

# Health Care Monitoring for the CVD Detection using Soft Computing Techniques

N.Sathiya Rani<sup>1\*</sup>, K.Vimala<sup>2</sup>, Dr.V.Kalaivani<sup>3</sup>

<sup>1</sup>PG Scholar, <sup>2</sup>Assistant Professor, <sup>3</sup>Associate Professor,  
Department of Computer Science and Engineering ( PG),  
National Engineering College, Kovilpatti, India.  
n.sathiya90@gmail.com<sup>1</sup>, vimala.arivalagan@gmail.com<sup>2</sup>,  
<sup>3</sup>kalai.engr@rediffmail.com

## ABSTRACT

*Now-a-days, many diseases are reducing the life time of the human. One of the major diseases is cardiovascular disease (CVD). It has become very common perhaps because of increasingly busy lifestyles. The rapid development of mobile communication technologies offers innumerable opportunities for the development of software and hardware applications for remote monitoring of cardiac disease. Compressed ECG is used for fast and efficient. Before performing the diagnosis, the compressed ECG must be decompressed for conventional ECG diagnosis algorithm. This decompression introduces unnecessary delay. In this paper, we introduce advanced data mining technique to detect cardiac abnormalities from the compressed ECG using real time classification of CVD. When the patient affect cardiac disease, at the time hospital server can automatically inform to patient via email/SMS based on the real time CVD classification. Our proposed system initially uses the data mining technique, such as Genetic algorithm for attribute selection and Expectation Maximization based clustering. In this technique are used to identify the disease from compressed ECG with the help of telecardiology diagnosis system.*

## KEYWORDS

*Cardiovascular disease (CVD), Data mining, Genetic algorithm, EM based cluster, Classification, Telecardiology.*

## 1. INTRODUCTION

Recent trends and developments in wireless technologies, mobile phones are play an important role to become an effective tool for cardiovascular monitoring. The Electro Cardio Gram (ECG) plays a key role in cardiac patient monitoring and diagnosis. In the industrialized world, it is estimated that millions of people die due to various cardiac heart disease annually. The key to treat these diseases is timely detection. Since ECG is the most commonly recorded signal in the patient monitoring [2] and examination Process, it becomes important to be able to reliably and quickly detect cardiac diseases from ECG analysis.

In the medical mining, huge amount of data generated by healthcare transactions are too complex and voluminous to be processed and analysed by traditional methods. Data mining provides methodology and technology to transform these data in to useful information for decision making.

The main aim of healthcare management, the data mining applications can be developed to better identify and track chronic disease states and high risk patients, design the appropriate interventions, and reduce the number of hospital admission and claims. Data mining applications are developed to evaluate the effectiveness of medical treatments. By the comparing and contrasting causes, symptoms and courses of the treatments, data mining can deliver an analysis of which treatments is best for particular diseases and cost effective. Remote Patient Monitoring (RPM) refers to a wide variety of technologies designed to manage and monitor a range of health conditions. Data security and optimization of wireless communication between devices of the system [12] are very important in the home monitoring. In the existing system compressed ECG must be decompressed before performing for every ECG packet introduces unnecessary delay and patient's mobile phone only received the analog signal from health care monitoring device to hospital server in continuous manner. However, this continuous mode of data transfer generates network traffic and causes internet expensive and computational burden on the hospital server.

Our proposed system plays an important role in reducing the delay and provides a fully automatic monitoring facility for the cardiac patient who requires on-going monitoring facility. The aim of this approach is to detect cardiac abnormalities from the compressed ECG using real time classification of CVD. Data mining techniques can be placed with in wireless monitoring facility to alert the patient. This technique running on the hospital server generates a set of constraints for representing each abnormality. Fig 1 depicts the proposed system where the attribute selection using genetic algorithm and EM based clustering are applied first to retrieve a set of constraints within the hospital server. The patient's mobile phone later uses these set of constraints (e.g., ventricular flutter/ fibrillation, premature ventricular contraction, atrial fibrillation, etc.) to classify abnormal beats in real time. With the help of real time classification, emergency message are informed automatically to the patient's mobile phone whenever a life-threatening cardiac abnormality is detected. Some of the related researches are explain in Section 2. Attribute selection using genetic algorithm and EM based clustering are provided in Section 3.

