Anomaly Detection in Medical Wireless Sensor Networks using SVM and Linear Regression Models

Osman Salem, LIPADE Laboratory, University of Paris Descartes, Paris, France

Alexey Guerassimov, LIPADE Laboratory, University of Paris Descartes, Paris, France

Ahmed Mehaoua, LIPADE Laboratory, University of Paris Descartes, France, Centre National de la Recherche Scientifique (CNRS), LaBRI, Bordeaux, France

Anthony Marcus, Department of Computer and Electrical, Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA

Borko Furht, Department of Computer and Electrical, Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA

ABSTRACT

This paper details the architecture and describes the preliminary experimentation with the proposed framework for anomaly detection in medical wireless body area networks for ubiquitous patient and healthcare monitoring. The architecture integrates novel data mining and machine learning algorithms with modern sensor fusion techniques. Knowing wireless sensor networks are prone to failures resulting from their limitations (i.e. limited energy resources and computational power), using this framework, the authors can distinguish between irregular variations in the physiological parameters of the monitored patient and faulty sensor data, to ensure reliable operations and real time global monitoring from smart devices. Sensor nodes are used to measure characteristics of the patient and the sensed data is stored on the local processing unit. Authorized users may access this patient data remotely as long as they maintain connectivity with their application enabled smart device. Anomalous or faulty measurement data resulting from damaged sensor nodes or caused by malicious external parties may lead to misdiagnosis or even death for patients. The authors' application uses a Support Vector Machine to classify abnormal instances in the incoming sensor data. If found, the authors apply a periodically rebuilt, regressive prediction model to the abnormal instance and determine if the patient is entering a critical state or if a sensor is reporting faulty readings. Using real patient data in our experiments, the results validate the robustness of our proposed framework. The authors further discuss the experimental analysis with the proposed approach which shows that it is quickly able to identify sensor anomalies and compared with several other algorithms, it maintains a higher true positive and lower false negative rate.

Keywords: *Healthcare Monitoring, Malicious Attacks, Medical Systems, Patient Anomaly Detection, Sensor Faults, Wireless Body Area Networks*

DOI: 10.4018/ijehmc.2014010102

1. INTRODUCTION

With the continual growth in expected duration of the average human lifetime (Kumar & Lee, 2012), the rise in population and number of elderly persons has led to inflated healthcare costs and shortage of professionals able to provide the care and treatment necessary to satisfy this increase in demand.

Today, healthcare professionals and caregivers are very interested in remote monitoring of elderly people and patient vital signs, as well as their surrounding environment. These requirements have sparked enormous interest in the utilization of Wireless Sensor Networks (WSNs).

Scientists and researchers have developed networks of wireless sensors, known as Wireless Body Area Networks (WBANs), which are composed of a set of small miniaturized sensors with wireless transmission capabilities, and may be externally attached or implanted. These devices are used to continuously gather physiological signals from patients or elderly people at home or in hospitals, and transmit collected data to a Local Processing Unit (LPU).

The LPU (e.g., smart phone, tablet, etc.) has superior processing power, batteries with increased energy resources and greater transmission range and bandwidth than the individual WBAN nodes. LPUs must be robust and able to process received measurements in real time, and raise medical alarms for caregivers upon sensing the deteriorating health state of patients to quickly react by taking appropriate actions (Otto, Milenkovi, Sanders, & Jovanov, 2005). Data may also be transmitted by the LPU to remote databases (DB) for storage and long term analysis.

WBANs have several advantages such as enabling doctors to monitor specific attributes of patients regardless of location, improving diagnosis accuracy and efficiency, and reducing the overall cost of health care by permitting doctors to constantly monitor patient health.

WBAN may also improve the chances of discovering diseases which further reduces risk and impacts the lifespan of individuals on a global scale. In this paper, we look to increase the usefulness of WBAN systems used in the healthcare industry by creating an application which is capable of "intelligently" discerning between patient health irregularities and sensor node failure.

There exist many medical WBAN systems which are publicly available for purchase including MICAz, MICA2, Tmote Sky, TelosB, IRIS, Imote2, and Shimmer. These types of WBANs are used to monitor and collect various physiological parameters of individuals such as Heart Rate (HR), pulse, oxygen saturation (SpO2), Respiration Rate (RR), Body Temperature (BT), ElectroCardioGram (ECG), ElectroMyoGram (EMG), Blood Pressure (BP), Blood Glucose Levels (BGL), Galvanic Skin Response (GSR), etc.

The ECG sensor, for example, is connected to three electrodes each of which is attached to the patients' chest for real time monitoring of the heart. Another type of sensor, the pulse oximeter, using infrared light and a photo sensor, simply clips to a patient finger and measures the pulse and blood oxygenation ratio (SpO2). While it may seem simplistic, the SpO2 sensor may detect asphyxia, insufficient oxygen (hypoxia), pneumonia and other blood oxygen related anomalies. The average human SpO2 ratio naturally exceeds 95%, but when this ratio drops below 90%, the pulse oximeter will trigger an alarm due to possible lung problems or respiratory failure. Prior to the assistance of these types of WBAN sensors, healthcare providers were reliant on big, expensive machines which were in short supply and required that the patient is observed directly while situated at the location of the machine.

The use of WBANs has been extended to monitor patients diagnosed with chronic illnesses and cognitive disorders such as Parkinson's, Diabetes, Alzheimer's, Asthma, and

Epilepsy. WBANs have proven to be great assets to both patients and healthcare providers in that they have reduced the costs associated with healthcare by solving problems such as overcapacity in hospitals, excessive waiting or sojourn times and the required number of nurses and doctors on call. WBANs also allow greater mobility for monitored individuals while constantly gathering and transmitting critical physiological data to their associated healthcare providers which may be useful for situations that require the long-term monitoring of a patient's recovery after leaving the hospital or assessing the impact of a patient's rehabilitation.

While WBANs have numerous advantages, their disadvantages range from poor reliability to the high susceptibility of security attacks after deployment. For us to exploit the strengths while reducing the probability of any weaknesses occurring, we must first look at what these weaknesses are in more detail so that we may find some means of mitigation. WBAN sensor nodes are prone to both hardware and software issues such as impaired components, sensor calibration, battery exhaustion or dislocation.

The sensor readings are themselves both unreliable and inaccurate (Ko, et al., 2010; Wang, Fang, Xing, & Chen, 2011; Zhang, et al., 2012), resulting from constrained hardware resources including reduced processing power, limited memory and energy resources, and transmission range. Individual sensor data gathering and transmission is also prone to several types of irregularities such as interference, noise, sensor misplacement, sweating patients, exhausted energy resources and external hacks and malevolent attacks such as data injection, modification or replay attacks that indirectly affect the LPU. This may lead to unexpected results, faulty alarms and diagnosis, and a reduction in public trust of these systems.

As a result, high false alarm rate and faulty measurements directly influence the public credibility of WBANs especially where dependability is exceedingly important as in the medical domain (Sahoo, 2012). If, for example, a pulse oximeter sensor is incorrectly attached or external fluorescent light radiates to the infrared sensor, erroneous measurements may result. In Chipara, Lu, Bailey, and Roman (2010), the authors found that the first source of unreliability in medical WSNs was the sensing components as opposed to some other problems (i.e. network failure, data transfer).

Nodes transmitting erroneous data have a negative impact on the accuracy of the gathered data which may have an effect on the patients' diagnosis. This may, in turn, lead to life threatening situations where emergency personnel receive false alarms based on node faults for a code blue. As a result, it becomes an extremely important task to detect erroneous measurements at the node level and differentiate between patient anomalies and node faults to minimize false alarms. Both patient anomalies and node faults produce abnormal measurements and require that each should be detected with the highest accuracy possible. We may only achieve this using an anomaly detection mechanism to recognise and extract abnormal patterns and correlations in the data and to differentiate between sick individuals and faulty sensors.

Anomaly based systems (Jurdak, Wang, Obst, & Valencia, 2011) typically look for irregular patterns in the data received from sensors as opposed to signature based intrusion detection systems where signatures are required to detect attacks. Signatures are neither available nor easy to write for healthcare monitoring applications. However, anomalies are defined as deviations from a dynamically updated normal model from the sensed data. Therefore an anomaly based detection approach is more adequate for WBANs given the absence of attack signatures. It is also important to note that anomaly based systems face challenges related to the training phase as it is difficult to find normal data in order to establish an appropriate normal profile.

Several anomaly-based detection techniques for sensor fault identification and isolation have been proposed and applied (Liu, Cheng, & Chen, 2007; Jurdak, Wang, Obst, & Valencia, 2011; Miao, Liu, He, Liu, & Papadias, 2011; Chen & Juang, 2012). These distributed techniques identify anomalies at the node level to prevent transmission of irregular values and reduce energy consumption. Using these distributed methods typically requires additional resources not found in most sensor node hardware. As a result, their accuracy is lower than centralized approaches which utilize a global representation for spatial-temporal analysis. To ensure reliable operation and accurate diagnosis, correlations between physical parameters, which exist in time and space, must be exploited in order to detect and extract irregular measurements. Usually, there is no spatial or temporal correlation among monitored attributes for faulty measurements.

Our primary focus in this paper is the detection of anomalous measurements in medical WBANs. We propose a novel machine learning based approach to detect abnormal values. First we use Support Vector Machine (SVM) (Bishop, 2006) to detect abnormal records, and when detected, we apply linear regression (Witten, Frank, & Hall, 2011) to pinpoint abnormal sensor measurements in an abnormal record. However, physiological attributes are heavily correlated, and changes occur typically in at least two or more parameters, e.g. in Atrial Fibrillation (AF) & Asthma, the heart rate and respiration rate increase simultaneously.

Our solution will increase the reliability of medical WBANs used for monitoring patients. Its primary task is to detect and extract anomalies in the WBAN data and, once found, differentiate between irregular patient vital signs and defective sensor measurements. Additionally, we seek to minimize false alarms triggered by anomalous sensors data.

