A new real-time resource-efficient algorithm for ECG denoising, feature extraction and classification-based wearable sensor network

Ali Fadel Marhoon*

Computer Science Department,
University of Basrah,
Basrah, Iraq
Email: ali_marhoon@yahoo.com
*Corresponding author

Ali Hussein Hamad

Information and Communication Engineering Department,
University of Baghdad,
Baghdad, Iraq
Email: ahamad@kecbu.uobaghdad.edu.iq

Abstract: Long-term patient monitoring is an important issue especially for the elderly. This can be done using a wearable wireless sensor network. These sensors have limited resources in terms of computation, storage memory, size and mainly in power. In this work, a real-time resource-efficient algorithm has been implemented and tested practically such that not all the Electrocardiography (ECG) data are transmitted to the server for later processing. The algorithm reads a sample window and processes it on the sensor node using an adaptive filter with a differentiator and then a fast and simple algorithm for feature extraction of the ECG signal to find P, Q, R, S and T waves. Finally, a classifier algorithm has been designed to distinguish between normal and abnormal ECG signals. The work has been implemented using Shimmer sensor nodes and uses the open source TinyOS 2.1.2 and Python 2.7.

Keywords: wearable sensor network; ECG signal; P, Q, R, S and T waves; feature extraction; TinyOS; real time; adaptive filter; classification; resource-efficient algorithm; patient monitoring; biomedical.

Reference to this paper should be made as follows: Marhoon, A.F. and Hamad, A.H. (2015) 'A new real-time resource-efficient algorithm for ECG denoising, feature extraction and classification-based wearable sensor network', *Int. J. Biomedical Engineering and Technology*, Vol. 18, No. 2, pp.103–114.

Biographical notes: Ali Fadel Marhoon is a Professor at the Computer Science University of Basrah, Iraq. His research interest is wireless sensor network and artificial intelligence.

Ali Hussein Hamad is a PhD Student at the Department of Electricity Engineering, University of Basrah, Iraq. His research interest is wearable sensor network for healthcare monitoring system.

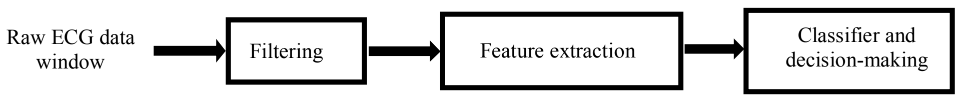
1 Introduction

Arrhythmia is a common medical condition which includes a broad range of heart-related pathologies. Although not all of them are permanent or require medical attention, they may provide hints to the development of serious heart diseases. ECG has been a cornerstone for the detection and diagnosis of such conditions for a long time (Gradl et al., 2012). Wearable Body Area Sensor Network (WBASN) can be used for long-term remote Electrocardiogram (ECG) monitoring. In general, this system consists of small motes, energy-constrained, limited computation capabilities and small size of storage memory. These sensors capture raw ECG data and transmit them to a Personal Digital Assistant (PDA) such as smartphone (gateway), in which the data are gathered and transmitted to the caregiver's server over the internet. In some cases, data processing and decision-making are performed in the gateway; the other is made in the server. Power consumption in wireless sensor network lies in three main parts from the least to the most: sensing, computation and transmission (Sendra et al., 2011). In e-healthcare system, long-term monitoring of ECG is required especially for the elderly. Due to limitations in the power of sensor node, where the wireless transmission consumes significant power, it is impractical to send all the data during the daily activity of the patient; also, a large data would be obtained, which may be normal ECG data. Energy consumption within wireless sensor network has the following components (Albu et al., 2010):

$$Energy = Sense\ () + Computation\ () + Communication\ ()$$

The sensing consumption depends on the sample rate, sensor hardware and the sampling duration. The consumption for computation depends on the microcontroller used, the code of the algorithm and power saving technique. The communication consumption depends on the used radio technology, the energy required to send 1 bit, the packet overhead, payload limit and data rate of the application. In this work, a lightweight resource-efficient algorithm has been implemented for ECG monitoring-based WBASN such that the processing of the patient data is performed inside the sensor node. If there is a significant abnormal data, then the sensor sends an alarm message to the caregiver's server and then starts transmission of ECG data without any processing. If the sample window is normal data, it would be ignored and another sample window tested. The algorithm consists of three modules as shown in Figure 1.