## **2. RELATED WORKS**

There are many researches are going in analyzing ECG signals in national and international level. For example, the telemedicine facilities have been developed for diagnosis through the Internet. Usually attribute selection [14] is done for selecting the features from the compressed ECG. Expected Maximization (EM) [10] clustering technique is used to create normal and abnormal ECG clusters. The threshold value of ECG signal determination is done using Wavelet transform coefficients [11].

For Remote Health Care monitoring, the compression of ECG signals are essential. For compression new wavelet-based [1] ECG compression technique is implemented and this technique is tested using several records taken from the MIT-BIH Arrhythmia database. Also, wavelet transform method is used to achieve high compression ratios (CR) and a low percent root mean square difference (PRD). Using smartphone-based wearable CVD-detection platform, real-time ECG acquisition and display, feature extraction, and beat classification are done.

Noise is present usually in ECG pattern and the noise reduction process is an important task in ECG signals. Such noises are difficult to remove using typical filtering procedures. So many noise reduction technique is developed in international level such as efficient analytical tool [5] which is a technique for averaging of cardiac cycles which is increasing signal to noise ratio.

Usually the parameters are given as the inputs from medical practitioner directly on to the mobile phone. The score [10] is calculated by miniature java based software running inside the mobile phone. Based on the score, level of urgency is determined by the intelligent program. At the end, specialists are contact automatically by messaging services. Moreover, the results of the scoring are transmitted to the hospital server.