The rest of this paper is organized as follows. In section 2, we review related work on anomaly detection and machine learning algorithms used in medical WSN. Section 3 briefly reviews SVM and linear regression used in our detection system. The proposed approach is presented in section 4. In section 5, we present our results from experimental evaluation, where we conduct a performance analysis of the proposed solution with real patient data. Finally, section 6 concludes the paper with a discussion of the results and plans for future work.

2. RELATED WORK

With the population of mankind ever increasing, medical facility vacancies are difficult to find, frustratingly lengthy waiting lines clogging emergency rooms, and the demand for doctors and staff seems to never be satisfied. These shortages result in the inability for many individuals to receive the care they needed. Due in part to the excessive congestion caused by many outpatients requiring minimal attention in these facilities and the evolution of WSN and smart devices, a new market was created for remote patient monitoring using small, wearable sensor systems. Researchers and scientists have worked hard to satisfy this demand, creating many novel systems which may alleviate, to some degree, the overcrowding issues for medical staff and healthcare facilities.

Novel architectures for monitoring patients, both in house and remotely, have been designed, developed and deployed in real world environments. One such system, MEDiSN (Ko, et al., 2010), CodeBlue (Malan, Fulford-jones, Welsh, & Moulton, 2004; Havard Sensor Networks Lab, 2013), LifeGuard (Montgomery, et al., 2004), AlarmNet (Wood, et al., 2006), Medical MoteCare (Navarro, Lawrence, & Lim, 2009), Vital Jacket (Cunha, et al., 2010). Some comprehensive survey studies of medical applications using WSNs are available in (Alemdar & Ersoy, 2010; Grgic, Žagar, & Križanovic, 2012). All of these systems are plagued with similar problems such as limited energy, faulty sensor hardware, and wireless transmission failure. As these networks often are responsible for monitoring a patient's livelihood, many researchers have created methods of autonomous fault detection for WSN and WBAN.

Authors in Zhang, Meratnia, and Havinga (2010) present a comprehensive analysis of modern fault and outlier detection techniques for WSNs. They present a comparative guideline detailing the steps necessary to appropriately select the best technique suitable for the characteristics of the data set. Several types of irregular readings have been captured and extracted from medical WSN data including single spikes, long duration spikes resulting from noisy environments, and continuously anomalous line fluctuations. To simplify the classification of WSN sensor fault types, the authors in Sharma, Golubchik, and Govindan (2010) categorize faulty measurements into short faults, faults resulting from noise and constant faults.

WSNs are plagued by a variety of issues that may endanger their functionality which stem from lack of quality and poor reliability (Zhang, Meratnia, & Havinga, 2010; Ying-xin, Xiang-guang, & Jun, 2011; Zhang, et al., 2012). Some of the more prevalent issues include hardware and software errors and faults, interference, widely variable environment dependent noise, dropped and lost packets, inconsistencies, and damaged sensors. New anomaly detection schemes for WSNs have been proposed which locate, extract, and classify atypical deviations in collected data to reduce false alarms generated as a result of faulty sensor measurements.

Authors in Banerjee, Xie, and Agrawal (2008) propose an algorithm to identify faulty sensors using the minimum and the maximum boundaries of the monitored parameters. Measurements which exceed the threshold of these boundaries are classified as inconsistent or outliers. Furthermore, medical WBAN systems may not assume all patients have the same attribute boundary intervals, as the min-max threshold values are dependent on an individuals' physiological characteristics including sex, age, weight, height, stress, and health condition.

Investigation and further study of machine learning algorithms for supervised classification and data mining algorithms for clustering has led to additional inspiration for our research team. Machine learning algorithms including Naïve Bayes (NB) (Yang, Dinh, & Chen, 2010), Bayesian Network (BN) (Farruggia, Giuseppe, & Ortolani, 2011), decision tree (C4.5) (Cheng, Xu, Pei, & Liu, 2010), Neural Networks (NN) (Bishop, 2006), K-Nearst Neighbor (KNN) (Bishop, 2006), Self-Organizing Map (SOM) (Siripanadorn, Hattagam, & Teaumroong, 2010) and Support Vector Machine (SVM) (Bishop, 2006) generate a variety of mathematical models based on correlational statistics from a training data set which are then applied to classify test instances as either normal or abnormal.

Several regression algorithms have been used in medical WBANs to build a model generated from time series data such as AutoRegression (Curiac & Volosencu, 2012), Least Square Error (Li, 2010), Non-seasonal Holt-Winters (Li, 2010). The authors in Xiaozhen, Hong, and Tong (2011) apply linear regression for missing data prediction and the experimental results validate their success, claiming low prediction errors. Another such project which applies logistic regression modelling (Huang, Jiang, Zhang, & Gao, 2010) evaluates the reliability of large scale industrial WSNs utilizing a static threshold. In Cheng, Xu, Pei, and Liu (2010) based on the J48 (decision tree) algorithm, the authors propose a large scale WSN diagnostic methodology which merges local classifier models into a single network spanning tree, responsible for the accuracy of the method and representative of the whole network.

To monitor an individual's physical activity, the authors in Yang, Dinh, and Chen (2010) use SunSpOT sensors attached to the thighs. Naïve Bayes is used to calculate values from the data to determine body position (i.e. sitting, standing, lying down, and walking). In a similar project which uses logistic regression (Choi, Ahmed, & Gutierrez-Osuna, 2012), a system is described which claims to use heart rate variability measurements to differentiate mental stress states from relaxation states.

In recent years SVM classification has become a more popular selection partially due to its simplistic numerical comparison for data classification and is often found to be the optimum solution for specific context. Several modern SVM based approaches have been proposed (Zhang, Meratnia, & Havinga, 2009; Rajasegarar, Leckie, Bezdek, & Palaniswami, 2010; Xu, Hu, Wang, & Zhang, 2012) for anomaly detection in WSNs. Furthermore, many non-linear versions (kernel based) of SVM have been investigated to find the optimum hyperplane that encompasses the majority of normal data in training phase. Once established, any data point landing outside the hyperplane boundary is classified as abnormal.

Often a major challenge in machine learning is that accurate model generation requires a training data set which has the classes labelled for each instance. The training data frequently requires close attention by researchers which must conduct extensive experiments to determine applicable pre-processing and balancing algorithms. We refer to Bishop (2006) for more details about these classification methods. Many attempts to resolve these challenges in training data set for machine learning led to methods of unsupervised learning or data mining.

Data mining algorithms group similar instances from the data into a single cluster and label smaller size clusters containing less than a given percentage of the total values, as abnormal. Some of the most popular and widely applied data mining algorithms include (Bishop, 2006) K-means, hierarchical clustering, Fuzzy C-means and GMM (Theodoridis, Pikrakis, Koutroumbas, & Cavouras, 2010). One challenge facing these clustering methods is that they assume anomalous data, which typically occurs much less frequently, is easily distinguished from normal data. We refer to Abduvaliyev, Pathan, Zhou, Roman, and Wong, (2013) for comprehensive classification of various detection techniques.

In Zhang, et al. (2010) a novel Outlier Detection and Countermeasure Scheme (ODCS) based on k-means, K-Nearest Neighbours (K-NN), static threshold and transmission frequency. K-NN is unsuitable for WSNs as it is computationally expensive and requires large amounts of memory space to store the training data as opposed to classification methods which discard the training data after building the model. Authors in Xie, Hu, Han, and Chen (2012) proposed a KNN-based anomaly detection method based on hyper-grid which has lower computational complexity than K-NN for WSNs. An unsupervised approach for anomaly detection in WSNs, Siripanadorn, Hattagam, and Teaumroong (2010) combines multiple models such as Discrete Wavelet Transform (DWT) and Self-Organizing Map (SOM). In this case, the DWT is used to reduce the size of input data for SOM clustering.

The authors in Liu, Cheng, and Chen (2007) proposed a distance based method to identify insider malicious sensors while assuming neighbour nodes are monitoring the same attributes. Each sensor monitors its one hop neighbours and measures the Mahalanobis distance between the calculated and actual values received in multivariate instances to detect anomalies. They discovered that it is not practical in medical WBAN applications to exploit promiscuous mode and increase network node redundancy which monitor the same parameters.

Authors in Yim and Choi (2010) propose a voting based approach to detect abnormal network events. In Miao, Liu, He, Liu, and Papadias (2011), the authors propose a failure detection approach for WSNs which exploits metric correlations to detect abnormal sensors and to uncover failed nodes. A simple prediction and fault detection method for WSNs was proposed in Yao, Sharma, Golubchik, and Govindan (2010) and has been evaluated on short, long, and constant fault classes. The proposed algorithm is based on the detection of deviations between reference and the collected measurements. The reference time series is built using the linear Segmented Sequence Analysis (SSA) and when the remainder between the reference and measured values is greater than a threshold, an alarm is triggered.

Rule-based, estimation-based, time series analysis and learning-based methods are four methods for fault detection discussed in Sharma, Golubchik, and Govindan (2010). They conduct experiments which investigate various fixed and dynamic thresholds for linear least squares estimation, Auto Regressive Integrated Moving Average (ARIMA), Hidden Markov Model (HMM), etc. No superior class of detection methods was found to be suitable for every type of anomalous event as the accuracy is dependent on both the size and quality of the data. Rule-based methods require precise calibration and tuning threshold parameters, learning methods require training phases, estimation methods are unable to classify faults, and time series analysis has the highest rate of false positives.

