rule is defined as follows:

It signifies the likelihood of item Y being possible when item X is possible. In our case let us assume that if blood sugar is high and exercise is low then there is a possibility of a heart disease

3. Association Rules Using association rules, a pattern is discovered based on a relationship of a particular item on other items in the same transaction. As one of the example where it is used, the association technique is used in heart disease prediction as it is used to tell us the relationship of different attributes used for analysis and sort out the patient with all the risk factor which are required for prediction of disease as is done in this.

and data mining technique like ANN, time series, clustering and association rules can be approached for this prediction (Soni& Sharma, 2011) [3]. Using of sensor devices for physiological monitoring of vital signs in healthcare services.

Banaee, Ahmed, &Loutfi, (2013) reviewed the latest methods and algorithms used to analyze data coming from wearable sensor devices [9]. They used some common data mining technique such as anomaly detection, prediction and decision making when considering in particular continuous time series measurements.

From literature review we concluded that not only classification or few data mining techniques make precise disease prediction system but 24 by 7 monitoring of patient in daily activities is essential Hence by fusing data mining techniques with wearable technologies we are trying to implement more precise " SMARTHEALTHCARE" .

III. BASIC ARCHITECTURAL DESIGN

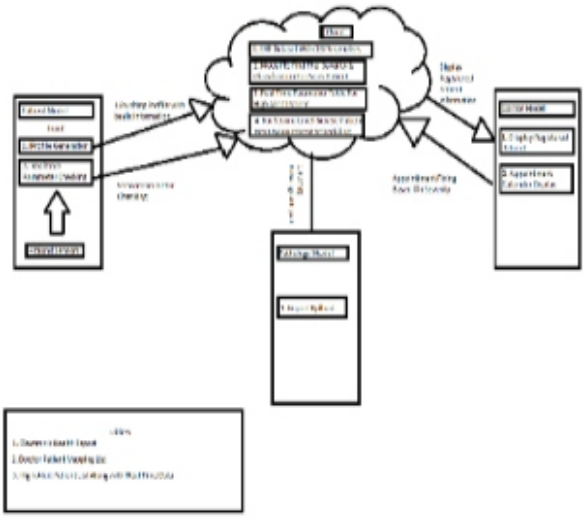

Fig. 3.1 Architectural Diagram

In this system we have made two main modules. In first module it is normal Machine learning classifier we have used to predict heart disease severity . Then use of wearable device we are doing after that. It is a mobile based machine learning application which is trai ed by a UCI dataset and real time data. For first time users we are collecting input from lab technician and hospitals. The user will register to our system resulting in generation of specific id and password for that user. Based on this id the doctor can view patient history and access the EHR of the user. The patient history can is stored on cloud by using time stamping which will help the doctor to get patient history in detail and easily. Based on the data collected and stored severity of the patient is being calculated using classification algorithm. By using fitband and

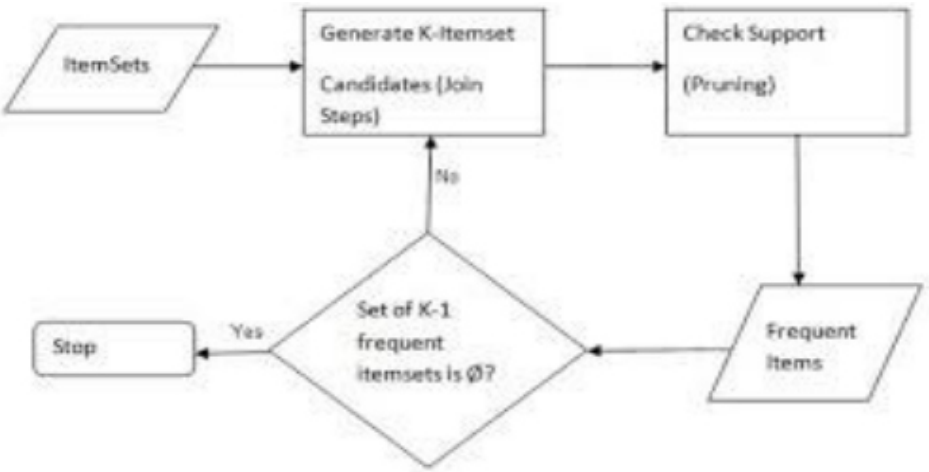

Figure 2: Association Rule Mining

V. SIMULATIONS AND EXPERIMENTAL RESULTS

The proposed solutions have been designed using Fitband. We are using Javascript to get the patients data in our application. We used Fitbit SDK