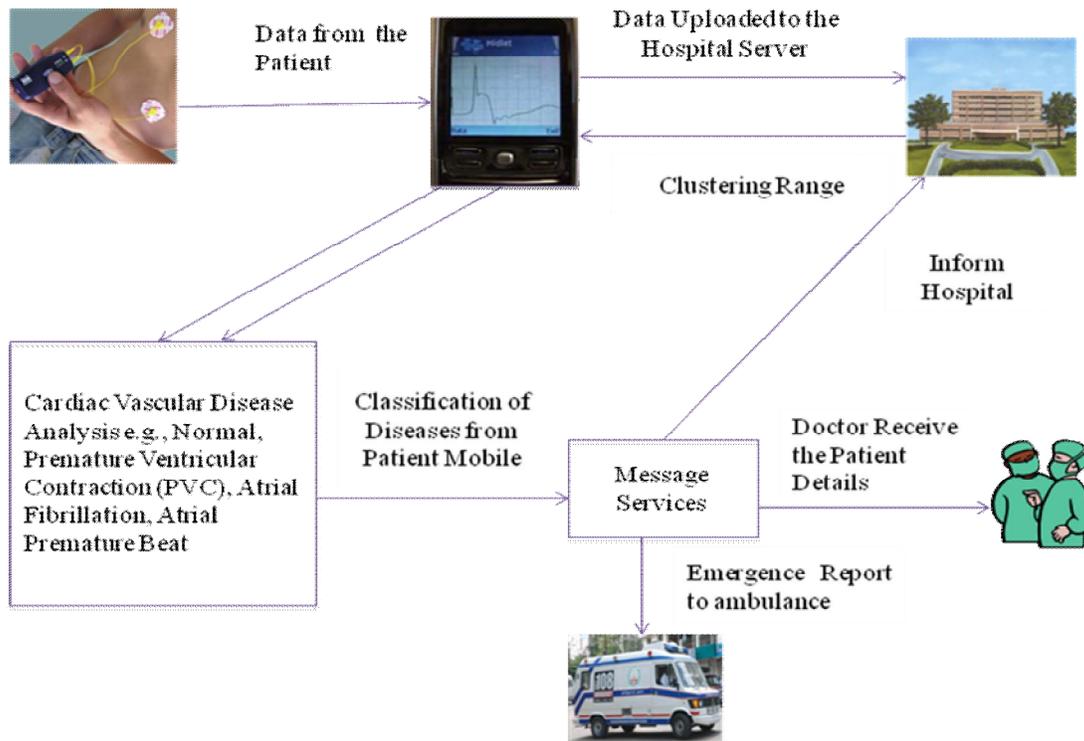


Fig 1: Cardiovascular Abnormality Diagnosis Process

Prototype Intelligent Heart Disease Prediction System (IHDP) [14] are developed in national level using data mining technique such as Decision trees, Navie Bayes, Neural Network. Also, biometric systems have been developed using ECE using data mining (DM) techniques like attribute selection and clustering. The biometric template has lesser in size compared to other forms of biometrics like face, finger, retina, etc.

The ECG signal is filtered using digital filtering techniques to remove power line interference and base line wander. Support Vector Machine (SVM) [18] is used as a classifier for detection of QRS complexes in ECG.

An Efficient DDC algorithm has been developed over existing modified Amplitude Zone Time Epoch Coding (AZTEC) technique, named as improved modified AZTEC [13] and tested on Common Standard for quantitative Electrocardiography (CSE) database. The performance has been evaluated on the basis of compression ratio (CR), percent root-mean-square difference (PRD) and fidelity of the reconstructed signal. A comparison of the wavelet-derived features of compressed and original signals has been used for performance evaluation of the compressed signal.

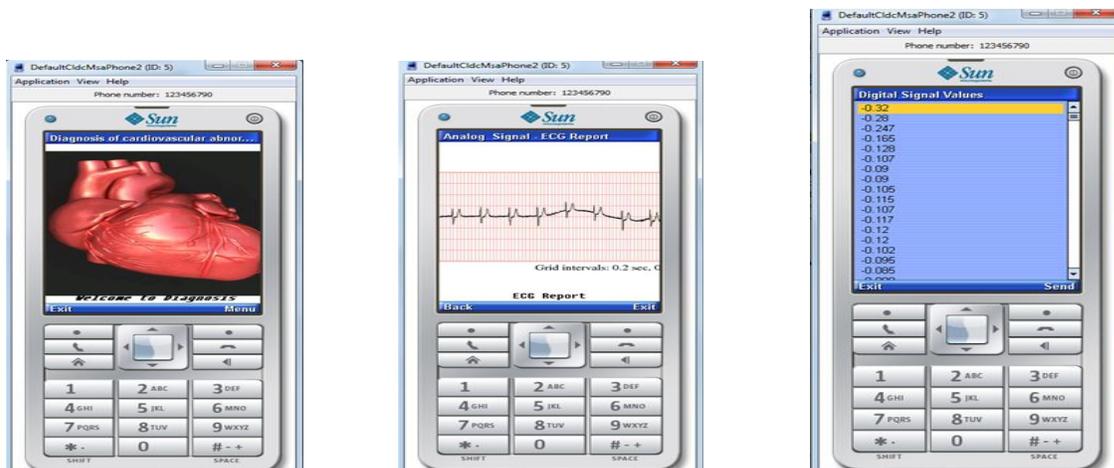
The ultimate aim of this proposed work is to provide a remote healthcare monitoring system for Faster Identification of Cardiac Abnormalities from Compressed ECG using advanced data mining approach. It is proposed to implement the data mining techniques on a hospital server that will generate a set of constraints for representing each of the abnormalities. Then, the patient's mobile phone receives these set of constraints and employs a rule-based system that can identify each of abnormal beats in real time.

### 3. SYSTEM AND METHOD

This system provides an efficient telecardiology diagnosis system for both patients and doctors for fast identification and treatment of the diseases. As we can see from Figure 1, the patient mobile phone receives the compressed ECG signal from the health care monitoring through bluetooth. The compressed ECG packets are sent to the hospital server through SMS/MMS/HTTP in digital format. After receiving the compressed ECG, the hospital server perform the disease recognition task with data mining techniques such as attribute selection and EM based clustering. Hospital server generates a range for particular disease. Based on the range, diseases are identified using rule based classification techniques.