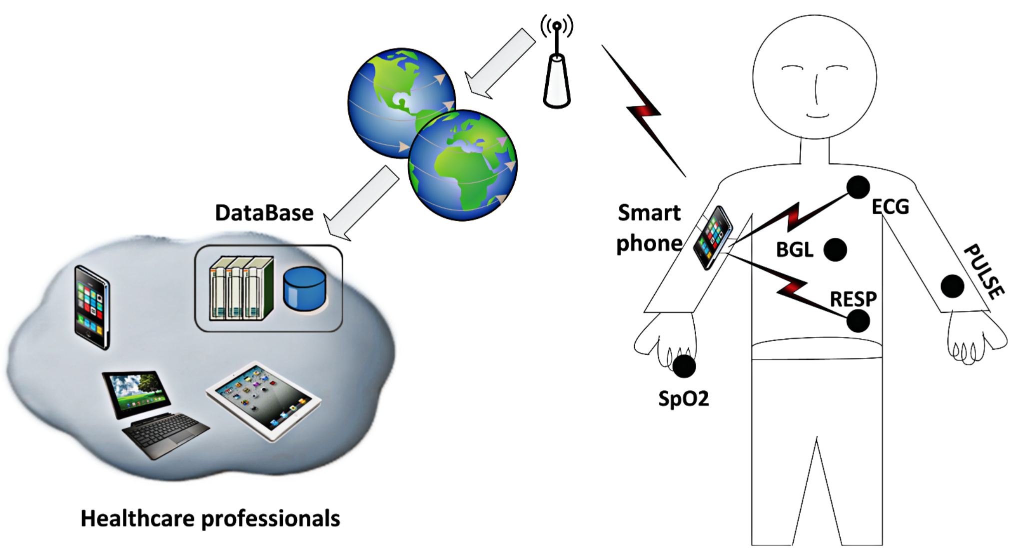
Healthcare applications for patient monitoring require strict reliability on gathered data. Usually, many physiological attributes are monitored in the same time, such as heart rate, blood pressure, respirations, pulse and oxygenation. Alarms set for each attribute are triggered whenever the associated value falls outside a predefined interval. There are however correlations which exist between physiological parameters and the spatio-temporal correlations amongst monitored physiological attributes which may be exploited to detect anomalies and distinguish between faulty sensor measurements and medical emergencies. Faulty sensor readings tend to show irregular, random values unrelated to other attributes in the instance. Once detected, the instances containing the irregularities may be discarded to reduce false alarms, clean the data, and increase the reliability and the accuracy of the monitoring system.

In this paper, we seek to enhance fault detection for current medical WBAN systems. We use SVM and linear regression algorithms to detect abnormal records and to pinpoint abnormal sensors readings in the LPU. SVM is utilized to reduce the temporal complexity and for binary record classification into either normal or abnormal. If a record is classified as abnormal, the linear regression is used to predict values for the current attributes which may also uncover irregular attributes. Our system can detect anomalies in the data of a patient and prior to triggering an alarm, will uniquely distinguish between human physiological abnormalities and sensor node failure utilizing both the spatial and temporal parameters.

3. BACKGROUND

In Figure 1, we consider a patient wearing N medical WBAN sensor nodes $(S_1, S_2, \ldots, S_N)$ to observe specific physiological parameters where the nodes collect and transmit these observations to a smartphone or LPU. The data is gathered by the smartphone where it analysed in real time and is able to send prompt alert notifications to healthcare providers. Further authenticated local and remote data storage may be enabled through and transmitted by the LPU or smartphone. Due to greater resources on board modern smart devices, the LPU analyses and mines the data for irregularities using lightweight

Figure 1. WSN for collecting vital signs and alerting caregivers



machine learning algorithms. It also maintains high accuracy when distinguishing between sensor errors and patient health irregularities and generating the appropriate alarm.

Physiological parameter measurements are collected which we declare as the data matrix $X=\left(X_{ij}\right)$ where i represents the temporal growth and j represents a sensor metric. All gathered values for all parameters are stored as a single record incrementally at time instant k we represent with $X_k = \left(x_{k1}, x_{k2}, \ldots, x_{kn}\right)$. X_k is the line k in the data matrix X given in Equation (1). We also denote by $A = \left(A_1, A_2, \ldots, A_n\right)$ the set of monitored attributes, where A_i is the column i in the matrix X:

$$X = \begin{array}{c} \\ X_1 \\ X_2 \\ \vdots \\ X_m \end{array} \begin{array}{c} \begin{array}{ccccc} A_1 & A_2 & A_3 & \cdots & A_n \end{array} \\ \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \cdots & x_{mn} \end{bmatrix} \end{array} \quad (1)$$

To provide instantaneous online anomaly recognition we process in real time, the collected data on the smart device. These measurements are likely to be of low quality and unreliable due to the sensor hardware constraints and resources as well as the physiological condition of the individual (i.e. bodily sweat, sensor detachment) and the environmental conditions (i.e. sensor damage, fading, disrupted communication). The WBAN monitor is reliant on the received data to maintain its accuracy and robustness, where erroneous instances are identified and it may trigger false alarm notifications when necessary to alert authorized medical personnel. To boost the analysis accuracy and reduce false alarms or misdiagnosis, abnormal instances must be found, analysed, and isolated.

To detect abnormal values, we use Support Vector Machine (SVM) to detect outliers and classify each instance (received attributes at time t) as normal or abnormal. Upon finding an anomaly, we apply a linear regression model to predict values for each attribute in the abnormal instance. When the variance exceeds a predefined threshold, between the predicted and actual value, we analyse data correlations to differentiate faulty sensors from irregular or degrading patient health. We discuss briefly the algorithms used in our approach, SVM and linear regression, in the remaining paragraphs of this section. For more in depth information about the algorithms, please refer to Witten, Frank, and Hall (2011).

3.1. Support Vector Machine

Support Vector Machine (SVM) (Bishop, 2006) is a widely used supervised machine learning method for binary classification which uses the training data to build a model for classification. The SVM then uses this model to classify, using attribute data, each instance in the test set.

The main concept behind linear SVMs is to maximize the distance between two parallel boundaries or hyperplanes which are defined by support vectors:

$$w^T X_i + w_0 = 1 \text{ and } w^T X_i + w_0 = -1 \quad (2)$$

The objective is to construct a separating hyperplane which achieves maximum separation between the 2 classes:

$$w^T X_i + w_0 = 0 \quad (3)$$

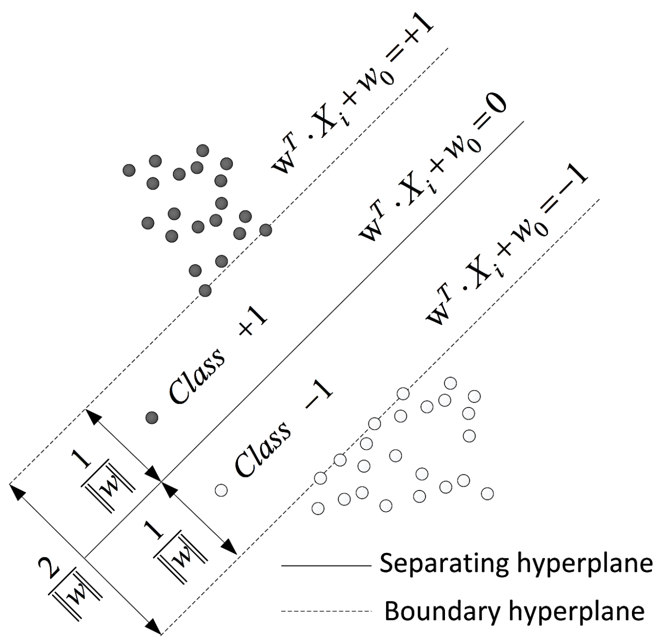
When generating the model for classification, SVM looks for the maximum margin hyperplane which divides the training data into two categories. Given the training data set $X = \left(X_1, X_2, \ldots, X_T\right)$, with their associated class $y_i \in \left\{-1,1\right\}$ (-1 for abnormal & 1 for normal or healthy patient), for $i \in \left(1,\ldots,T\right)$, the hyperplane is the solution of the optimization problem shown in Box 1.

The margin width (or distance) is equal to $2 / \left\|w\right\|$ as shown in Figure 2. The margin er-

Box 1.

$$\begin{array}{lll} \underset{w,w_0}{Minimize} & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\xi_i & \\ Subject\ to & w^T X_i + w_0 \geq 1 - \xi_i & if\ X_i \in y_1 \\ & w^T X_i + w_0 \leq -1 + \xi_i & if\ X_i \in y_2 \\ & \xi \geq 0 & \end{array} \tag{4}$$

Figure 2. Linear SVM and data separation



rors, ξ_i, are used to prevent over fitting problems, and are positive for points inside the margin, or outside the margin on the wrong side of the classifier, and 0 for points in the correct side of the classifier. C is a user-defined constant. SVM uses the coordinates of the nearest training data points in both classes in order to create the largest possible separation between border values in each class. Those specific data points are called the support vectors and they have to satisfy:

$$y_i(w^T \cdot X_i + w_0) = 1 \tag{5}$$

where w is the normal vector to the hyperplane, and w_0 is the bias of the hyperplane function. The normal vector w is calculated using the values in the training data set:

$$w = \sum_{i=1}^{n} \alpha_i y_i X_i \tag{6}$$

where α_i are the Lagrange multipliers of the optimization task, and they are different than zero only for points outside the margin and inside the correct side of the classifier. The classification of X_i is based on the sign of $h\left(X_i\right)$:

$$\begin{aligned} y_i &= sign\left(h(X_i)\right) \\ &= sign\left(w^T \cdot X_i + w_0\right) \end{aligned} \tag{7}$$

3.2. Linear Regression

The linear regression (Bishop, 2006; Witten, Frank, & Hall, 2011) is a statistical modeling method used to predict the current value of monitored attribute. For a given attribute A_j, it exploits spatial correlation to predict the current value $(\hat{x}_{ij})$, as linear combination of measured values for other attributes $x_{ik|k \neq j}$. The model of the predicted attribute value is given by:

$$\begin{aligned} \hat{x}_{ij} &= a_0 + a_1 x_{i1} \\ &+ a_2 x_{i2} + \cdots + a_n x_{in} \end{aligned} \tag{8}$$

where a_k are the coefficients of the regressors (weights). These coefficients are obtained during the training phase as the result of division of the covariance of A_i and A_j attributes, on the variance of A_j:

$$\begin{aligned} a_k &= \frac{Cov(A_i, A_j)}{Var(A_i)} \\ &= \frac{\sum_k \left(x_{ki} - \bar{A}_i\right)\left(x_{kj} - \bar{A}_j\right)}{\sum_k (x_{ki} - \bar{A}_i)} \end{aligned} \tag{9}$$

Once the model is computed from training data, it is used to predict the value of each attribute $(\hat{x}_{ij})$ at instance i. Afterward, we compare the predicted value $(\hat{x}_{ij})$ with the actual value (x_{ij}) to find if it fits within a small margin error and to classify p_i as normal or abnormal.