In user guide you will get code to take the data from fitband . We are using Pulse, BP, sleeping pattern to get the association rule. The Functional Requirement for the Application is to get the exact situation of patient and doing the recommendation as per rule.

User interface of this program is the common, a mobile application with section for filling required medical details for prediction purpose and relevant suggestions. HELP button will appear on screen which is a documentation component of a software program that explains the features of the program and helps the user understand its capabilities and also a video tutorial for better understanding. In our previous paper we discussed mathematical model in detail. Here we want to focus on use of FP – Growth algorithm for Association rule generation.

FP growth algorithmic program is an efficient algorithm for producing the frequent item sets without generation of candidate item sets. It adopts a divide and conquer strategy and it needs two database scans to seek out the Support count. It can mine the items by using lift, leverage and conviction by specifying minimum threshold.

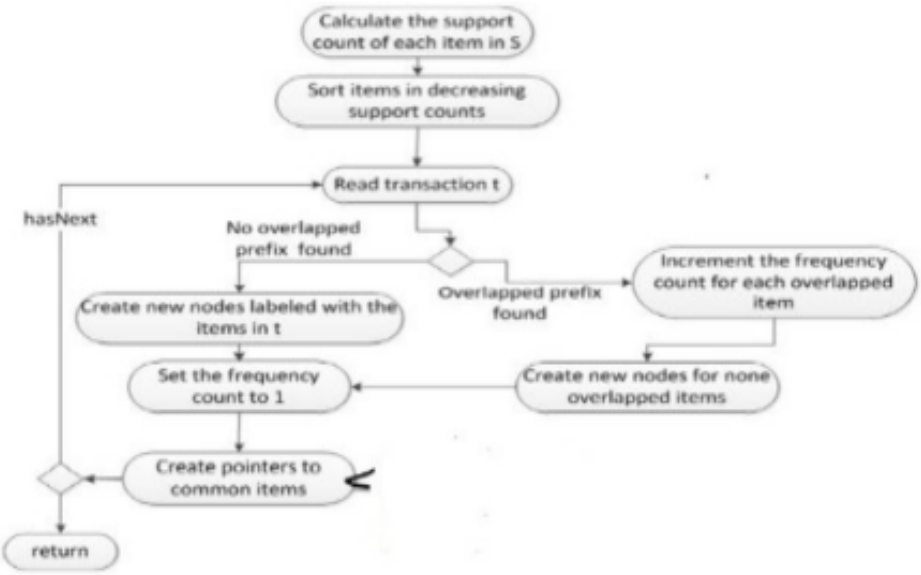

Fig.2 Activity Diagram for FP-Growth

Python implementation for FP growth is as follows: #variables: #name of the node, a count #nodelink used to link similar items #parent vaiable used to refer to the parent of the node in the tree #node contains an empty dictionary for the children in the node class tree Node: def init (self, nameValue, numOccur, parentNode): self.name = nameValue self.count = numOccur self.nodeLink = None self.parent = parentNode #needs to be updated self.children = {} #increments the count variable with a given amount def inc(self, numOccur): self.count += numOccur #display tree in text. Useful for debugging def disp(self, ind=1): print (' '*ind, self.name, ' ', self.count) for child in self.children.values(): child.disp(ind+1)

VI. CONCLUSION

Main aim of paper is leaverging cloud computing environment in Smart health care and making use of Edge computing. Edge computing is the technology where we can use the machine learning techniques on the smart devices . Wearble devices mainly now a days are with sufficient amount of memory and computing power. We can easily implement Association Rule Algorithms on these devices which in turn can be effectively used in recommendation systems and Emergencies can be handled.