#### 3.1 Patient's mobile phone

In the proposed system plays an important role in reducing the delay by performing diagnosis directly from the compressed ECG and mobile phone receive the digital format ECG from health care monitoring device to hospital server using HTTP. The telecardiology framework provides a fully automatic monitoring facility for the cardiac patient who requires ongoing monitoring facility. Healthcare monitoring to mobile phone communication was performed using bluetooth. Java 2 Micro Edition in net beans language was used for programming the small miniature program called midlet. Figure 2 shows different snapshots of our Midlet implementation of mobile phone based telecardiology application.



(a) Main Screen on Patient's Mobile

(b) Drawing ECG in Patient's Mobile

(c) Viewing ECG Samples in Patient's Mobile

Fig 2: Our Implementation of mobile phone based wireless telecardiology

### 3.2 Attribute Selection on Hospital Server

In the proposed system, Genetic Algorithm is used for the attribute selection. The attribute selection process is used to reduce the dimensionality of the analysed data, it will speed up learning algorithms, improves the performance of data mining techniques (e.g., learning time, predictive accuracy, etc.), and improves the comprehensibility of the output. The genetic algorithm (GA) is an optimization and search technique based on the principles of genetics and natural selection. A GA allows a population composed of many individuals to evolve under specified selection rules to a state that maximizes the fitness. A genetic algorithm mainly composed of three operators: selection, crossover, and mutation.

Table 1. Calculate the fitness value for each phenotype

Genotype	Phenotype	Fitness
Parent 1 : 1001010	0.331	0.109
Parent 2 : 0100101	- 0.835	0.697
Parent 3 : 1101010	1.339	1.790
Parent 4 : 0110110	- 0.300	0.090
Parent 5 : 1001111	0.488	0.238
Parent 300 : 0001101	- 1.591	2.531

Table 2. Generate the new generation (Child)

Genotype	Genotype	Phenotype	Fitness
Parent 6: 00 011 01	Child1:0001001	-1.685	2.839
Parent 3:11  010 10	Child2:1101110	1.433	2.054
Parent 5: 10 011 11	Child3:1001111	0.898	0.806
Parent 6: 00 011 01	Child4:0001101	-1.654	2.736
Parent 2: 01 001 01	Child5:0101001	1.339	1.793
Parent 1: 10 010 10	Child6:1000110	0.677	0.458

In selection, a good string (on the basis of fitness) is selected to breed a new generation from compressed ECG digital values. Roulette wheel selection method is used to select the parent in randomly. Individuals are given a probability of being selected that is directly proportionate to their fitness. Two individuals are chosen randomly based on these probabilities and produce offspring. Table 1 shows that fitness value for each chromosome. The fitness function evaluated for each individual, providing fitness values, which are the normalized. Normalization means dividing the fitness value of each individual by the sum of all fitness values, so that the sum of all resulting fitness values to 1. The population is sorted by the descending fitness value. After the selection process, crossover is performed.

Crossover is a process of taking more than one parent solutions and producing a child solution from them. Two-Point Crossover will generate a two cut-point in given two parents and recombines the first part of first parent with the second part of the second parent to create one

offspring and then recombines the second part of the first parent with the first part of the second parent to create a second offspring.

Mutation alters a string locally to maintain genetic diversity from one generation of a population of chromosomes to the next. In each generation, the population is evaluated and tested for termination of the algorithm. Mutation occurs during evolution according to a user-definable mutation probability. This probability should be set to low. Table 2 shows that generate a new generation using crossover and mutation. Based on the genetic algorithm best fitness values are generated from the total fitness value.

#### **Algorithm 1: Genetic Algorithm**

- Step 1 : Randomly select two parents : P1,P2 from R,and place them in to matting pool
- Step2 : Apply the suggested two-point crossover operation on the selected two parents P1,P2 to generate two new offspring say : O1,O2.
- Step 2.1 : Insert two offspring ( O1, O2) into new population.
- Step 3 : For mutation, select one individual based on the fitness value.
- Step 3.1: Insert mutation into new population.
- Step 4 : Combine the crossover and mutation, produce the offspring.
- Step 5 : If the desired number of generations is not completed, go to Step 1. Otherwise end.