4. PROPOSED APPROACH

We consider a general scenario for remote patient monitoring, as shown in Figure 1, where many wireless motes with restricted resources are used to collect data, and a portable collection device (e.g. smart phone) with higher resources and higher transmission capabilities than motes, is used to analyze collected data, and to raise alarms for emergency team when abnormal patterns are detected. We seek to detect abnormal values in order to reduce false alarms resulted from faulty measurements, while differentiating faults from patient health degradation.

The proposed approach is based on decision tree and linear regression. It builds a decision tree and looks for linear coefficients from normal vital signs that fall inside restricted interval range of monitored attributes. In the rest of this paper, we focus only on the following vital signs:

$$BP \in \left[60 - 110\right]$$

$$HR \in \left[60 - 100\right]$$

$$pulse \in \left[60 - 100\right]$$

$$respiration\, rate \in \left[12 - 30\right]$$

$$SpO2 \in \left[90 - 100\right]$$

Attributes values that fall outside these (restricted) normal intervals are considered abnormal. HR and pulse reflect the same attribute from different sensors, where pulse is obtained from the pulse oximeter and HR is measured as the number of interbeat intervals (R-R) in ECG signal.

The proposed approach is based on two phases: training and detection. In the training phase, we build the classification models for SVM and linear regression methods, and in the testing phase, inputs are classified as abnormal if they deviate from established model. The linear SVM is used in our approach to classify each received record as normal or abnormal. The SVM is used due to its accuracy and low complexity, where the classification requires only the sign of $h\left(X_i\right)$ in Equation (7).

Abnormal instances detected by SVM will only activate the forecasting procedure using the linear regression, where we recursively assume that an attribute $\left(x_{ik}\right)$ is missing, and the coefficients of linear regression are used to estimate the current value for this attribute ($\hat{x}_{ik}$) with respect to the others $\left(x_{ij|j\neq k}\right)$ as given in Equation (10) for heart rate estimation:

$$\widehat{HR}_i = C_0 + C_1 Pulse_i + C_2 RESP_i + \cdots + C_5 BP_i \qquad (10)$$

If the Euclidean distance between current $\left(HR_i\right)$ and estimated $\left(\widehat{HR}_i\right)$ values is larger than the predefined threshold (10% of estimated value) for only one attribute, the measurement is considered faulty and replaced by estimated value with linear regression. Equation (11) shows the residual threshold used to detect abnormal measurement:

$$e_i = \left|x_{ik} - \hat{x}_{ik}\right| \geq 0.1 * \hat{x}_{ik} \qquad (11)$$

However, if at least k readings are higher than the threshold, we trigger a medical alarm for response caregiver emergency team to react, e.g. heavy changes in the HR and reduced rate of SpO2 are symptoms of patient health degradation and requires immediate medical intervention. We assume that the probability of many attributes ($k = 2$ in our experiments) being faulty is very low. The pseudo-code for our proposed algorithm is given in Algorithm 1.

The SVM is used to reduce the computation complexity, and to prevent the estimation of each attribute for each instance on the base station. SVM is based on sign comparison for classification, and the combination of both approach for fault detection and classification is used. Sliding window is not used in our experiments to reduce the complexity. When the model is specified with the training data, updating or rebuilding the model requires additional complexity (temporal & spatial) and reduce the accuracy of the classification model. Most of the time, the gathered measurements are normal, and updating the classification model using skewed data (normal data only for training) leads to erroneous classification.

5. EXPERIMENTAL RESULTS

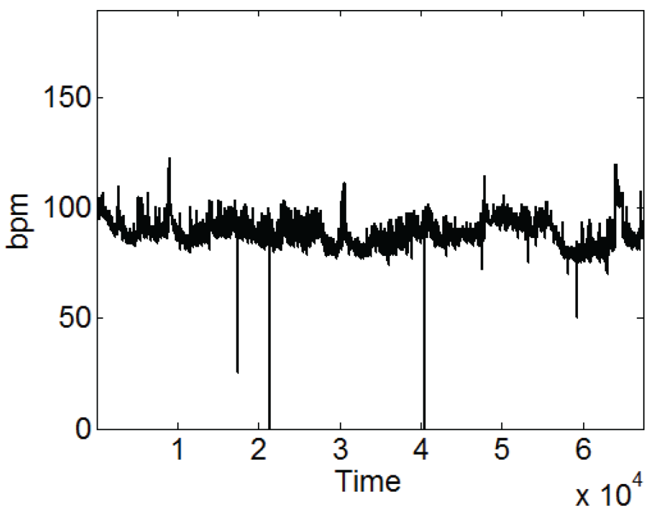
In this section, we present the performance analysis results of the proposed approach for anomaly detection in medical WSN. Afterward, we conduct analysis to study the impact of decision threshold on true positive and false alarm ratio. We used real medical data set from the Physionet database (Physionet, 2013), which contains 7 attributes (ABPmean, ABPsys, ABPdias, HR, PULSE, RESP and SpO2). We only focus on 5 attributes in each record: ABPmean, HR, PULSE, RESP and SpO2.

The variations of Heart Rate (in beats per minute — bpm) are shown in Figure 3. We take notice of 4 abnormal measurements (spikes) where 2 between them falling down to zero. Other variations associated with a clinical change of the monitored patient can be clearly distinguished in Figure 3. The variation of Blood Pressure (in millimeters of mercury — mmHg), Pulse (in bpm), respiration rate (in respirations per minutes — rpm), and oxygenation ratio (in percentage) are presented in Figures 4, 5, 6 and 7 respectively.

Algorithm 1. Detection algorithm

for each received record X_i during T **do**
  Classify X_i using SVM
  if $Class(X_i) == "ABNORMAL"$ **then**
    for each x_{ik} **do**
      $\hat{x}_{ik} = \sum_{j=1, j \neq k}^{n} a_j . x_{ij}$
      $ctr += \left(\left(\left| x_{ik} - \hat{x}_{ik} \right| \right) \geq 0.1 \times \hat{x}_{ik} \right) ? 1 : 0$
    end for
    if $(ctr \geq k)$ **then**
      Raise alarm for healthcare
    end if
  end if
end for

Figure 3. Heart rate



In fact, HR and Pulse measure the same physiological parameter using two different devices, and usually they must present the same variations. However, when comparing Figures 3 and 5, they exhibit some differences especially for spikes at different time instant. The difference results from abnormal values reported by the sensor.

To prove the correlation between monitored attributes, we show the variation curves of the 5 parameters in Figure 8, where we can notice that clinical emergency induces changes in many parameters at the same time instant. However, there is no spatial correlation among monitored attributes for faulty measurements. It is important to note that some curves in Figure 8 are

Figure 4. Blood pressure

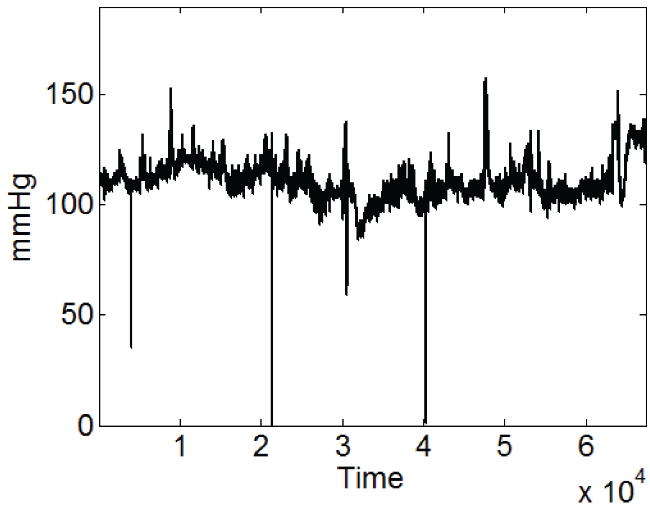
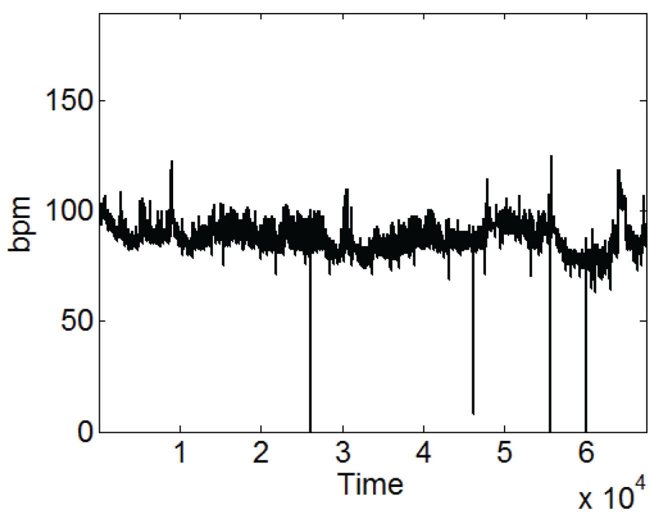


Figure 5. PULSE



shifted to clarify the shape of their variations. We can visually distinguish zones of clinical change, where many attributes change at the same time instant.

As physiological parameters vary by individual and they are dependent on many physical characteristics (sex, age, weight, activity, etc.), the use of a static interval for anomaly detection is heavily reliant on additional dynamic parameters (environmental, ages, activities, etc.) which are difficult to set dynamically.

