

Anomaly Detection in Medical Wireless Sensor Networks

Osman Salem*

LIPADE Laboratory, University Paris Descartes, Paris, France
osman.salem@parisdescartes.fr

Yaning Liu

JCP-Consult, Cesson-Sevigne, France
yaning.liu@jcp-consult.com

Ahmed Mehaoua

Division of IT Convergence Engineering, POSTECH, Pohang, Korea ahmed@postech.ac.kr
LIPADE Laboratory, University Paris Descartes, Paris, France ahmed.mehaoua@parisdescartes.fr
Centre National de la Recherche Scientifique (CNRS), LaBRI, Bordeaux, France ahmed.mehaoua@labri.fr

Abstract

In this paper, we propose a new framework for anomaly detection in medical wireless sensor networks, which are used for remote monitoring of patient vital signs. The proposed framework performs sequential data analysis on a mini gateway used as a base station to detect abnormal changes and to cope with unreliable measurements in collected data without prior knowledge of anomalous events or normal data patterns. The proposed approach is based on the Mahalanobis distance for spatial analysis, and a kernel density estimator for the identification of abnormal temporal patterns. Our main objective is to distinguish between faulty measurements and clinical emergencies in order to reduce false alarms triggered by faulty measurements or ill-behaved sensors. Our experimental results on both real and synthetic medical datasets show that the proposed approach can achieve good detection accuracy with a low false alarm rate (less than 5.5%).

Category: Embedded computing

Keywords: Healthcare monitoring; Wireless sensor networks; Security; Anomaly detection; Fault detection; Mahalanobis distance

I. INTRODUCTION

In medical applications, implementations of specialized wireless sensor networks (WSNs), known as wireless body area networks (WBANs), are composed of numerous small wireless devices attached to or implanted in the body of a patient to collect various vital signs and

to transmit collected data to a central device (i.e., base station, smart phone, etc.) for processing. It triggers medical alarms for emergency medical services upon detection of anomalies in the gathered physiological data, to quickly react by taking the appropriate actions [1, 2]. This allows real-time monitoring and early detection of clinical deterioration [2-4].

Open Access <http://dx.doi.org/10.5626/JCSE.2013.7.4.272>

<http://jcse.kiise.org>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 31 July 2013, Accepted 3 October 2013

*Corresponding Author

The medical applications of WSNs are closely related to vital-sign monitoring and real-time detection of life-threatening emergencies within a few seconds, like heart attacks or sudden falls by elderly people, or to monitor individuals for early detection of chronic illnesses and cognitive disorders (e.g., cardiovascular, Alzheimer disease, Parkinson disease, diabetes, epilepsy, asthma). For example, high blood pressure is an important indicator for cardiovascular diseases. WBANs are also used in kinematic for rehabilitation assessment and to collect environmental parameters (temperature, humidity, light, exposure to radiation, etc.) of the monitored patient.

At present, many existing medical wireless devices are available in the market (i.e., MICAz, MICA2, Tmote Sky, TelosB, IRIS, Imote2, Shimmer, etc.) and can be used to collect various vital signs [5], such as heart rate (HR), pulse, oxygen saturation (SpO_2), respiration rate, body temperature (T°), electrocardiogram (ECG), electromyogram, blood pressure, blood glucose levels, galvanic skin response, etc. These small devices will improve the life quality of patients by allowing in-home and remote monitoring for the elderly, the immobile, and people with long-term diseases. Wearable and invasive medical sensors provide mobility and freedom by allowing monitored persons to continue their daily life activities while being monitored. They also reduce the healthcare costs (overcapacity, waiting, sojourn time, number of nurses, etc.) through reducing the number of occupied beds in hospital by patients under monitoring.

Patients in hospitals or elderly at home are under-monitored (with about 3 checkups per day) [6] and 6,000 a year die due to poor patient monitoring [7]. The use of WSNs will reduce the healthcare costs by triggering an alarm for caregivers when the health of remotely monitored patients enters a critical phase. For example, when the blood pressure of a diabetes patient is above 130/80, he must be treated immediately because of the high risk of heart attack. Similarly, low SpO_2 is a symptom of heart and lung problems.

However, with the small size and weight of these devices, their underlying constrained resources (such as limited energy, small memory, reduced processing power, limited transmission range, etc.) make them susceptible to various sources of environmental noise—e.g., communication interference, transmission malfunctions, signal fading, short hardware fault, errors, malicious attacks through data injection/modification, replaying attacks [1], or simply the energy outage of the used sensor. These sources of environmental noise may lead to unreliable measurements [8], faulty diagnosis results, false alarms, and an unreliable monitoring [9, 10]. Medical applications have strict requirements for reliability, security, and privacy [2]. The sensor measurements should be accurate to avoid false alarms and misdetections. Anomalous data (also called outliers) from badly attached or compromised sensors must be identified and isolated to ensure

reliable operations. A medical WSN will be rejected by healthcare personnel and patients if results are not reliable [3].