### **3.3 Cluster Formation on Hospital Server**

This techniques deal with the missing or erroneous feature set, handle high dimensional data and accept the desired number of clusters as input. Expectation Maximization (EM) is basically iterates between two core steps, called the E-step and the M-step. In the E-Step, assignment of the probabilities for each instances being a member of a particular class is performed. On the other hand, M-step updates or reestimates the probability values by calculating the log likelihood data. As shown in Algorithm 2, our implementation of EM revolves around the idea that every single ECG packet Corresponding to a Particular CVD is assigned to a cluster for identify the disease. We assume that there are N number of clusters (C1, C2, CN) representing N number of diseases. At the beginning of the algorithm2, step1 represents the initial value of the mean, standard deviation for each of the patient Step 2 denote the value of the standard deviation and the mean are refined and updated. Step 3 produce the final value for mean and standard deviation. At the end of this process, we will have the same number of mean and standard deviations as the number of cardiac anomalies.

<b>Algorithm 2: EM based Clustering</b>
<p>Step 1: Start with initial value of mean (<math>\mu</math>), standard deviation (<math>\sigma</math>) and probability (P).</p> <p>Step 2: Loop (For each iteration <math>i</math> - Number of instances) Calculate the probability of each element belong to each cluster <math>C_1, C_2 \dots C_N</math>.</p> $P(C_k/x_i) = \frac{P(C_k)P(x_i/C_k)}{P(x_i)}$ <p>The Probability of <math>P(x_i/C_k)</math> can be modeled using any distribution function for the commonly used normal (Gaussian) distribution; it can be given by,</p> $P(x_i/C_k) = N(\sigma_{C_k}, \mu_{C_k})$ <p>Update the mixture parameters on the basis of the new estimates.</p> $P_{C_k} = \sum_{i=1}^n \sum_{k=1}^N \frac{P(C_k/x_i)}{n}$ $\mu_{C_k} = \sum_{i=1}^n \sum_{k=1}^N \frac{x_i P(C_k/x_i)}{P(C_k/x_i)}$ $\sigma_{C_k} = \sum_{i=1}^n \sum_{k=1}^N \frac{P(C_k/x_i)(x_i - \mu_{C_k})^2}{P(C_k/x_i)}$ <p>End loop</p> <p>Step 3: Final value of mean (<math>\mu</math>), standard deviation (<math>\sigma</math>) and probability (P).</p>

### 3.4 Rule Based Technique on a mobile phone

Using the proposed technique, one or more diseases can be successfully identified.  $\Gamma$  represents the constraint set for  $N$  number of diseases. Each element within  $\Gamma$  (i.e., an individual disease) contains the actual frequency constraints. If a patient is recommended for ongoing monitoring, then during the holter monitoring the doctors could have already obtained the patient's normal and multiple abnormal ECG traces, from where the constraints are computed within the hospital server (before the patient is monitored wirelessly). Then, normal and CVD-affected ECG traces can be fed to our Data Mining (DM) model (i.e., the model that executes on hospital server) to obtain the constraint set ( $\Gamma_i$ ) for all the cases (i.e., normal and all the abnormal cases). Equation (1) shows the constraint set ( $\Gamma_i$ ) for an individual case, where  $i = 0, 1, 2, \dots, N$ ,  $N$  is the maximum number of diseases, and  $M$  is the total number of attributes for that particular disease. The number of attributes depends on the attribute selection procedure on each patient. Equation (2) shows the mean values of each attribute (i.e., for all  $M$  attributes). On the other hand, Equation (3) shows the standard deviations against the mean values  $\mu_i$ . Therefore, Equation (4) demonstrates the valid ranges of each cluster (e.g., normal beat, ventricular tachyarrhythmia, premature ventricular beats, etc.). These constraint set can efficiently be calculated from the proposed DM.