REFERENCES

[1] Keyue Ding and Kent R Bailey et al, "Genotypeinformed estimation of risk of coronary heart disease based on genome- wideassociation data linked to the electronic medical record", International journal of BMC cardiovascular Disorders, Vol.11,2011

[2] AndrewKusiak, Bradley Dixonb, Shital Shaha, (2005) "Predicting survival time for kidney dialysis patients: a data mining approach", Elsevier Publication, Computers in Biology and Medicine ,Vol.35, pp 311–327

[3] Sadik Kara, AysegulGuvenb, Ayse OztUrk Onerc, (2006) "Utilization of artificial neural networks in the diagnosis of optic nerve diseases", Elsevier Publication, Computers in Biology and Medicine,Vol. 36, pp 428–437.

[4] Weiser, M., Gold, R., Brown, J.S.: The origins of ubiquitous computing research at PARC in the late 1980s. IBMSYSTEMSJOURNAL 38(693-696 (1999)

[5] Moore, G.E.: Cramming More Components Onto Integrated Circuits. Electronics 38(8), 114-117 (1965)

[6] Wu, F.J., Kao, Y.F., Tseng, Y.C.: From wireless sensor networks towards cyber physical systems. Pervasive and Mobile Computing 7(4), 397-413 (2011).

[7] Daisy JA Janssen and Emiel FM Wouterset al, "Selfperceived symptoms and care needs of patients with severe to very severechronic obstructive pulmonary disease, congestive heart failure or chronic renal failure and its consequences for their closest relatives:the research protocol", Journal of BMC palliative care, Vol.7, 2008

[8] Shou-En Lu and Gloria L Beckleset al, "Evaluation of risk equations for prediction of short-term coronary heart disease events inpatients with long-standing type 2 diabetes: the Translating Research into Action for Diabetes", International J

[9] M.A.Jabbar,"Heart Disease Prediction System using Associative Classification and Genetic Algorithm", ICECIT, pp 183-192, Elsevier, vol 1(2012)

[10] M.A.Jabbar et.al" Prediction of heart disease using Random forest and Feature subset selection ",AISC SPRINGER, vol 424,pp187-196(2015)

[11] M.A.Jabbar, " Classification of heart disease using artificial neural network and feature subset selection", GJCST, Volume 13 Issue 3 , Ver 1.0, 15-25,(2013)

[12] Sun J., Reddy C.K. Big Data Analytics for Healthcare. Tutorial presentation at the SIAM International Conference on Data Mining, Austin, TX, 2013.

ISSN (Print):2321–404X, (Online):2321–4384

Automated Fire Fighting Robot

[1] Monisha. A, [2] Prabha. R, [3] Suruthi. V. P, [4] Swathi Suriyanarayanan, [5] M. Aasha
[1,2,3,4] Ug Scholars,
[5]Assistant Professor Department of Computer Science And Engineering Avinashilingam Institute For Home Science And Higher Education For Women. Coimbatore, India
E-mail: [1]monishaandy06@gmail.com, [2]prabha97raju@gmail.com, [3]suruthiprakasan@gmail.com, [4]saiswathi9707@gmail.com, [5]arathil800@gmail.com

ABSTRACT

Fire-fighting robot works on behalf of humans in hazardous area. This system is used to protect human lives and surroundings from fire mishaps. The objective of this work is to develop a fire fighting robot which aids the society to protect them from fire accidents. The robot helps in emancipation and helps people by suggesting emergency medical assistance to the person suffering from injuries caused during fire accidents.

Keywords - Dousing, Centrifugal pump, Thermocouple, SSH Interface, Sensors.