Consequently, faulty measurements heavily affect the monitoring and medical diagnosis results. The false alarms may threaten the life of monitored patient and affect the credibility of such monitoring application, where reliability is extremely important to ensure accuracy in the medical domain [11]. For example, an improperly attached pulse oximeter clip or an external fluorescent light may produce erroneous readings.

The sensing components are the first source of unreliability in medical WSNs, not networking issues [3]. Therefore, abrupt deviations in collected data must be detected and processed in real-time to distinguish between a clinical emergency and faulty measurements, in order to reduce false alarms. Both cases induce anomalous measurements and should be accurately detected. Therefore, an anomaly detection mechanism is required to identify abnormal patterns and to distinguish between sick patients and faulty measurements, thus reducing false alarms and unnecessary intervention by healthcare professionals. The physiological parameters are heavily correlated, where changes occur in at least two or more parameters, and the spatial correlation between monitored attributes can be exploited to distinguish faulty measurements from patient health degradation state.

Anomaly detection algorithms in sensor measurement can be classified into two approaches: parametric and non-parametric. Parametric methods assume a known underlying distribution of collected measurements. The parameters of the distribution function are calculated in a training phase and are used in a test phase to determine if the observation has been emitted by the associated distribution function. However, this assumption is unrealistic in medical applications for monitoring the variations of physiological attributes. Many physiological parameters are highly dynamic and do not have a matching statistical distribution, e.g., the HR can vary from person to person, and even for the same individual, the HR changes with physical activities.

The non-parametric approach does not require any prior knowledge (or assumptions) on the data distribution and uses the distance between test instances (or observations) and the established model to detect deviations in data patterns through the use of thresholds. The most widely used approaches are the kernel density estimator (KDE) and histograms. KDE uses kernel functions to estimate the probability density function (PDF). The test instance with low probability with respect to established PDFs is considered as abnormal. The histogram method is based on the frequency of occurrence of data and determines which category the test instance belongs to. The accuracy of these methods heavily depends on the used threshold.

These approaches assume the existence of training data

without anomalies. In reality, training data is a challenging problem. Usually, unlabeled samples are used to build an initial model, and data from a sliding window are used to update the model. The training data is not free from anomalies, which may induce masking (hiding second outliers) and swamping (where normal values are considered as anomalies).

In this paper, we aim to accurately identify abnormal measurements in the data gathered by medical WSNs. We consider a scenario where many sensors are attached to the patient, in order to monitor some physiological parameters, and transmit the data to a smart phone which must analyze the collected data, and raise alarms to the caregiver only when the patient health degrades. We seek to detect and to remove outliers in order to reduce false alarms triggered by inconsistent sensor readings which significantly deviate from the normal data measurements. The objective is to raise alarms only when the patient's health is abnormal (illness).

The proposed anomaly detection framework is based on the Mahalanobis distance (MD) [12] and the KDE [13]. The MD takes advantage of the correlation between monitored attributes to detect deviations. Only when the MD is greater than a pre-defined threshold, the KDE is activated to detect temporal outliers and to pinpoint responsible attributes. We have applied our anomaly detection approach on real physiological datasets with anomalies. Our experimental results show the effectiveness of our proposed approach for accurate detection with low false alarm rate.

The objective of our proposed framework is to provide reliability in medical WSNs and to distinguish between faulty measurements and critical health degradation. We seek to reduce the false alarm rate triggered by inconsistent sensor readings. Data processing is realized on the base station (smart phone), which has a global view for spatio-temporal analysis.

The rest of this paper is organized as follows. In Section II, we review the related work. In Section III, a brief overview of the techniques used in our approach is presented. Section IV describes the proposed approach for the intrusion detection system. In Section V, experimental results are presented to demonstrate the effectiveness of the proposed approach. Finally, Section VI concludes the paper.

II. RELATED WORKS

Several medical applications for WSNs have been proposed for health monitoring. An accelerometer-based method was used to detect patient inactivity at home and to trigger an alarm for an immobile patient for a long time [14]. Another approach [15] deals with a wearable accelerometer to detect falls by elderly people under remote monitoring.

Many other architectures for medical sensor networks have been proposed and deployed to monitor patients and to raise alarms in case of medical emergencies, such as MEDiSN [4] and CodeBlue [16, 17] for monitoring HR, ECG, SpO₂, and pulse, LifeGuard [18] for ECG, respiration, pulse oximeter, and blood pressure, AlarmNet [19] and Medical MoteCare [20] for physiological (pulse and SpO₂) and environmental parameters (temperature and light), Vital and Jacket [21] for ECG and HR. Surveys of medical applications using WSNs are available [22, 23].