$$\Gamma = \{\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_N\} \quad (1)$$

$$\Gamma_i = \{f_1^i, f_2^i, f_3^i \dots f_M^i\} \quad (2)$$

$$\mu_i = \{\mu_1^i, \mu_2^i, \mu_3^i \dots \mu_M^i\} \quad (3)$$

$$\sigma_i = \{\sigma_1^i, \sigma_2^i, \sigma_3^i \dots \sigma_M^i\} \quad (4)$$

$$f_M^i = \mu_M^i \pm \sigma_M^i \quad (5)$$

Patient's mobile can execute a rule-based system following the condition specified by Equation (5) to detect particular abnormality.  $s_m$  is the value of the  $m_{th}$  attribute.  $F$  is the set containing the frequency counts of encoding ECG signal. Based on the condition set out by Equation (6), the compressed ECG packet of the patient can be classified as belonging to a particular class  $i$  (where,  $i = 1, 2, 3, \dots, N$ ).

$$\forall s_m \text{ in } F : (\mu_m^i - \sigma_m^i) \leq s_m \leq (\mu_m^i + \sigma_m^i) \quad (6)$$

Patient's mobile phone transmits the compressed ECG packets to the hospital at a convenient time. On receipt of the compressed ECG, the hospital server calculates the constraints and forwards it to the patient's mobile phone. The patient's mobile phone then performs continuous monitoring or beat classification. When abnormal beat is detected, the mobile phone activates the messaging services.

#### 4. RESULTS AND DISCUSSIONS

In our experimentation the compressed ECG samples are collected from MIT-BIH (Massachusetts Institute of Technology - Beth Israel Hospital) physiobank database.

Table 3: Sample output

Physionet Database and Entry	Number of Samples	Best Fitness value using GA	Cluster	Beat Classification
CU 01	300	278	Class 1	Ventricular Flutter
CU09	300	274	Class 2	Atrial Fibrillation
200	300	280	Class 3	Atrial Premature beat
208	300	292	Class 4	Premature Ventricular Contraction

The programs for patient to doctor communication were performed with Midlet implementation using netbeans IDE 7.0.1. Healthcare monitoring to mobile phone communication was performed using bluetooth. On the other hand, patient to doctor / patient to server / server to doctor mobile phone communication via MMS, SMS and HTTP were performed. Data mining techniques are running on hospital server using netbeans. Our experimentation attribute selection using genetic algorithm produce a 278 best fitness value from 50820 total fitness for ECG samples.

The quality and performance of the EM based algorithm becomes very good when a huge dataset. EM based clustering calculate the mean and standard deviations from the best fitness value. The diseases are classified from the clustering range. After the disease classification, the hospital server inform to patient's mobile phone. Figure 3. shows the hospital server send the message to the patient's mobile phone via HTTP and SMS. If the patient condition critical, automatically call to the ambulance.

### Send SMS/ Email

Name:	Rani
Email:	rani@gmail.com
Disease Diagonized:	Diagonised Disease is:PVC
	<input type="button" value="Post"/>

Name:	Rani
Phone:	9876543210
MSG:	PVC
	<input type="button" value="Send"/>

Fig 3: Hospital Server Transmit the Message to Patient Mobile Phone

## 5. CONCLUSION

Patient's mobile phone transmits the compressed ECG packets to the hospital at a convenient time (e.g., may be once a week, when Internet connectivity is available). This model is suitable whenever there is bandwidth restriction on the Internet (i.e., continuous monitoring by the hospital server is not possible). On receipt of the compressed ECG segments, the hospital server calculates the constraints (i.e.,  $\Gamma$  values) and forwards it to the patient's mobile phone. The patient's mobile phone then performs continuous monitoring or beat classification. This work could be enhanced as identification of the patient and retrieval of the patient's previous records for doctor analysis.

## REFERENCES

- [1] Alshamali A , Ghaith T, "Combined coding and adaptive thresholding algorithms for ECG compression", 5th IEEE transaction on GCC Conference March 2009, pp.1-3.