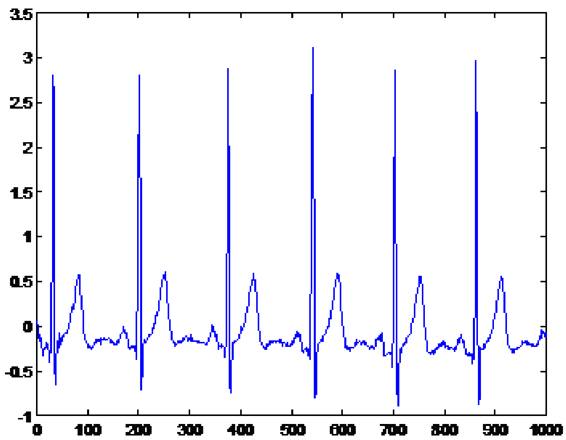
Figure 1 Module of the proposed algorithm modules



2 Literature survey

Recent study shows some techniques for ECG feature extraction. Zhao and Zhan (2005) proposed a wavelet transform and vector machines for feature extraction where a new approach for heart rhythm recognition and classification of ECG has been included. Liu et al. (2011) proposed an integrated ECG signal processing scheme using a systematic wavelet transform algorithm which can realise multiple functions in real time, including baseline drift removal, noise suppression, QRS detection, heart rate prediction and classification. Nabar et al. (2011) proposed GeM-REM, a resource-efficient ECG monitoring method for body area network which uses a generative ECG model at the base station and its lightweight version at the sensor. The sensor transmits data only when the sensed ECG deviates from the model-based value. Jayasumana et al. (2010) present a new system developed to continuously monitor the ECG of patients and analyse them in real time to identify any abnormalities where the system consists of two parts: a portable device on the patient and the server side. Cordier et al. (2010) formalise the concept of learning symbolic rules from multisource data in a cardiac monitoring context. The sources, electrocardiograms and arterial blood pressure measures, describe cardiac behaviours from different viewpoints. Many researchers such as Batra and Kapoor (2013) and El Mimouni and Karim (2013) have adopted the work of Pan and Tompkins (1985), where various filters are involved in the analysis of the ECG signal. The signal passes through a band pass filter composed of cascaded high pass and low pass filters. Subsequent processes are differentiation and time averaging of the signal. The drawback of this approach is that it detects QRS complex only but cannot detect P and T waves.

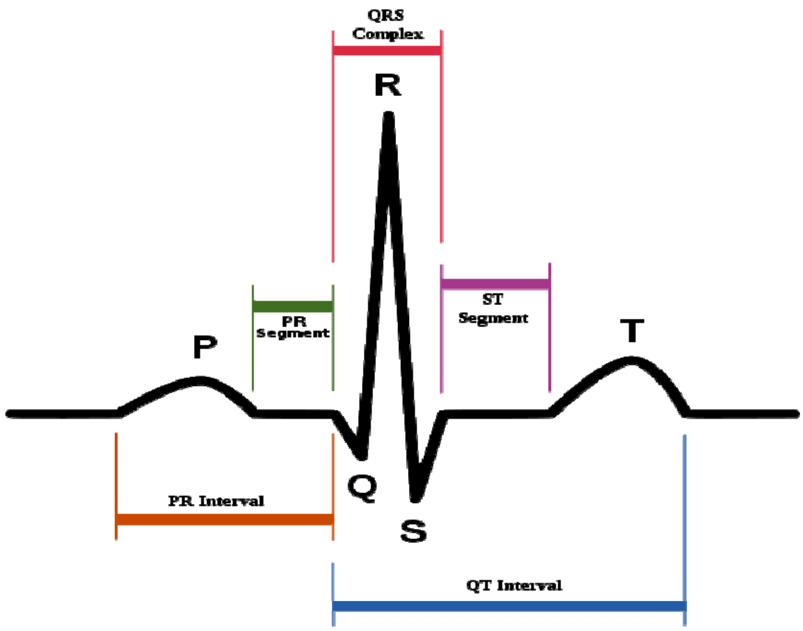
Figure 2 Raw ECG data for 1000 samples (approximately 5 seconds) (see online version for colours)