However, the data collected by WSNs have low quality and poor reliability [10] due to their limited resources. Data filtering techniques are used to reduce noise level and retain good data. However, it may change and reduce the shape of variation rather than only cleaning datasets by removing outliers. Therefore, anomaly-based intrusion detection systems are used to build a normal data model and detect unusual deviations. Different approaches for anomaly detection have been proposed and applied in WSNs to detect abnormal deviations in collected data and have been analyzed in terms of their detection accuracy and false alarm ratio [24, 25]. Machine learning [26] algorithms for classification and data mining [26] techniques for clustering have been widely applied, such as neural networks, Naive Bayes, decision trees (C4.5) [27], support vector machines (SVMs) [28-31], self-organizing maps [32], k-means [33], k-nearest neighbor [34], expectation maximization, hierarchical clustering, fuzzy C-means, and Gaussian mixture model [35], etc. Various classification and clustering techniques have been comprehensively studied [26].

Logistic regression modeling with a static threshold was used to evaluate the reliability of a WSN in the industrial field with a large number of sensors [36]. However, they do not update the parameters of the established model to be able to identify the cause of potential loss of reliability. On the same scale of large sensor networks, a diagnosis method based on the enhanced C4.5 (J48 or decision tree algorithm) was proposed, which merges the local classifiers into a large spanning tree to answer for the whole network accuracy [27]. The physical activity of a person was monitored using Sun SPOT sensors attached to the thighs [37]. They used the Naive Bayes algorithm to determine if the person is sitting, standing, lying down, or walking. However, they did not take into consideration that the values can be corrupted due to faulty hardware. Similarly, another system was able to distinguish between mental stress states and relaxation states using logistic regression based on the HR variability [38].

The SVM classifier has gained popularity due to its optimum solution and its simple numerical comparison for data classification. Several SVM-based approaches have been proposed [29, 30, 39] for anomaly detection in WSNs. Moreover, many nonlinear versions of SVMs (kernel-based) have been investigated to find a boundary (or hyperplane) that encompasses the majority of normal

data in the training phase. When the decision boundary is established, any new data outside the boundary is classified as abnormal.

However, machine learning algorithms need a pre-classified (or labeled) training data set, which is often skewed or unavailable in the real world. Skewed (unbalanced) labeled data occurs when one class is over-represented (e.g., 99% of data are normal), and anomalies are almost not available in the training data set. Constructing a labeled training set is often a laborious and expensive task. To resolve these problems of training data in machine learning methods, data mining (or unsupervised) techniques group similar data in one cluster, and flag the small-size clusters (containing less than $t\%$ of total values) as abnormal. However, these techniques assume that anomalous data can be clearly distinguished from normal data, and they are rare when compared to the size of a normal data cluster. They also require the prior knowledge of the number of clusters.

A survey of different techniques for outlier detection in WSNs was proposed, with a comparative guideline to select a suitable technique based on the characteristics of the used data set [25]. Linear regression was used to predict missing data with low error in WSNs [40]. Different approaches were used for anomaly detection in WSNs, such as fixed and dynamic thresholds, linear least squares estimation, auto regressive integrated moving average, hidden Markov model, etc. [41].

A distance-based method was used to identify insider malicious sensors, while assuming neighbor nodes monitor the same attributes [42]. Each sensor monitors its one-hop neighbors and uses the MD between measured and received multivariate instances to detect an anomaly. However, it is impractical in medical applications to exploit promiscuous modes and to place redundant sensors for monitoring the same parameters. MD has been used to classify electronic products as healthy or unhealthy [43].

A score-based approach was used for anomaly detection in collected data by sensors [44]. The proposed approach was based on a Hampel filter and KDE to identify outliers, but it did not take into account the correlation between attributes. Only limited research has used spatial and temporal correlation for outlier detection [10]. The temporal dependency means that the current attribute measurement depends on readings at the previous time instants, while the spatial dependency means that the observations from different attributes are correlated.

In health monitoring, the physiological parameters are heavily correlated. To increase the accuracy of anomaly detection systems, our proposed approach exploits the spatial and temporal dependencies among the monitored physiological parameters to distinguish between faulty measurements and medical emergencies. The objective is to ensure reliable operations of sensors and accurate medical diagnosis results. Sensor measurements tend to be correlated in time and space, and errors are usually uncor-

related with other attributes.

The first attempt to capture spatio-temporal correlations was introduced [10] using regression to build two models using previous observations. However, as the model keeps a sliding window of the past collected instances, the model is subject to false alarms if outliers are not discarded from the training data.

Our proposed framework measures the spatial dissimilarity between multivariate vectors (p -dimensional), through the use of MD to detect abnormal instances. When an abnormal instance is detected, the KDE is activated to detect the change point (temporal deviation). As the physiological parameters are heavily correlated, clinical emergency induces changes in many attributes (at least k), and faulty/abnormal measurements are uncorrelated with other attributes. Therefore, based on the number of deviated attributes, we can distinguish between faulty measurements and a patient entering in an emergency situation.