Figures 9 and 10 respectively show the predicted and error values for HR using linear regression. The measured values of HR (actual) are presented in Figure 3. The error represents the difference between actual and predicted values of HR. To test the efficiency of the used algorithms, we compare the results

Figure 6. Respiration rate

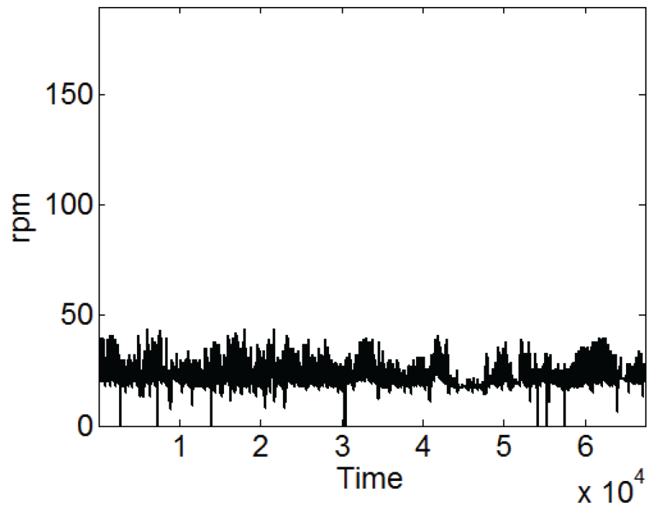
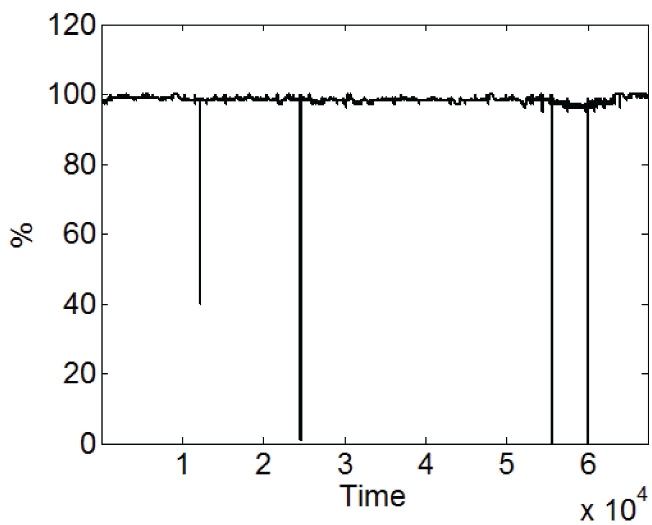


Figure 7. Oxygenation ratio



(predicted and error) with different classifiers using the WEKA (Hall, et al., 2009; WEKA, 2013) toolkit: Decision Stump, Decision Table, Additive Regression and K-NN for $K = 1$.

Figures 11 and 12 show similar results (predicted and error respectively) using additive regression tree, where the prediction error is higher than linear regression. Figures 13 and 14 show the results of the decision stump classifier. Figures 15 and 16 show the results of the decision stump classifier. The results using KNN, which has a slower runtime due to the greater computational complexity, are shown in Figures 17 and 18 to have a lower error rate in comparison to additive regression, decision stump and decision table.

Figure 23 shows the mean absolute error for each of these classifiers, where decision

Figure 8. Variations of the 5 parameters

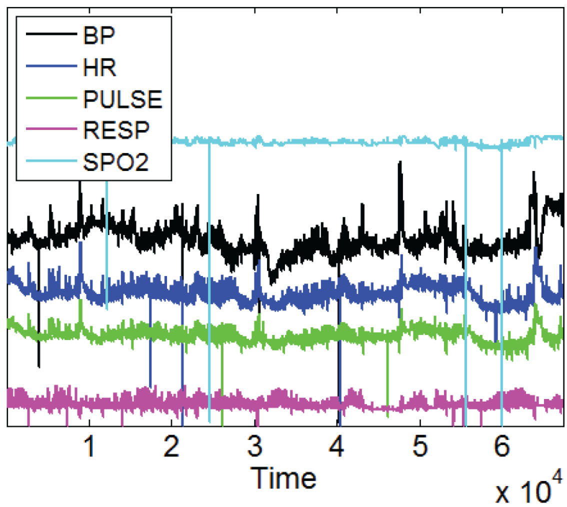


Figure 9. Predicted heart rate using linear regression

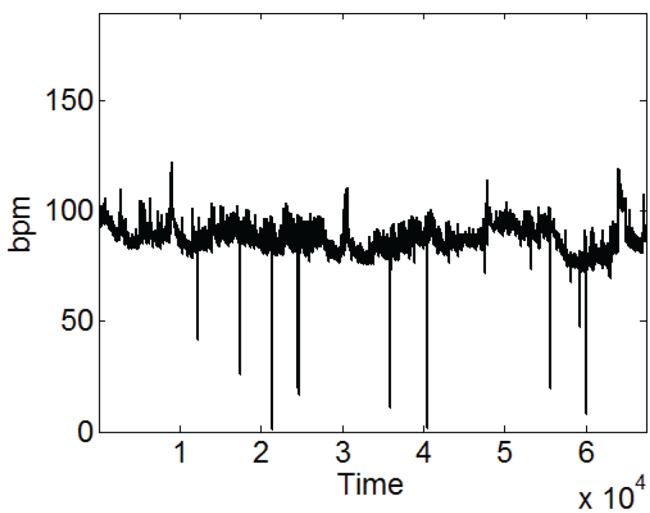


table achieves the prediction with the highest mean error rate, followed in descending order by decision stump, additive regression, K-NN and linear Regression. A slight difference, in terms of the mean prediction error, between K-NN and additive regression is presented in Figure 23. During the experiment, we discover that the result of additive regression sometimes is better than K-NN when using other data set. That is to say, the accuracy of the prediction algorithm depends also on the data in training phase. Linear regression had the lowest error percentage and the best overall performance out of the four classifiers, which is also why we use this classifier in the rest of this paper.

Figure 19 shows the raised alarms by the application using SVM. In our previous work (Salem, Guerassimov, Mehaoua, Marcus, &

Figure 10. Prediction error using linear regression

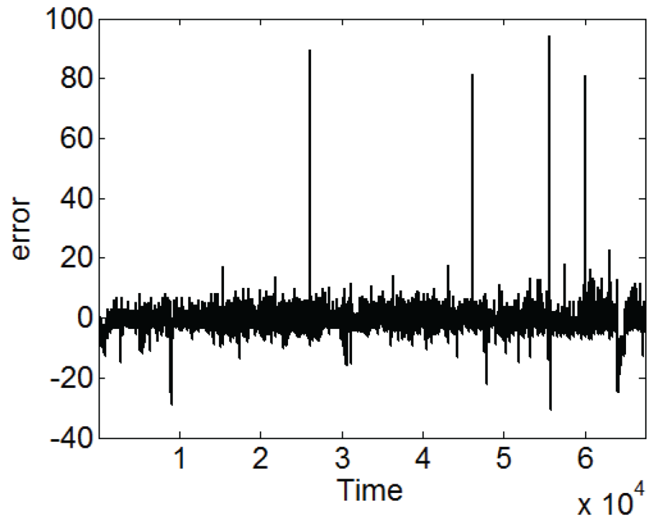
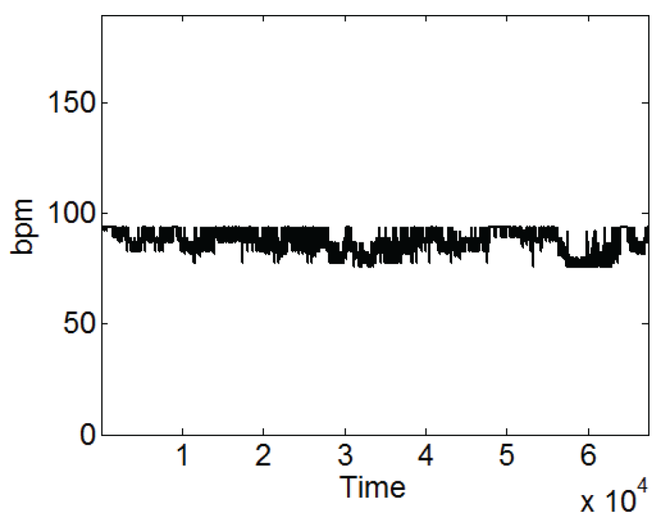


Figure 11. Predicted heart rate using additive regression



Furht, 2013), our approach applied J48 on real patient data. To compare the performance of both classifiers, the alarms triggered by J48 are shown in Figure 20. These results confirm that SVM slightly outperforms J48 in terms of detection accuracy, but J48 builds the classification model (decision tree) faster than SVM.