3 Methods

The algorithm states that instead of transmitting all the raw ECG data to the server, process these data there; the process of data and disease diagnosis has been made in the sensor node. To do this, a problem of limited RAM of the sensor node (10 Kb for Shimmer2r) has been solved by taking a sample window of 5 seconds (1000 sample, where the sampling rate was 200 Hz). The process involved in this window including the filtering takes only 9 Kb of RAM. The proposed system reads these sample windows continuously; then it would be filtered due to the noisy nature of the ECG signal and the effect of the artefact of the patient during daylight. Figure 2 shows the raw ECG data. A simple and fast search algorithm for feature extraction is performed to find P, Q, R, S and T peaks. Figure 3 shows the properties of the ECG signal. The measurement of the duration of PQ, QRS and QT wave forms and analysis of their variations in the ECG are used in the next module to monitor electrical activity in the heart to detect any damage or disease.

Figure 3 Properties of the ECG signal (see online version for colours)



In this work, ECG capturing, filtering and classification were developed to process the signal generated from the wearable sensor network. The processing algorithm can be divided in four main modules: (i) sensor node, (ii) filter design, (iii) feature extraction and (iv) beat classification.

3.1 Sensor module

In this work, the target embedded sensor system is the Shimmer platform (Shimmer, 2014). From the hardware viewpoint, this platform includes a low-power 16-bit microcontroller, Texas Instrument MSP430F1611, a low-power radio supported with 802.15.4 radio and an extension module for Electrocardiography (ECG), Electromyograph (EMG), which measures and records the electrical activity associated with skeletal muscle contractions, and Galvanic Skin Response (GSR), which monitors skin resistance. Also, the platform has a built-in 3D accelerometer acquisition. The MSP430 microcontroller runs at 8 MHz, has 10 KB of RAM, 48 KB of flash and includes a fast hardware multiplier. In this work, a three-lead two-channel ECG sensor extension module provided by Shimmer has been used (Figures 4 and 5) where

- Lead II is the ECG vector signal derived from the RA to LL vector; this is derived from Channel 1 on the Shimmer ECG board.

- Lead III is the ECG vector signal derived from the LA to LL vector; this is derived from Channel 2 on the Shimmer ECG board.
- Lead I is the ECG vector signal derived from the RA to LA vector; this is derived by subtracting Lead III from Lead II.

Figure 4 d ECG leads connections on body

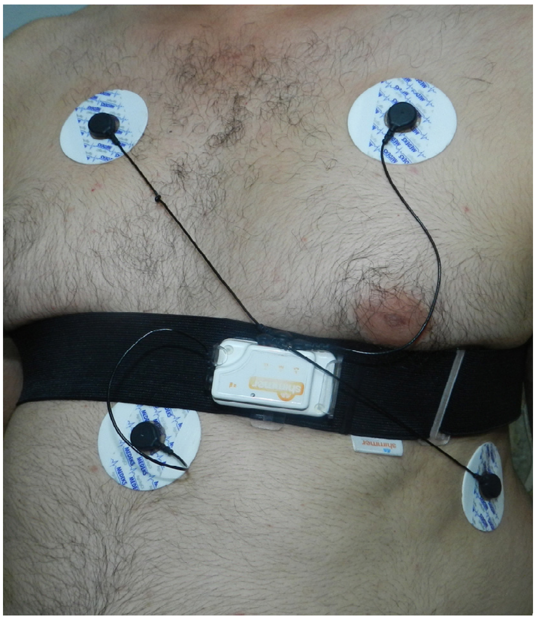
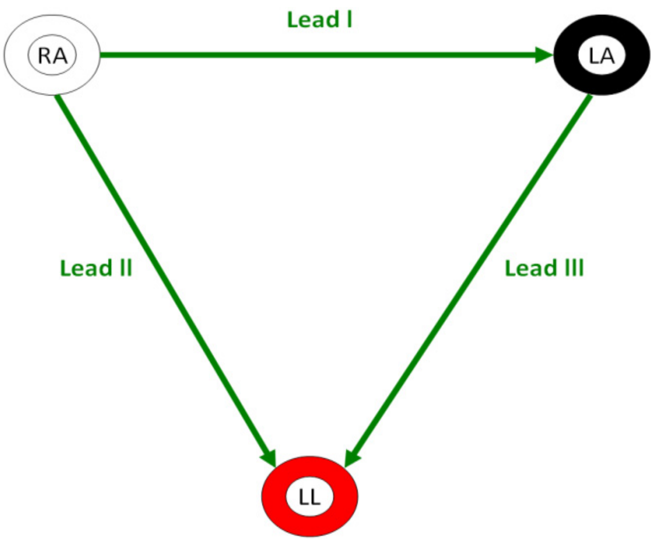


Figure 5 ECG signal vectors (see online version for colours)



Source: Batra and Kapoor (2013)

An open source TinyOS 2.1.2 has been used for senor board programming; also, the open source Python 2.7 under Linux has been used for data collection and presentation.