In this paper, we propose a simple and lightweight approach for online anomaly detection in collected data by medical wireless sensors. The proposed approach is based on MD for spatial analysis and KDE for temporal analysis. The objective is to reduce false alarms resulting from faulty measurements, thus enhancing the reliability and the accuracy of the monitoring system.

III. BACKGROUND

In this section, we briefly review the MD and the KDE used in our framework.

A. Mahalanobis Distance

The MD is a commonly used method for outlier detection in multivariate data. Let $X = (A_1, A_2, \dots, A_p)$ be a multivariate data, where $A_k = (x_{1k}, x_{2k}, \dots, x_{nk})$ is a set of n observations of the k^{th} attribute, and X_i represents an instance vector $X_i = (x_{i1}, x_{i2}, \dots, x_{ip})$

$$X = \begin{matrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{matrix} \begin{pmatrix} A_1 & A_2 & A_3 & \cdots & A_p \\ x_{11} & x_{12} & x_{13} & \cdots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \cdots & x_{np} \end{pmatrix} \quad (1)$$

The MD measures the distance between attributes while taking into account the correlation between them:

$$MD_i = \sqrt{(X_i - \mu)^T \Sigma^{-1} (X_i - \mu)} \quad (2)$$

where $\mu = (\mu_1, \mu_2, \dots, \mu_p)^T$ is the mean vector ($1 \times p$) and Σ is the covariance matrix ($p \times p$) of these p attributes, calculated as:

$$\mu_k = \frac{1}{n} \sum_{i=1}^n x_{ik} \quad (3)$$

$$\Sigma_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \mu_i)(x_{kj} - \mu_j) \quad (4)$$

A large value of MD_i means a large deviation between attributes. MD_i^2 follows a chi-square distribution $\chi_{p,0.975}^2$ with p degrees of freedom (p is the number of attributes) and the 97.5% quantile is used as a threshold for anomaly detection by MD^2 (0.025 significance level for cutoff value). The alarm decision function is given by:

$$A_i = \begin{cases} 1 & \text{if } MD_i \geq \sqrt{\chi_{p,0.975}^2} \\ 0 & \text{Otherwise} \end{cases} \quad (5)$$

B. Kernel Density Estimator

The KDE is a nonparametric method used to estimate the PDF for statistical analysis. Let $A_k = \{x_{1k}, x_{2k}, \dots, x_{nk}\}$ be i.i.d. random variables having a common PDF $f(x)$. The cumulative distribution function $F(x)$ can be estimated by the empirical cumulative distribution function $\hat{F}(x)$:

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n 1(x_{ik} \leq x) \quad (6)$$

To estimate the density $\hat{f}(x)$, we consider the discrete derivative (for a small h):

$$\hat{f}(x) = \frac{\hat{F}(x+h) - \hat{F}(x-h)}{2h} \quad (7)$$

which can be written as:

$$\begin{aligned} \hat{f}(x) &= \frac{1}{2nh} \sum_{i=1}^n 1(x-h \leq x_{ik} \leq x+h) \\ &= \frac{1}{2nh} \sum_{i=1}^n 1\left(\left|\frac{x-x_{ik}}{h}\right| \leq 1\right) \\ &= \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_{ik}}{h}\right) \end{aligned} \quad (8)$$

where:

$$K(u) = \begin{cases} \frac{1}{2} & |u| \leq 1 \\ 0 & |u| > 1 \end{cases} \quad (9)$$

$K(\cdot)$ is the kernel of uniform density function on $[-1,1]$, and h is the bandwidth. $\hat{f}(x)$ counts the probability that the point x is close to observations (x_{ik}) . A large value (near 1) indicates that many observations are near the point x , and a low value indicates that x is an outlier. The uniform kernel in Eq. (9) is a special case of kernel estimator, and in the rest of this paper, we use the Gaussian kernel given by:

$$K(x-x_{ik}) = \frac{1}{\sqrt{2\pi}} e^{-0.5(x-x_{ik})^2} \quad (10)$$

and the optimal bandwidth:

$$h_{opt} \approx \frac{1.06\sigma}{n^{1/5}} \text{ and } \sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ik} - \mu)^2 \quad (11)$$

where μ and σ are the mean and the variance of the vector X_k . In hypothesis testing, KDE is used to estimate the probability of new observations, and when the p-value (probability value) is less than threshold α ($\alpha \in [0.01 - 0.05]$), the null hypothesis is rejected and the observation is considered as abnormal.