We used the Receiver Operating Characteristic (ROC) in the performance evaluation of the proposed approach to show the relationship

Figure 12. Prediction error using additive regression

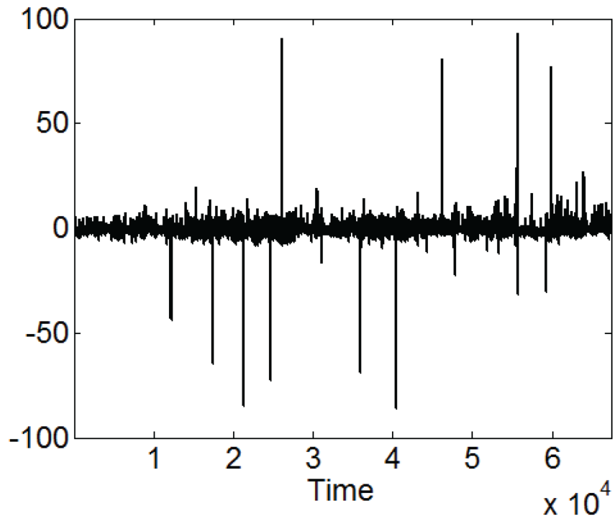
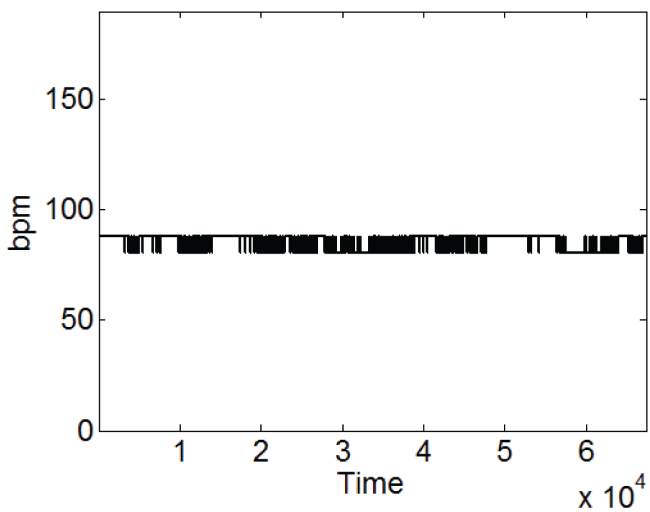


Figure 13. Predicted heart rate using decision stump



between the true positive rate (Equation (12)) and the false positive rate (Equation (13)).

$$TPR = \frac{TP}{TP + FN} \tag{12}$$

where TP represents the number of true positives, and FP is the number of false positives. The false positive rate (FPR) is defined as:

$$FPR = \frac{FP}{FP + TN} \tag{13}$$

ROC curves are used for accuracy analysis where it represents, graphically, the true positive

Figure 14. Prediction error using decision stump

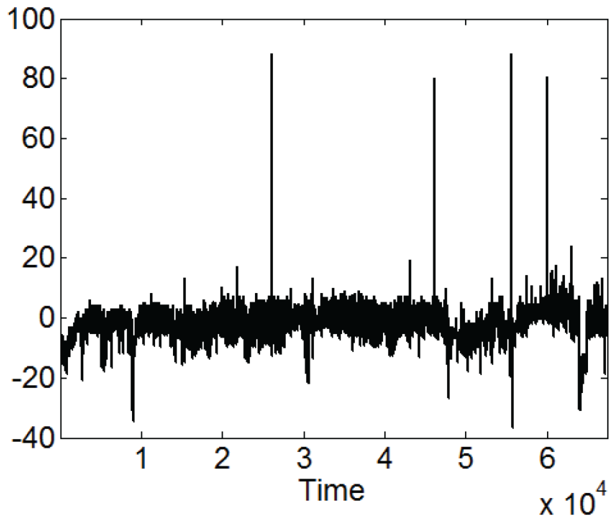
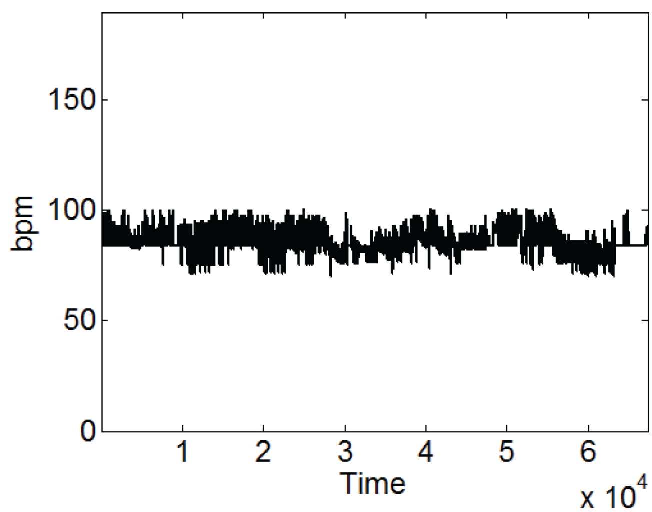


Figure 15. Predicted heart rate using decision table



rate versus the false positive rate when varying the value of the decision threshold. In general, a good detection algorithm must achieve a high detection ratio with the lowest false alarm rate.

Figure 21 shows the raised medical alarms by the proposed approach. The raised alarms are triggered by heavy changes in at least k attributes. We can clearly notice in Figure 21 that faulty measurements (spikes), without correlated changes between physiological parameters (BP, HR, Pulse, SpO2 and RESP), don't trigger medical alarms. Figure 22 shows the raised alarms while replacing SVM by J48 in the proposed approach. The results are similar, and the raised alarms occur at the same time instants. We can notice some differences

Figure 16. Prediction error using decision table

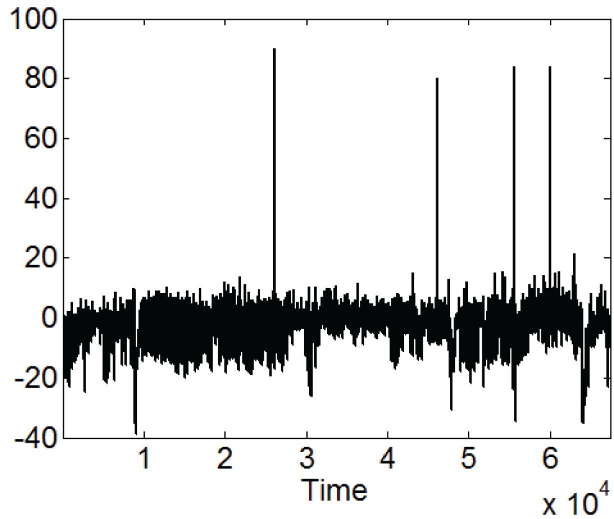
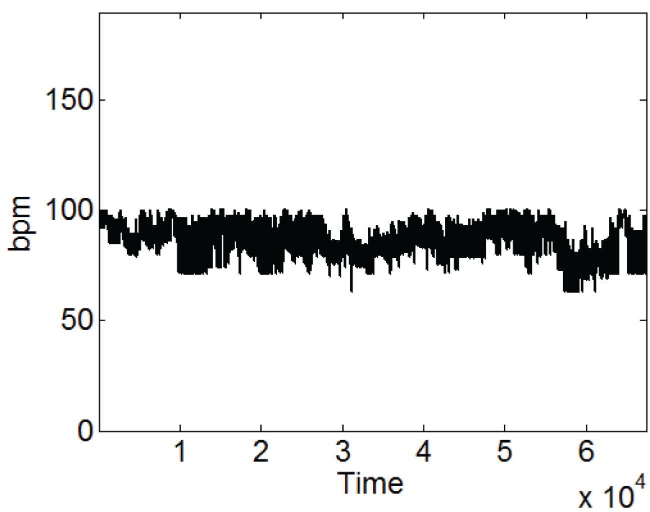


Figure 17. Predicted heart rate using K-NN



between Figures 21 and 22, where we get two additional alarms by J48 in Figure 22 (before 3.10^4 and 4.10^4 respectively).

Figure 24 shows the ROC for the proposed approach where the first nominal classifier is SVM, followed by J48, Logistic regression, Naïve Bayes and Decision Table respectively. SVM and J48 classifiers were the two most accurate algorithms which achieved the best performances with TPR=100% for both, FPR=6.5% for SVM and FPR=7.4% for J48 respectively. The ROC for SVM and J48 is very similar having only minor difference. However, they achieve better performance compared to the other classifiers. The ROC validates that our claim that the proposed approach achieves high accuracy for detecting mote anomalies.

Figure 18. Prediction error using K-NN

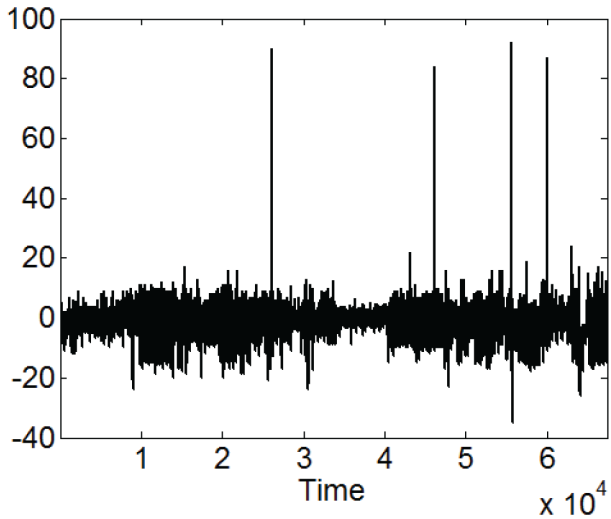
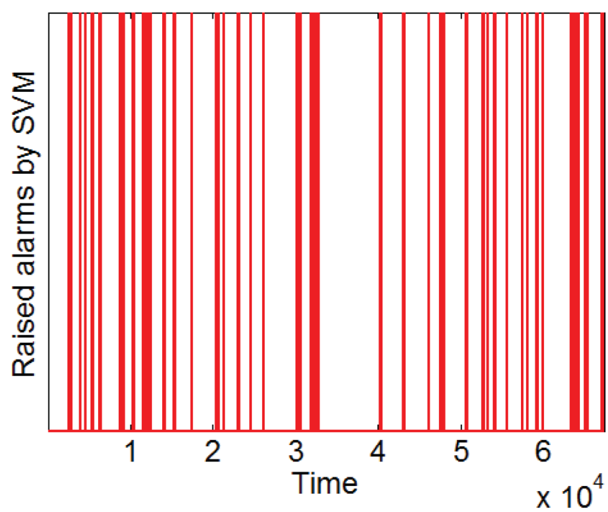


Figure 19. Raised alarms by SVM



6. CONCLUSION AND PERSPECTIVES

Medical WBAN is a new emerging technology in the field of healthcare, providing vital care and access to patients, elderly, and infants. It allows continuously monitoring patients without restrictions in the movements and keeping the healthcare professional informed of any evolution of patients' condition.