3.2 Filtering

ECG signal contains two types of noise: biological noise and environmental noise. Biological noise occurs due to physiological interference, such as low-frequency noise (baseline wanders < 0.03 Hz). Respiration and muscle movement of the human body may produce a high-frequency EMG noise (1–5000 Hz). Also, another noise may occur due to ECG electrodes' movement from their places or due to motion artefact. Environmental noise originated from instruments and circuit components like power line interference (50–60 Hz), electrode contact noise, electrosurgical noise, radio frequency noise. The frequency range of ECG wave varies from 0.1 to 250 Hz (Gautam et al., 2008). Majority of the existing related research use multi-stage filters such as low pass, high pass and notch filter (Pan and Tompkins, 1985). By this approach, only QRS complex is detected, while the P and T waves are omitted since they behave as low frequency. Another approach is to use wavelet (Zhao and Zhan, 2005) such that it is good in denoising and can find all the features of the ECG signal. Both approaches are not well suited in wearable sensor networks for two reasons. First, it requires a high amount of storage memory, which is not available for all sensor types. Secondly, it requires a huge computation processing which in turn consumes more power of the sensor. The high-frequency part in the ECG signal is the QRS complex, a derivate action has been added first to amplify the high-frequency characteristics of the QRS complex and to provide information about the slop of the complex. A five-point derivative has the transfer function (Pan and Tompkins, 1985)

$$H(z) = 0.1 * \left(2 + z^{-1} - z^{-3} - z^{-4}\right)$$

The discrete formula is as follows:

$$\begin{aligned} ecg\_der(nT) = 0.1 * 2 * \big(&ecg\_raw(nT) + ecg\_raw(nT - T) \\ &- ecg\_raw(nT - 3T) - ecg\_raw(nT - 4T)\big) \end{aligned} \tag{1}$$

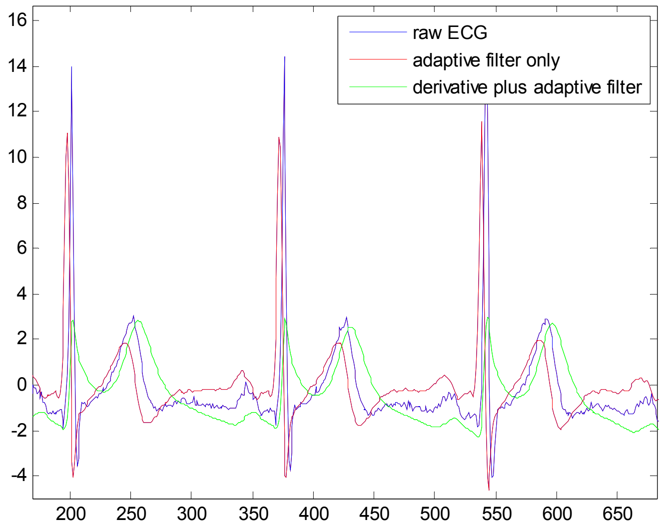
where *ecg_raw*(*nT*) is the raw ECG recorded data and *ecg_der*(*nT*) is the differentiation of the raw input ECG signal. Filters such as low pass, high pass and notch filters can reduce the interferences on ECG signal, while muscles noise, muscle artefact and baseline wander cannot be eliminated because of their irregular occurrences and irregular morphological attributes. In this work, an adaptive filter has been used which gives full features of the ECG signal, and it is efficient in reducing the motion artefacts and muscle noise, muscle artefact, AC noise and baseline wander. The resultant output series named *ecg_adp*(*nT*) are generated by performing an adaptive filter to the differentiated series named *ecg_der*(*nT*). The adaptive filter expression is characterised by (Zhou et al., 2009) the following:

$$ecg\_adp(nT) = \alpha * ecg\_adp(nT - T) + (1 - \alpha) * ecg\_der(nT) \, n = 1, \ldots, \infty \tag{2}$$

where α is the balance coefficient that is relative to the signal sampling frequency (default is 0.95). This type of filter is good in denoising, but the QRS complex does not appear clear. Figure 6 shows the raw input data and a comparison between using adaptive filter only and adaptive filter with derivative action where the effect of the derivative on

QRS complex appearance as well as the P and T waves is clear. Figure 6 shows the filtered ECG signal obtained where P, Q, R, S and T peaks are found using the feature search algorithm.