I. INTRODUCTION

A robot is needed to help people in dangerous situations. The fundamental requirement of fire fighting robot is to douse fire during fire mishaps replacing the fire fighters. Fire mischances have turned out to be normal in the recent days and in some cases, lifesaving turns into a troublesome task for the fire-fighters who are locked in. The absence of some essential fire-fighting abilities among the general population leads to the issue of an enormous catastrophe. In such cases, a fire-fighting robot can protect numerous casualties in the place of catastrophe. This robot can be implemented for shielding firefighters from extraordinary peril in petrochemical, synthetic unsafe items, lethality or during exploder fire mischance. The study of this work is to design and test a robot capable of dowsing fire by effectively replacing a firefighter during fire accidents. The implementation of this system will increase the safety of fire fighters and therefore help in mitigating unsafe conditions. This system mainly focuses on detecting fire in disaster-prone areas and thereby extinguishes fire on detection. This system provides visual indications and suggests first-aid to injured people. The rest of the paper is structured as follows. Section 2 discusses about the literature review related to the area of firefighting robots based on the proposed system.

II. RELATED WORK

Hisato Ando et al.[1] (2017) submitted a paper on "Aerial Hose Type Robot by Water-Jet for Fire Fighting" where they designed a prototype to validate the feasibility of the amount of water required to extinguish a fire, and then evaluate the reaction force and its effect on the robot.

B. Swetha Sampath [2](2011) presented a paper on "Automatic Fire Extinguishing Robot" which is a hardware-based model and is designed to automatically extinguish fire during fire accidents.

The robot moves in the direction of fire depending on its intensity. Calcium Silicate boards are used to coat the surface of the robot which can withstand very high temperature. The temperature detecting capacity of the robot is changed by warming the Thermocouple closures to a cut-off temperature, above which the robot begins reacting to the fire. The robot is used in rescuing operations during fire accidents where the possibility for service men to enter the fire prone areas is very less. The advantage of the robot is that it can turn ON automatically using the thermocouple when it detects the fire.

Teh Nam Khoonet al.[3](2012) presented a paper on "Autonomous Fire Fighting Mobile Platform" which illustrates the advancement work depending on the stage and plausibility of being a self-governing unit to screen a recommended territory, identify for flame and smother the fire. At the point when the fire sources being recognized the fire will be speedily quenched utilizing fire smothering framework that is mounted on its stage to distinguish the fire source the contribution from fire sensors are finely tuned in connection to the encompassing zone.

Sahil S.Shahet al.[4] (2013) presented a paper on "Fire Fighting Robot" that has been designed using embedded systems to stimulate household fire. It is capable of migrating through a modelled floor and is has the ability to scan the flame. The robot can even go about as a way guider in ordinary case and as a fire quencher in crisis. Robot douses the fire before raging out of control.

Chee Fai Tan[5] (2013) presented a paper on "Fire Fighting Mobile Robot" which is a remote-controlled robot to replace the fire fighters and reduce the risk of working in hazardous area and perform firefighting task. Different types of currently available fire fighting machines such as LUF60, FIREROB, FFR-1, FIREMOTE 4800, MVF-5, JMX-LT50, SACI2.0, ArchiBot-M, ThermiteT2, MyBOT2000 are described and compared with the present fire fighting robot technology.

AlHaza Tet al.[6](2015) published a paper on "New Concept for Indoor Fighting Robot" that has been designed for indoor fire fighting. This robot has the ability to climb stairs and also negotiate with several types of flow materials. Outer surface of the robot is coated with Ag to withstand temperature up to 700o C for 60 minutes. Multiple thermal insulation technique has been used for this purpose. It can send videos and audio information to control unit by communicating with trapped and injured victims inside the buildings. Ahmed Hassaneinet al.[7] (2015) submitted a paper on "An Autonomous Fire Fighting Robot" to locate and extinguish fire in the given environment. The robot migrates to the scene of activity

and avoids obstacles it faces in its excursion. This is mainly because they can be used in dangerous situation s for an individual to involve themselves in it.

Kadam.Ket al.[8](2018) submitted a paper on "Fire Fighting Robot"designed robot which extinguishes fire using remote control. These robots are for the enterprises where possibility of unplanned fire accidents take place. The proposed vehicle can identify the flame when it goes closer to it and stifling it consequently by utilizing gas sensor and temperature sensor. It includes equip engines and engine driver to control the development of robot. Transfer circuit controls the pump when it identifies fire at that point, it states with microcontroller (Arduino UNO R3) through a Bluetooth module. The proposed robot consists of a water source to sprinkle water. It will give GUI to Arduino task utilizing android. It might confront obstructions while distinguishing fire, at that point it has deterrent staying away from capacity. It distinguishes snags utilizing ultrasonic sensors up to scope of 80 m.