These types of monitoring systems are tasked with providing humanity an outstand-

Figure 20. Raised alarms by J48

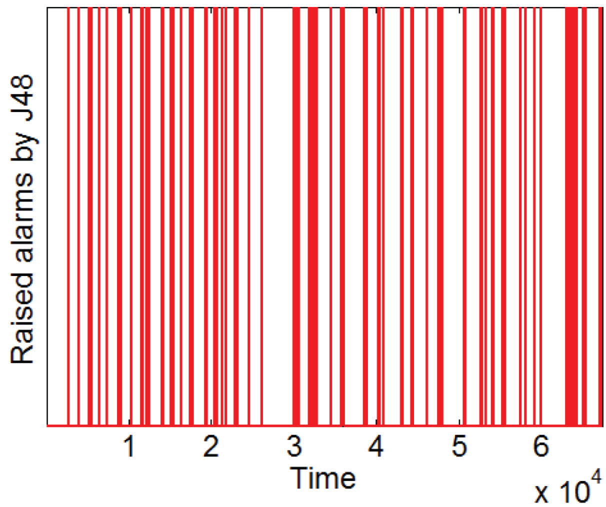
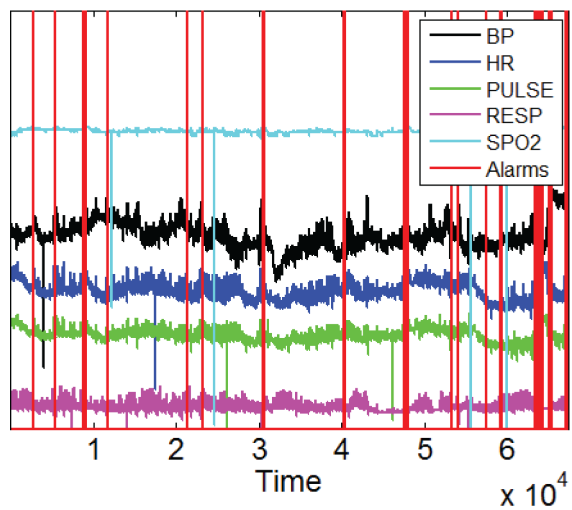


Figure 21. Raised medical alarms by SVM and linear regression



ing instrument for patient observation and autonomous diagnostic, alarm, and emergency services. They also provide simple, remote patient data management and allow greater freedom for healthcare professionals which may, as a result, better serve clients from practically any location where network connectivity exists.

We described our architecture conceptually and detailed the results from preliminary experimentation conducted with our applications'

Figure 22. Raised medical alarms by J48 and linear regression

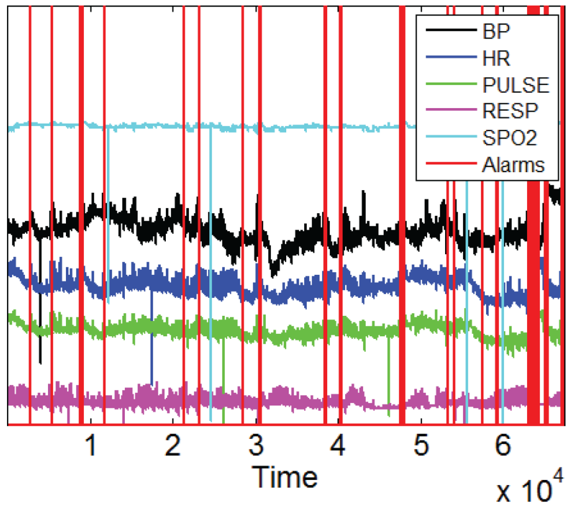
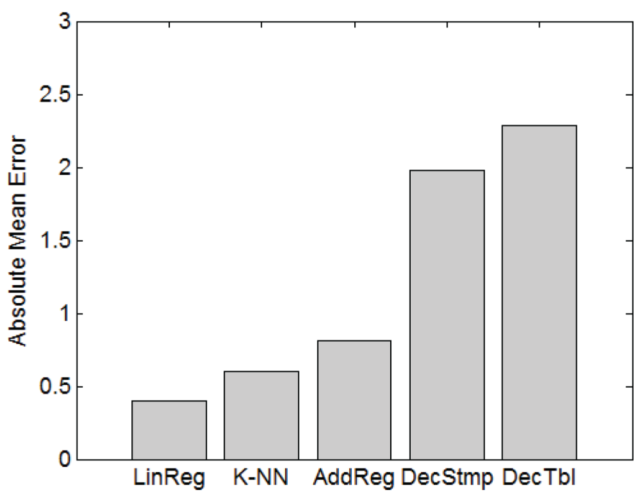


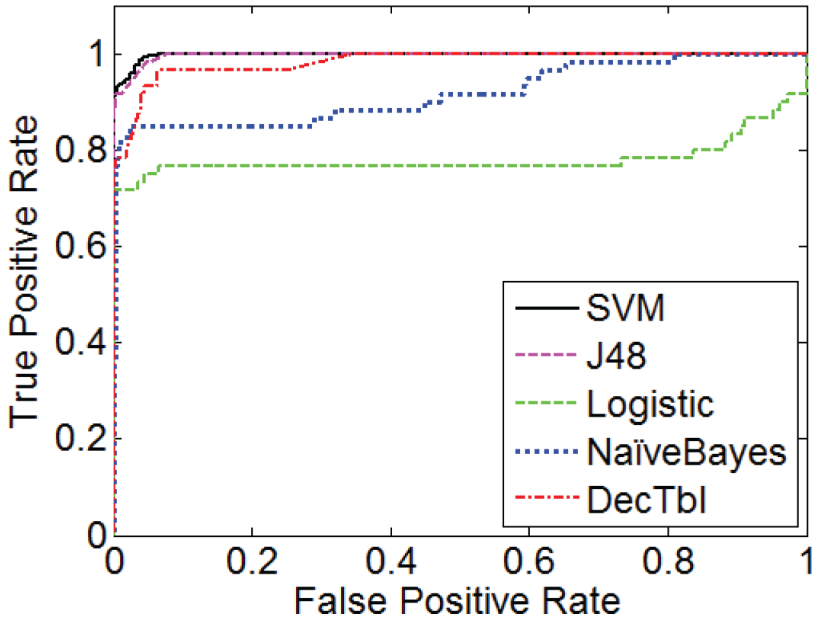
Figure 23. Mean error rate with different classifiers



analysis of real patient data. Furthermore, we discussed the issues and justified the need for medical WBANs for ubiquitous patient monitoring and authenticated remote patient data access for healthcare professionals.

The application, after mining the incoming data incrementally, applies machine learning algorithms to generated models based on algorithmically located correlations in the data. It is also able to distinguish between irregular patient

Figure 24. Receiver Operating Characteristic (ROC)



attributes and faulty sensor data to maintain robustness and high accuracy. Considering the limitations associated with WSN technology and the spontaneity of our environment, we have constructed a reliable application for real time global patient monitoring from modern smart devices. The experimental results confirm the applications' high detection accuracy, low false alarm ratio and its' ability to quickly identify and differentiate between sensor faults and irregularities in a patient's health.

In the future, knowing that most collected sensor measurements are normal, we look to reduce the amount of exchanged data between the wireless sensors and the sink node, by transmitting only abnormal values on the sensor motes to reduce energy consumption by wireless transmissions.

ACKNOWLEDGMENT

This research was supported by Korea Science and Engineering Foundation, under the World Class University (WCU) program with additional support from NSF grants CCF-0545488 and OISE-0730065.

REFERENCES

Abduvaliyev, A., Pathan, A.-S. K., Zhou, J., Roman, R., & Wong, W.-C. (2013). On the vital areas of intrusion detection systems in wireless sensor networks. *IEEE Communications Surveys & Tutorials*, *99*, 1–15.

Alemdar, H., & Ersoy, C. (2010). Wireless sensor networks for healthcare: A survey. *Computer Networks*, *54*(15), 2688–2710. doi:10.1016/j.comnet.2010.05.003

Banerjee, T., Xie, B., & Agrawal, D. P. (2008). Fault tolerant multiple event detection in a wireless sensor network. *Journal of Parallel and Distributed Computing*, *68*(9), 1222–1234. doi:10.1016/j.jpdc.2008.04.009

Bishop, C. (2006). *Pattern recognition and machine learning (Information science and statistics)*. Springer-Verlag New York, Inc.

Chen, Y.-C., & Juang, J.-C. (2012). Outlier-detection-based indoor localization system for wireless sensor networks. *International Journal of Navigation and Observation*, 2012.

Cheng, X., Xu, J., Pei, J., & Liu, J. (2010). Hierarchical distributed data classification in wireless sensor networks. *Computer Communications*, *33*(12), 1404–1413. doi:10.1016/j.comcom.2010.01.027

Chipara, O., Lu, C., Bailey, T. C., & Roman, G.-C. (2010). Reliable clinical monitoring using wireless sensor networks: Experiences in a step-down hospital unit. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems (SenSys'10)* (pp. 155-168).

Choi, J., Ahmed, B., & Gutierrez-Osuna, R. (2012). Developpement and evaluation of an ambulatory stress monitor based on wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, *16*(2), 279–286. doi:10.1109/TITB.2011.2169804 PMID:21965215

Cunha, J. P., Cunha, B., Pereira, A. S., Xavier, W., Ferreira, N., & Meireles, L. (2010). Vital-Jacket®: A wearable wireless vital signs monitor for patients' mobility in cardiology and sport. In *Proceedings of the International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth.*

Curiac, D.-I., & Volosencu, C. (2012). Ensemble based sensing anomaly detection in wireless sensor networks. *Expert Systems with Applications*, *39*(10), 9087–9096. doi:10.1016/j.eswa.2012.02.036

Farruggia, A., Giuseppe, L. R., & Ortolani, M. (2011). Probabilistic anomaly detection for wireless sensor networks. In *Proceedings of the 12th International Conference on Artificial Intelligence Around Man and Beyond* (pp. 438-444).