Figure 6 The filtered data without the derivative action (see online version for colours)



3.3 Feature extraction algorithm

A segment of the ECG of 1000 samples (5 second window where the sampling frequency is 200 Hz) of data is sampled to obtain the features. We choose the size of the window such that multiple beats are present (five beats at least). Also, this is within the available RAM of the sensor node (10 Kbyte). Measuring ECG features such as R-R intervals, QRS complex PQ and width requires detecting Q, R, S, P and T peaks. In order to perform peak detection at low computational overhead, we develop a lightweight feature extraction algorithm. The main part of this algorithm is the detection of the R peaks. Other features depend on the detection of these peaks. The search algorithm is summarised as Algorithm 1, while Figure 7 shows the results of the search algorithm.

3.4 Classifier design (Surawicz, 2009; Rautaharju, 2009)

The normal ECG characteristics of sinus rhythm are as follows:

- *Rate*: 60–100 bpm.
- *Rhythm*: regular.
- *P-wave*: uniform, upright, normal shape, one before each QRS complex.
- *QRS duration*: 60–100 msec.
- *PQ interval*: 120–200 msec.
- *QT_c interval*: 390–450 msec.

Algorithm 1

1- *Search for a maximum value within the sampled window which represents one of the R peaks (MAX).*

2- *Search for a minimum value within the sampled window which represents one of S peaks (MIN).*

3- *Obtain a threshold such that:*

Threshold R = MAX/2 and threshold S = MIN/2

4- *Find R_i peaks overall the sampled window which should be above threshold R*

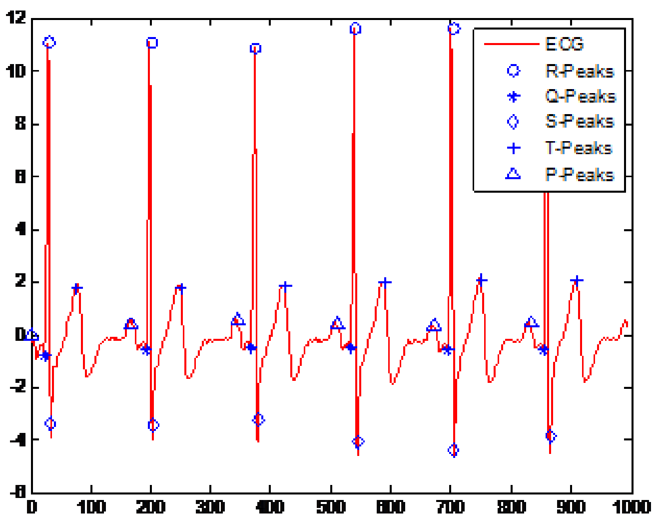
5- *For each R_i, create a sub-window [R_i to R_i+10] (S peaks appear normally after the R peaks by a few samples) samples to search for S peaks which should be below threshold S.*

6- *For each R_i create a sub-window [R_i R_i–10] (Q peaks appear normally before the R peaks by a few samples) samples to search for Q peaks. Now the QRS complex has been obtained.*

7- *To find P peaks, create a window [R_i–100 to R_i – 25] (P peaks appear before the Q peaks) and search for the maximum value.*

8- *To find T peaks, create a window [R_i +25 to R_i+100] (T peaks appear after the S peaks) and search for the maximum value.*

Figure 7 The results of the feature extraction search algorithm (see online version for colours)



3.4.1 The PQ interval

The PQ interval starts at the beginning of the atrial contraction and ends at the beginning of the ventricular contraction. The PQ interval indicates how fast the action potential is transmitted through the AV node (atrioventricular) from the atria to the ventricles. Measurement should start at the beginning of the P-wave and end at the beginning of the

QRS segment. A prolonged PQ interval is a sign of a degradation of the conduction system or increased vagal tone. This is called the *first, second or third degree AV block.* A short PQ interval can be seen in the WPW syndrome in which faster-than-normal conduction exists between the atria and the ventricles.