III. METHODOLOGY

An Arduino based algorithm is used to detect fire and measure distance from fire source while the robot is on its way to douse fire. When the fire is detected and the robot is at a distance close to the fire, a centrifugal pump throws water for extinguishing fire. A water spreader is used for effective extinguishing. It is noted that the velocity of water is reduced on using the water spreader in a process.

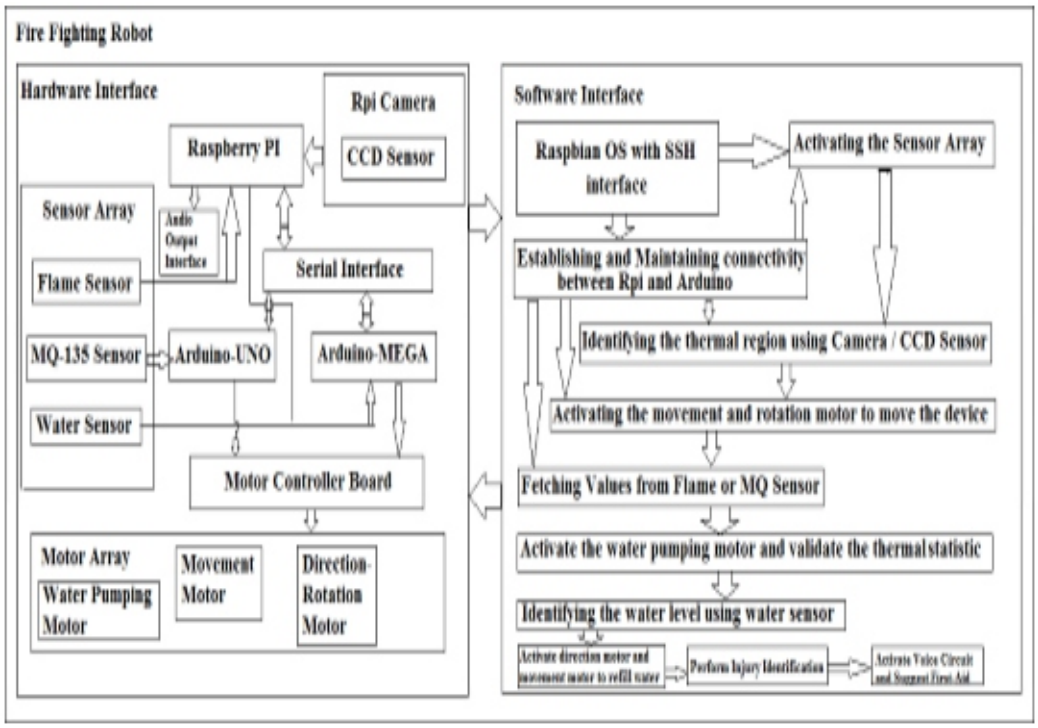

Fig 1. Hardware and Software Interface

Fig 1.illustrates to setup a firefighting robot the construction process is divided into two types, hardware interface and software interface. The software interface is a primary interface which consist of Raspbian OS with SSH interface. Secured shell interface (SSH interface) sets up the local network and wireless connectivity. It is also known as network interface. SSH interface controls the Raspberry pi via connection. To run the software, the hardware connections are made simultaneously. In hardware interface Raspberry pi stretch is used and it is connected with RPI (Raspberry Pi) camera, Sensors, Audio Output interface(speaker) and serial interface. Serial interface is used to connect Arduino with Raspberry pi. RPI camera consist of 20 megapixel which is used to capture the video of the surrounding of about 30meters.Sensor array which is connected in raspberry pi consist of three different sensors namely flame sensor, MQ 135 sensor and water sensor. Flame sensor is used to identify the sensitivity of the flame. MQ Sensor is composed of three sensors. They are gas sensor, flame sensor and smoke sensor. The flame sensor present in MQ 135 is used to check the concentration of the flame. Water sensor is used to pump water for dousing the fire. The Arduino is used to construct fire fighting robot is Arduino UNO and Arduino MEGA .12 DC Motors are fixed in Arduino which is controlled by Micro Controller board. These motors help in various motions like movement, rotation and pumping water. At software interface side by hardware connection it maintains and establishes the connectivity between RPI and Arduino and it identifies the thermal region using camera which is fixed in Raspberry pi. Based on the activities motors are fixed in order to activate the movement and rotation. Once the flame is identified using flame sensor the values are fetched. The raspberry pi sends the commands and the Arduino receives those command and response accordingly. Thereby it activates the water pumping motor and validate the thermal static. Once the water level is decreased the water sensor identifies the water level and activates direction to refill motor. RPI camera also performs identification of the person and suggests first aid through audio output interface.

IV. INNOVATION