IV. PROPOSED APPROACH

We consider a general medical deployment scenario, where N ($N \leq p$) wireless nodes (S_1, \dots, S_N) with restricted resources are placed on the patient body (as shown in Fig. 1). These sensors are used to collect vital signs and transmit the collected data at regular time interval to a sink device. A portable smart phone is placed on the patient arm to collect data from sensors. The collected data are processed in real time on the smart phone to detect an anomaly and raise alarms for caregivers only when patient health degrades (respiratory failure, cardiac arrest, etc.). Faulty measurements must be detected and isolated in order to reduce false alarms and prevent fault diagnosis.

Sensor measurements are sent periodically every discrete time interval T (e.g., 1 minute) to the smart phone, which has more processing power and storage resources than sensors. The real-time analysis of the gathered data on the smart phone is required for the early detection of clinical deterioration, and to alert healthcare profession-

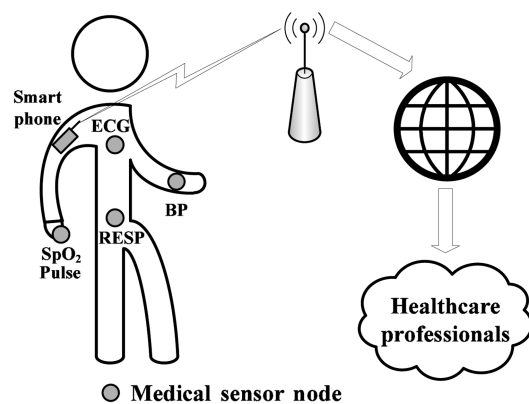


Fig. 1. Remote collection of vital signs in real-time. SpO₂: oxygenation ratio, ECG: electrocardiogram, BP: blood pressure, RESP: respiration rate.

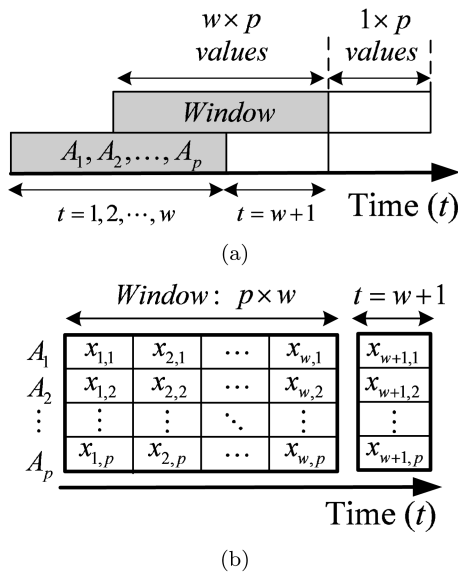


Fig. 2. Sliding window: (a) sliding window used to estimate μ and Σ and (b) reference window and testing instance.

als upon detection of a clinical emergency. To detect abnormal patterns with unsupervised models, a sliding window of the last observations (as shown in Fig. 2a and as detailed in Fig. 2b) is used as training data for estimating the mean μ and covariance matrix Σ . The size of the window has a tradeoff between accuracy and complexity of processing and storage.

After the arrival of a new instance, MD is calculated between the training data in the sliding window and the current attributes values. If MD is greater than $\chi^2_{p,0.975}$ with p degrees of freedom, the univariate KDE is used to pinpoint the abnormal attribute(s), and the window slides one slot by removing the oldest first instance and adding the new one. The architecture of the proposed approach is shown in Fig. 3.

However, the data in the sliding window is not reliable and may contain outliers, which disrupt the estimated values for these statistical parameters. When outliers are in the training set, they dominate and pull the statistical parameters toward them, and this inappropriately leads to a large value of MD for normal data (swamping), or a small value of MD for outliers (masking or misdetection). To provide accurate results of MD, the used data in the sliding window must be cleaned to guarantee anomaly-free training data.

Many robust estimation methods for mean and covariance matrixes of multivariate data have been proposed and used to remove outliers, e.g., minimum volume ellipsoid, orthogonalized Gnanadesikan-Kettenring [45], minimum covariance matrix (MCD) [46], fast-MCD [47], and deterministic MCD [47]. These methods seek to find a subset h out of w instances not contaminated by outli-

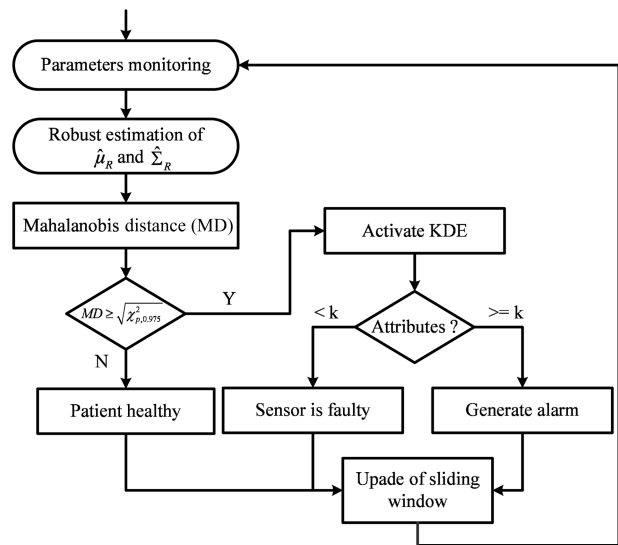


Fig. 3. Flow diagram of the implementation. KDE: kernel density estimator.