3.4.2 The QRS duration

The QRS duration indicates how fast the ventricles depolarise. The ventricles depolarise normally within 100 msec. When this is longer than 110 msec, this is a *conduction delay*. Possible causes of a QRS duration of more than 110 msec include the following:

- *Blocked left bundle branch.*
- *Blocked right bundle branch.*
- *Electrolyte disorders.*
- *Idioventricular rhythm and paced rhythm.*

3.4.3 The QT interval

The normal QT_c (corrected) interval indicates how fast the ventricles are repolarised, becoming ready for a new cycle. If QT_c is less than 340 ms, *short QT syndrome* can be considered. One difficultly of QT interpretation is that the QT interval gets shorter as the heart rate increases. This problem can be solved by correcting the QT time for heart rate using the Bazett formula:

$$QT_c = \frac{QT}{\sqrt{RR\ interval(\mathrm{sec})}} \tag{3}$$

The classifier module is capable of identifying arrhythmic episodes, namely tachycardia, bradycardia, first, second or third degree AV block, left bundle branch and right bundle branch. The procedure for beat segmentation and classification for heart disease diagnosis is described in Algorithm 2.

Algorithm 2

Compute P, Q, R, S and T indices for each beat in the window.

Find the average heart rate HR = 60/(average R-R intervals/sampling rate) beat per minute

if HR < 60 bpm, then send alarm message for bradycardia

if HR > 100 bpm, then send alarm message for tachycardia

if no. of P-wave ≠ no. of QRS wave, then send alarm message for AV block

if QRS duration > 110 msec, then send alarm message for conductance delay

if QT_c *<340 msec, then send alarm message for short QT syndrome*

4 Discussion

In this work, we propose a new real-time lightweight algorithm that maintains the constraints of wearable wireless sensor network in storage memory and power consumption, such that not all ECG data are transmitted to the server for later processing. The ECG data are sampled, filtered, processed for feature extraction, and diagnosis for cardiac care on node. The proposed algorithm is a pure standalone sensor processing and is compared with Nabar et al. (2011), where a generative ECG model is used at the base station and its lightweight version at the sensor. The sensor transmits data only when the sensed ECG deviates from model-based values. In this work, the algorithm is developed for continuously monitoring the patient's ECG data and analyse them on the node in real time, and can detect the abnormality. The first step of this work involves a design of filter which was initially done using the algorithm proposed by Zhou et al. (2009) where only adaptive filter has been used which cannot show a clear QRS complex alone. We were able to improve this algorithm by incorporating a differentiator along with an adaptive filter which shows a good result as shown in Figure 5. The differentiator has a drawback such that it is particularly sensitive to high frequency. To overcome this disadvantage, the adaptive filter has been used after the differentiator to eliminate the high-frequency noise occurring from the differentiator. The algorithm is also capable of data classification for cardiac disease diagnosis, which is important in patient monitoring. Other possible extensions of the algorithm can be made for more disease diagnosis. The MSP430 does not support a floating point in the sending payload message, so the overall computation was implemented inside the sensor node which will send only the diagnostic status as a code as shown in Table 1.

Table 1 Code sample for some diagnostic cases

Diagnostic	*Case status*	*code*
Heart rate < 60 bpm	Bradycardia	00
Heart rate > 100 bpm	Tachycardia	01
No. of P wave ≠ no. of QRS	AV block	02
QRS duration > 110 msec	Conductance delay	03
QTc < 340 msec	Short QT syndrome	04

5 Conclusion

The aim of this work is to design and implement a real-time algorithm which maintains the limitation of the resources in the used hardware. The algorithm is designed for online automatic ECG diagnosis such that it is used for long-term patient monitoring. The available memory in Shimmer unit was (10 KB) of RAM, the algorithm was optimised to be implemented in that platform such that the total amount of data used is about (9 KB) of RAM. Since the MSP430 does not have a floating point unit, TinyOS 2.1.2 was used for implementation of the algorithm which supports floating point computation. Our goal was to reduce the amount of streamed wireless data sent to the server for computation and decision-making. This goal was achieved through maintenance of the limited processing and storage resources of the sensor platform unit. This approach would increase significantly the lifetime of the sensor node and the hence the overall monitoring system. Finally, for additional future work, we would like to incorporate other medical sensors such as EMG, GSR and body temperature. Also, studying the effect of the patient motion activity and body posture may be useful with ECG monitoring.