Fire fighting robot is used in disaster prone areas where fire fighters find difficulties in rescuing people caught in fire accident. In order to overcome this kind of problems firefighting robots are proposed for the welfare of the society. During a fire extinguishing activity, it is recommended to emit water directly on the fire source by fire fighters positioned away from the fire. However, it is difficult to directly access the fire source.

The main motto of the robot is to directly access the fire source on behalf of firefighters and rescue the people caught in fire accident. Thus, Mobile fire fighting robots are introduced to detect fire and safeguard the people from harm.

V. CONCLUSION

Firefighting robot can be used in place of humans reducing the risk of life of the firefighter. Therefore, it can be used in areas where fire accidents are caused more predominantly. They provide greater efficiency to detect the flame and it helps in extinguishing and also suggests first-aid in case of harm. Robot detects temperature, smoke and flame at the spot where the robot exists. This robot is helpful in areas where natural calamity and bomb explosions occur.

REFERENCES

[1] Ando, H., Ambe, Y., Ishii, A., Konyo, M., Tadakuma, K., Maruyama, S., & Tadokoro, S. (2018). Aerial Hose Type Robot by Water Jet for Fire Fighting. IEEE Robotics and Automation Letters, 3(2), 1128-1135.

[2] Sampath, B. S. (2011, November). Automatic fire extinguisher robot. In Ubiquitous Robots and Ambient Intelligence (URAI), 2011 8th International Conference on (pp. 215-218). IEEE.

[3] Khoon, T. N., Sebastian, P., & Saman, A. B. S. (2012). Autonomous fire fighting mobile platform. Procedia Engineering, 41, 1145-1153.

[4] Shah, S. S., Shah, V. K., Mamtora, P., & Hapani, M. (2013). Fire fighting robot. Int. J. Emerg. Trends Technol. Comp. Appl, 2(4), 232-234.

[5] Tan, C. F., Liew, S. M., Alkahari, M. R., Ranjit, S. S. S., Said, M. R., Chen, W., & Sivakumar, D. (2013). Fire fighting mobile robot: state of the art and recent development.

[6] AlHaza, T., Alsadoon, A., Alhusinan, Z., Jarwali, M., & Alsaif, K. (2015). New Concept for Indoor Fire Fighting Robot. Procedia-Social and Behavioral Sciences, 195, 2343- 2352.

[7] Hassanein, A., Elhawary, M., Jaber, N., & El-Abd, M. (2015, July). An autonomous firefighting robot. In Advanced Robotics (ICAR), 2015 International Conference on (pp. 530- 535). IEEE.

[8] Kadam, K., Bidkar, A., Pimpale, V., Doke, D., & Patil, R. (2018). Fire Fighting Robot. International Journal Of Engineering And Computer Science, 7(01), 23383-23485.

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).

2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);

3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);

4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Address of the Editorial Office:

Enriched Publications Pvt. Ltd.
S-9,IInd FLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR -5, DWARKA, NEW DELHI, INDIA-110075,
PHONE: - + (91)-(11)-45525005