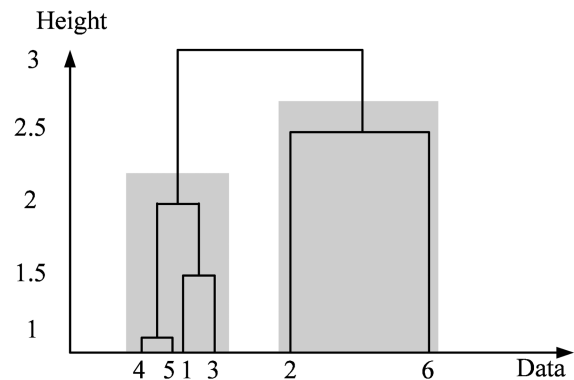


Fig. 4. Dendrogram formed from 6 instances and 2 clusters associated with the cutting level.

ers. In this paper, we look for a simpler method with low computational complexity and storage requirements to derive a subset h of instances without outliers ($h \leq w$).

To achieve this objective, we use the hierarchical agglomerative clustering to aggregate data points with low distance (or resemblance coefficient) in one cluster. The resemblance matrix containing the distance between each point and whole others is used to identify the minimal coefficient and merge the two data points into one cluster. This procedure is repeated to build the dendrogram (or cluster tree) shown in Fig. 4. Clustering is obtained by cutting the dendrogram at a desired level. In this paper, when the distance between clusters (intercluster distance) becomes large enough (at least 3 times the previous distance), we stop the aggregation procedure. The stop point will determine the number of clusters, and

the cluster containing the majority of data points is used to robustly estimate the statistical parameters ($\hat{\mu}_R$ and $\hat{\Sigma}_R$).

After the robust estimation of mean ($\hat{\mu}_R$) and covariance ($\hat{\Sigma}_R$), MD is used to detect deviations using the threshold in Eq. (5). If MD is larger than a threshold ($MD(\hat{\mu}_R, \hat{\Sigma}_R) \geq \sqrt{\chi_{p,0.975}^2}$), an alarm is triggered without any indication about underlying attribute(s). Therefore, we apply the univariate KDE (iff $A_i = 1$) on each attribute to pinpoint suspicious attributes before raising any medical alarm to alert caregivers or emergency teams.

However, KDE is also sensitive to outliers through the bandwidth, which is directly proportional to standard deviation ($\hat{\sigma}$ in Eq. (11)). To overcome this problem and to provide training data without anomalies and without additional computational complexity, we use the subset of values in a sliding window with weight equal to one in the reweighted estimator as a reference:

$$w_i = \begin{cases} 1 & (X_i - \hat{\mu}_R)^T \hat{\Sigma}_R^{-1} (X_i - \hat{\mu}_R) \leq \chi_{p,0.9}^2 \\ 0 & \text{Otherwise} \end{cases} \quad (12)$$

For a new observation x_{new} in each attribute, KDE is used to calculate the probability of the new observation PDF ($\hat{f}(x_{new})$). The p-value (probability value) test is used to detect outliers. If the probability is smaller than the pre-defined significance level α , the observation is considered as abnormal. Strong outliers have a significance level between 0 and 0.01, and weak outliers between 0.01 and 0.05.

When only one attribute is anomalous, the measurement is considered faulty, and no alarm will be raised. However, if at least k attributes are abnormal, we trigger an alarm for healthcare professionals to react; e.g., heavy changes in the HR and reduced rate of SpO₂ are symptoms of patient health degradation and require immediate medical intervention. We assume that the probability of many attributes ($k = 2$ in our experiments) being faulty is very low.

V. EXPERIMENTAL RESULTS

In this section, we present the application results of the proposed framework for online anomaly detection in gathered data by medical WSNs. We use a real medical dataset from the PhysioNet database (MIMIC Database) [48]. The dataset contains 7 attributes: mean values of blood pressure (BPmean), systolic blood pressure, diastolic blood pressure, HR, pulse, respiration rate (RESP), and oxygenation ratio (SpO₂). We only focus on five attributes ($p = 5$): BPmean, HR, pulse, RESP, and SpO₂. We assume no prior knowledge about existing anomalies or faulty measurements in this dataset. We use a sliding window of 24 ($w = 24$) and $k = 2$ attributes.