References

Albu, D., Lukkien, J. and Verhoeven, R. (2010) 'Energy effect of on-node processing of ECG signals', *Proceedings of the Digest of Technical Papers International Conference on Consumer Electronics*, 9–13 January, Las Vegas, NV, pp.7–8.

Batra, P. and Kapoor, R. (2013) 'A novel algorithm for ECG signal processing', *International Journal of Computer Science and Telecommunications*, Vol. 4, No. 2, pp.103–107.

Cordier, M-O., Fromont, E. and Quiniou, R. (2010) 'Learning rules from multisource data for cardiac monitoring', *International Journal of Biomedical Engineering and Technology*, Vol. 3, Nos. 1–2, pp.133–155.

El Mimouni, E.H. and Karim, M. (2013) 'Novel real-time FPGA-based QRS detector using adaptive threshold with the previous smallest peak of ECG signal', *Journal of Theoretical and Applied Information Technology*, Vol. 50, No. 1, pp.33–43.

Gautam, A., Lee, Y-D. and Chung, W-Y. (2008) 'Noise cancellation of electrocardiogram signal measured by wearable USN node', *Proceeding of the IEEE International Conference on Multisensory Fusion and Integration for Intelligent Systems*, 20–22 August, Seoul, South Korea, pp.154–157.

Gradl, S., Kugler, P. and Eskofier, B. (2012) 'Real-time ECG monitoring and arrhythmia detection using Android-based mobile devices', *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 28 August–1 September, San Diego, CA, pp.2452–2455.

Jayasumana, H.M.K.G.S., Thennakoon, T.M.U.A.S., Chandrasekara, C.M.R.B. and Sandaruwan, M.T. (2010) 'A stand-alone ECG abnormality detector', *Proceedings of the 5th International Conference on Information and Automation for Sustainability*, 17–19 December, Colombo, pp.489–492.

Liu, X., Zheng, Y, Phyu, M.W., Zhao, B., Je, M. and Yuan, X. (2011) 'Multiple functional ECG signal processor for wearable applications of long-term cardiac monitoring', *IEEE Transactions on Biomedical Engineering*, Vol. 58, No. 2, pp.380–389.

Nabar, S., Banerjee, A., Gupta, S.K.S. and Poovendran, R. (2011) 'GeM-REM: generative model-driven resource efficient ECG monitoring in body sensor networks', *Proceedings of the International Conference on Body Sensor Networks*, 23–25 May, Dallas, TX, pp.1–6.

Pan, J. and Tompkins, W.J. (1985) 'A real-time QRS detection algorithm', *IEEE Transactions on Biomedical Engineering*, Vol. 32, pp.230–236.

Rautaharju, P.M. (2009) 'AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: Part IV: the ST segment, T and U waves, and the QT interval. A scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society, endorsed by the International Society for Computerized Electrocardiology', *Circulation*, Vol. 119, pp.e241–e250.

Sendra, S., Lloret, J., García, M. and Toledo, J.F. (2011) 'Power saving and energy optimization techniques for wireless sensor networks', *Journal of Communications*, Vol. 6, No. 6, pp.439–459.

Shimmer (2014) *Shimmer small wireless sensor platform designed to support wearable applications*. Available online at: http://shimmer-research.com (accessed on 20 June 2014).

Surawicz, B. (2009) 'AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: Part III: intraventricular conduction disturbances: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society, endorsed by the International Society for Computerized Electrocardiology, *Circulation*, Vol. 119, pp.e235–e240.

Zhao, Q. and Zhan, L. (2005) 'ECG feature extraction and classification using wavelet transform and support vector machines', *Proceedings of the International Conference on Neural Networks and Brain*, Vol. 2, 13–15 October, Beijing, pp.1089–1092.

Zhou, H., Hou, K-M. and Zuo, D. (2009) 'Real time automatic ECG diagnosis method dedicated to pervasive cardiac care ', *Wireless Sensor Network*, Vol. 1, pp.276–283.