Grgic, K., Žagar, D., & Križanovic, V. (2012). Medical applications of wireless sensor networks — current status and future directions. *Medicinski Glasnik*, *9*(1), 23–31. PMID:22634904

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *SIGKDD Explorations*, *11*(1). doi:10.1145/1656274.1656278

Havard Sensor Networks Lab. (2013). *CodeBlue: Wireless sensors for medical care*. Retrieved from http://fiji.eecs.harvard.edu/CodeBlue

Huang, F., Jiang, Z., Zhang, S., & Gao, S. (2010). Reliability evaluation of wireless sensor networks using logistic regression. In *Proceedings of the 2010 International Conference on Communications and Mobile Computing (CMC'10)* (pp. 334-338).

Jurdak, R., Wang, X. R., Obst, O., & Valencia, P. (2011). Wireless sensor network anomalies: Diagnosis and detection strategies. In D. A. Tolk, & L. C. Jain (Eds.), *Intelligence-based systems engineering* (pp. 309–325). Springer. doi:10.1007/978-3-642-17931-0_12

Ko, J., Lim, J., Chen, Y., Musvaloiu-E, R., Terzis, A., & Masson, G. et al. (2010). MEDiSN: Medical emergency detection in sensor networks. [TECS]. *ACM Transactions on Embedded Computing Systems*, *10*(1), 1–29. doi:10.1145/1814539.1814550

Kumar, P., & Lee, H.-J. (2012). Security issues in healthcare applications using wireless medical sensor networks: A survey. *Sensors (Basel, Switzerland)*, *12*(1), 55–91. PMID:22368458

Li, Y. (2010). *Anomaly detection in unknown environments using wireless sensor networks*. University of Tennessee.

Liu, F., Cheng, X., & Chen, D. (2007). Insider attacker detection in wireless sensor networks. *Proceedings - IEEE INFOCOM*, *07*, 1937–1945. doi:10.1109/INFCOM.2007.225

Malan, D., Fulford-jones, T., Welsh, M., & Moulton, S. (2004). CodeBlue: An ad hoc sensor network infrastructure for emergency medical care. In *Proceedings of International Workshop on Wearable and Implantable Body Sensor Networks.*

Miao, X., Liu, K., He, Y., Liu, Y., & Papadias, D. (2011). Agnostic diagnosis: discovering silent failures in wireless sensor networks. *Proceedings - IEEE INFOCOM*, *11*, 1548–1556.

Montgomery, K., Mundt, C., Thonier, G., Thonier, A., Udoh, U., Barker, V., et al. (2004). Lifeguard – A personal physiological monitor for extreme environments. In *Proceedings of the IEEE 26th Annual International Conference on Engineering in Medicine and Biology Society* (pp. 2192-2195).

Navarro, K. F., Lawrence, E., & Lim, B. (2009). Medical MoteCare: A distributed personal healthcare monitoring system. In *Proceedings of the International Conference on eHealth, Telemedicine, and Social Medicine (eTELEMED'09)* (pp. 25-30).

Otto, C., Milenkovi, A., Sanders, C., & Jovanov, E. (2005). System architecture of a wireless body area sensor network for ubiquitous health monitoring. *Journal of Mobile Multimedia*, *1*(4), 307–326.

Physionet. (2013). Retrieved from http://www.physionet.org/cgi-bin/atm/ATM

Rajasegarar, S., Leckie, C., Bezdek, J. C., & Palaniswami, M. (2010). Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks. *IEEE Transactions on Information Forensics and Security*, *5*(3), 518–533. doi:10.1109/TIFS.2010.2051543

Sahoo, P. (2012). Efficient security mechanisms for mHealth applications using wireless body sensor networks. *Sensors (Basel, Switzerland)*, *12*(9), 12606–12633. doi:10.3390/s120912606 PMID:23112734

Salem, O., Guerassimov, A., Mehaoua, A., Marcus, A. M., & Furht, B. (2013). Sensor fault and patient anomaly detection and classification in medical wireless sensor networks. In *Proceedings of the IEEE ICC 2013 (ICC'13)* (pp. 4338-4343).

Sharma, A. B., Golubchik, L., & Govindan, R. (2010). Sensor faults: Detection methods and prevalence in real-world datasets. *ACM Transactions on Sensor Networks*, *6*(3), 1–39. doi:10.1145/1754414.1754419

Siripanadorn, S., Hattagam, W., & Teaumroong, N. (2010). Anomaly detection in wireless sensor networks using self-organizing map and wavelets. *International Journal of Communication*, *4*(3), 74–83.

Theodoridis, S., Pikrakis, A., Koutroumbas, K., & Cavouras, D. (2010). *Introduction to pattern recognition: A Matlab approach*. Academic Press.

Wang, H., Fang, H., Xing, L., & Chen, M. (2011). An integrated biometric-based security framework using wavelet-domain HMM in wireless body area networks (WBAN). In *Proceedings of the IEEE International Conference on Communications (ICC'11)* (pp. 1-5).

WEKA. (2013). The university of Waikato. *WEKA data mining tool*. Retrieved from http://www.cs.waikato.ac.nz/~ml/weka/

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques* (3rd ed.). Morgan Kaufmann Publishers Inc.

Wood, A., Virone, G., Doan, T., Cao, Q., Selavo, L., Wu, Y., & Stankovic, J. (2006). *ALARM-NET: Wireless sensor networks for assisted-living and residential monitoring*. University of Virginia.

Xiaozhen, Y., Hong, X., & Tong, W. (2011). A multiple linear regression data predicting method using correlation analysis for wireless sensor networks. In *Proceedings of the Cross Strait Quad-Regional Radio Science and Wireless Technology Conference* (pp. 960-963).

Xie, M., Hu, J., Han, S., & Chen, H.-H. (2012). Scalable hyper-grid k-NN-based online anomaly detection in wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, *99*, 1–11.

Xu, S., Hu, C., Wang, L., & Zhang, G. (2012). Support vector machines based on K nearest neighbor algorithm for outlier detection in WSNs. In *Proceedings of the 8th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM'12)* (pp. 1-4).

Yang, X., Dinh, A., & Chen, L. (2010). Implementation of a wearerable real-time system for physical activity recognition based on naïve Bayes classifier. In *Proceedings of the International Conference on Bioinformatics and Biomedical Technology (ICBBT'10)* (pp. 101-105).

Yao, Y., Sharma, A., Golubchik, L., & Govindan, R. (2010). Online anomaly detection for sensor systems: A simple and efficient approach. *Performance Evaluation*, *67*(11), 1059–1075. doi:10.1016/j.peva.2010.08.018

Yim, S.-J., & Choi, Y.-H. (2010). An adaptive fault-tolerant event detection scheme for wireless sensor networks. *Sensors (Basel, Switzerland)*, *10*(3), 2332–2347. doi:10.3390/s100302332 PMID:22294929

Ying-xin, X., Xiang-guang, C., & Jun, Z. (2011). Data fault detection for wireless sensor networks using multi-scale PCA method. In *Proc. of International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC'11)* (pp. 7035-7038).

Zhang, Y., Chao, H.-C., Chen, M., Shu, L., Park, C.-h., & Park, M.-S. (2010). Outlier detection and countermeasure for hierarchical wireless sensor networks. *IET Information Security*, *4*(4), 361–373. doi:10.1049/iet-ifs.2009.0192

Zhang, Y., Hamm, N., Meratnia, N., Stein, A., Van de Voort, M., & Havinga, P. (2012). Statistics-based outlier detection for wireless sensor networks. *26*(8), 1373-1392.

Zhang, Y., Meratnia, N., & Havinga, P. (2009). Adaptive and online one-class support vector machine-based outlier detection techniques for wireless sensor networks. In *Proceedings of the International Conference on Advanced Information Networking and Applications Workshops* (pp. 990-995).

Zhang, Y., Meratnia, N., & Havinga, P. J. (2010). Outlier detection techniques for wireless sensor networks: A survey. *IEEE Communications Surveys and Tutorials*, *12*(2), 159–170. doi:10.1109/SURV.2010.021510.00088

Osman Salem graduated in computer network engineering in 2001. He received his PhD from Paul Sabatier University, Toulouse, France, in 2006. He was a researcher in the department of computer science at Telecom Bretagne (2006-2007). Since September 2008, he has been an Associate Professor at University of Paris Descartes. His research activities includes security and anomaly detection in medical wireless sensor networks.

Alexey Guerassimov born in Moscow in 1988, he is a master degree student in network and telecommunication security at the University of Paris Descartes. His main research interests are on the anomaly detection in medical wireless sensor networks. He performed various internships especially in the IT laboratory of Paris Descartes (LIPADE), Paris, France, and the Division of IT Convergence Engineering (ITCE) in the University of POSTECH, Pohang, South Korea.

Ahmed Mehaoua is currently a professor at the Univ. of Paris Descartes, and the Head of the Multimedia Communications and Security Group at the LIPADE. He was an associate professor at the Univ. of Versailles, and research fellow in the Univ. of Cambridge. He was with the Computer Science Research Institute of Montreal, Canada, and was involved in various R&D projects. His research interests include the QoS optimization for audiovisual communications and WBANs for medical and health applications.

Anthony Marcus was born in New Haven, Connecticut in 1970. After serving the Military Police in the U.S. Army for eight years, Anthony earned a B.S. in Computer Science, an M.S. in Computer Engineering and is currently pursuing a Ph. D. in Computer Science from Florida Atlantic University. His research interests include wireless and sensor networking technologies, mobile application development and data mining and machine learning.

Borko Furht is a professor and chairman of the Department of Electrical & Computer Engineering and Computer Science at Florida Atlantic University (FAU) in Boca Raton, Florida. He is also Director of the NSF-sponsored Industry/University Cooperative Research Center on Advanced Knowledge Enablement. Before joining FAU, he was a vice president of research and a senior director of development at Modcomp (Ft. Lauderdale), a computer company of Daimler Benz, Germany, a professor at University of Miami in Coral Gables, Florida, and a senior researcher in the Institute Boris Kidric, Yugoslavia.