The variations of BPmean, HR, pulse, RESP, and SpO₂ are presented in Figs. 5–9, respectively. BP is measured

in millimeters of mercury (mmHg) with normal values ($\in [90 - 140]$). HR and pulse are in beats per minute (bpm) with normal values for a healthy adult in rest $\in [60 - 100]$. The RESP is measured in respiration per minute (rpm) and SpO₂ is the percentage of oxygen in the blood with respect to normal values $\in [95\% - 100\%]$. As the physiological parameters are usually not the same for all people and they depend on many parameters (sex, age, weight, activity, etc.), the use of a static interval for anomaly detection heavily depends on many additional dynamic parameters (environmental, ages, activities: rest, moving, awake, sleep, etc.), and these parameters are not easy to set dynamically.

Clearly in Fig. 5, there are two abnormal values of BP falling to 30 and 55 bpm, and other variations associated with clinical change of the monitored patient can be visually distinguished. Furthermore, some values in HR and

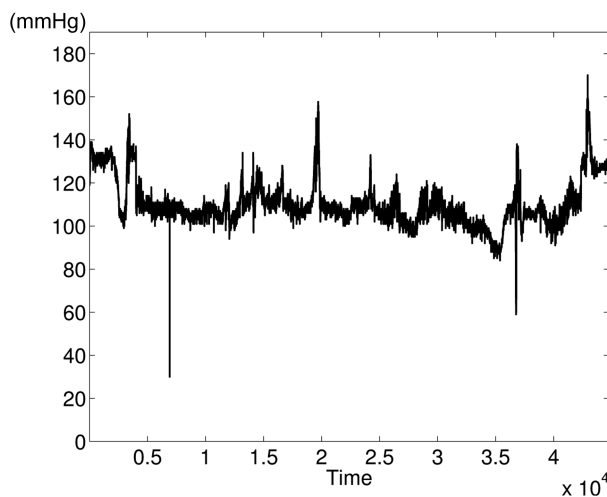


Fig. 5. Blood pressure.

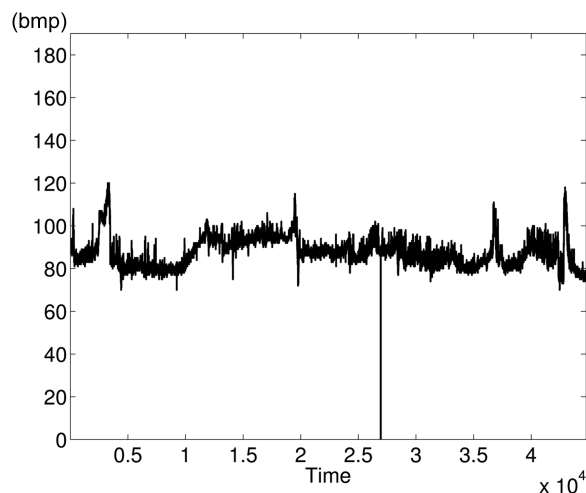


Fig. 6. Heart rate.

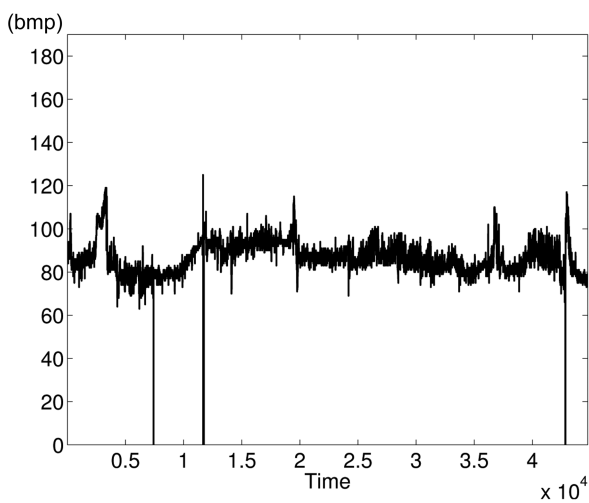


Fig. 7. Pulse.

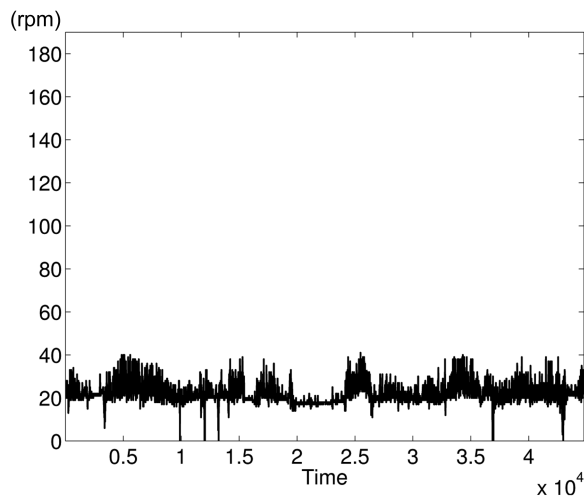


Fig. 8. Respiration rate.