- [2] Bount M. Batra M. Capella N. Ebling R. Jerome R. Martin M. Nidd M. Niemi M. and Wright S. (2007), "Remote health care monitoring using Personal Care Connect", IBM Systems Journal Vol. 46 No 1.
- [3] DiHu Chen, Sheng Yang, "Compression of ECG signal using video codec technology-like scheme", Journal of Biomedical Science and Engineering Vol.1, No:1, May 2008, pp.22-26.
- [4] Ebrahim Nemati, M. Jamal Deen, and Tapas Mondal, McMaster "A Wireless wearable ECG Sensor for Long-Term Applications", IEEE Communications Magazine, January 2012.
- [5] Fahim Sufi, Ibrahim Khalil "Faster person identification using compressed ECG in time critical wireless telecardiology applications", Journal of Network and Computer Application, Vol.34, Issue 1, January 2011.
- [6] Fahim Sufi, Ibrahim Khalil, Abdun Naser Mahmood, "clustering based system for instant detection of cardiac abnormalities from compressed ECG", Journal Expert Systems with Applications: An International Journal, Vol.38, Issue 5, May, 2011.
- [7] Hamid SadAbadia, Masood Ghasemia, Ali Ghaff aria "Mathematical algorithm for ECG signal denoising using window analysis", IEEE transaction Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub. 2007, 151(1):73-78.
- [8] Jose García, Ignacio Martínez, Leif Sörnmo, Member, IEEE, Salvador Olmos, Associate Member, IEEE, Angel Mur, and Pablo Laguna, "Remote Processing Server for ECG-Based Clinical Diagnosis Support", IEEE Transactions On Information Technology in Biomedicine, Vol. 6, No. 4, December 2002, pp.277-284.
- [9] Joseph J. Oresko, "Real-Time Cardiovascular Disease Detection Via electrocardiogram Processing", IEEE transactions on information technology in biomedicine, Vol. 14, No. 3, May 2010, pp.734-740.
- [10] Karupothula Madhavi Latha 1, Shiramshetty Gouthami2, P. Ranjith Kumar, "Scoring System Approach for Assessment of Critical Illness using Mobile Phones", International Journal of Distributed and Parallel Systems (IJDPS) Vol.3, No.1, January 2012.
- [11] Mamaghanian H, Khaled N, Atienza D, Vandergheynst P, "Compressed Sensing for Real-Time Energy-Efficient ECG Compression on Wireless Body Sensor Nodes", IEEE Transactions on Biomedical Engineering Volume: 58 Issue: 9, Sept. 2011, pp.2456-2466.
- [12] Marcin Bajorek. and Jędrzej Nowak. (2011), "The role of a mobile device in a home monitoring healthcare system", Conference On Computer Science and Information Systems.
- [13] Nibedit Dey, Tripada Prasad Dash, Sriram Dash, "ECG Signal denoising by Functional Link Artificial Neural Network (FLANN)", International Journal of Biomedical Engineering and Technology, Vol. 7, No.4 December 2011, pp. 377 - 389.
- [14] Palaniappan S, Awang R, "Intelligent heart disease prediction system using data mining techniques", IEEE/ACS International Conference on Computer Systems and Applications, 2008. AICCSA 2008, ISBN:978-1-4244-1967-8.
- [15] Physiobank: Physiologic signal archives for biomedical research (2009). [online]. Available: <http://www.physionet.org/physiobank/>
- [16] Ranveig Nygaard, Gerry Melnikov, and Aggelos K. Katsaggelos, "A Rate Distortion Optimal ECG Coding Algorithm", IEEE Transactions On Biomedical Engineering, Vol. 48, No. 1, January 2001, pp.28-40.
- [17] S. S. Mehta and N. S. Lingayat, "Comparative study of QRS detection in single lead and 12-lead ECG based on entropy and combined entropy criteria using support vector machine", Journal of Theoretical and Applied Information Technology.
- [18] Vinod Kumar a, S.C. Saxena b, V.K. Giri a, Dilbag Singh, "Improved modified AZTEC technique for ECG data compression: Effect of length of parabolic filter on reconstructed signal", Computers and Electrical Engineering, pp. 334-344, 2005.
- [19] Ying Zhang and Hannan Xiao, "Bluetooth-Based Sensor Networks for Remotely Monitoring the Physiological Signals of a Patient", IEEE Transactions On Information Technology In Biomedicine, Vol. 13, No. 6, November 2009, pp.1040-1048.