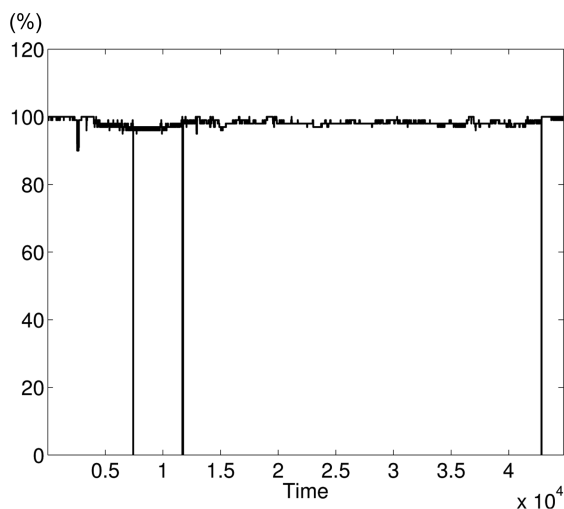


Fig. 9. Oxygenation ratio.

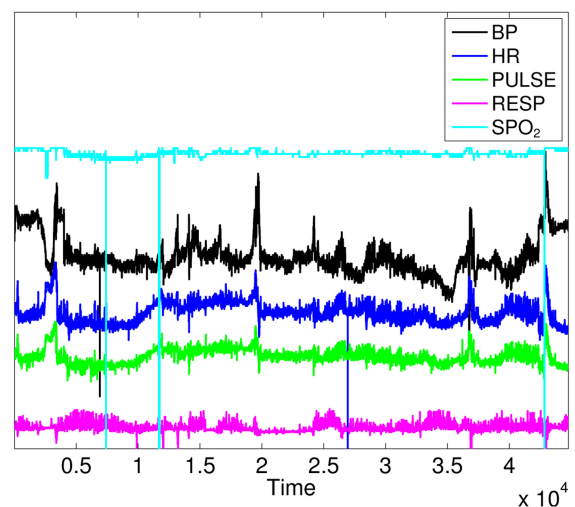


Fig. 10. All parameters. BP: blood pressure, HR: heart rate, RESP: respiration rate, SpO₂: oxygenation ratio.

pulse fall to zero in different time instants (shown in Figs. 6 and 7). HR and pulse measure the same physiological parameter using two different devices, and usually, both curves must superpose. This is not the case when comparing both figures. The same goes for the RESP and the SpO₂ in Figs. 8 and 9. We can visually identify abnormal variations in Fig. 9, where we can see some abnormal readings of SpO₂ with zero values (3 spikes).

To prove the correlation between monitored attributes, we show the variation curves of the 5 parameters in Fig. 10, 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, where one attribute heavily changes independently from others. It is important to note that some variation curves in Fig. 10 are shifted for clarifying the shape of their variations. We can visually identify 4 zones of clinical changes, where either the values of many attributes increase at the same time, or some attributes increase and others decrease.

First, we apply MD over the five physiological attributes (with robust estimation of mean and covariance) to show the utility of KDE used in the second phase. The variations of squared MD (without the applications KDE) are presented in Fig. 11, with a threshold $\chi^2_{5, 0.975} = 12.83$ (horizontal line). Most raised alarms by squared MD in Fig. 12 are false alarms and result from benign deviations or faulty measurements, as shown in Fig. 13, which contains the raised alarms by robust MD and the variations of the 5 attributes.

The raised alarms by the sequential execution of both methods (MD followed by KDE) and the inspections of k deviated attributes are shown in Fig. 14, where the raised alarms are triggered by simultaneous variations in at least

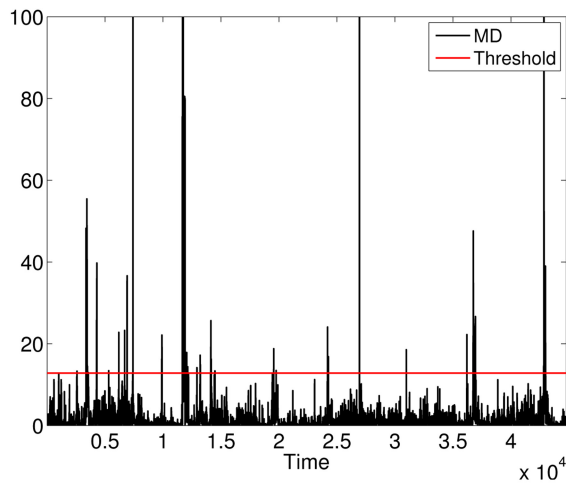


Fig. 11. Squared Mahalanobis distance (MD) & threshold.

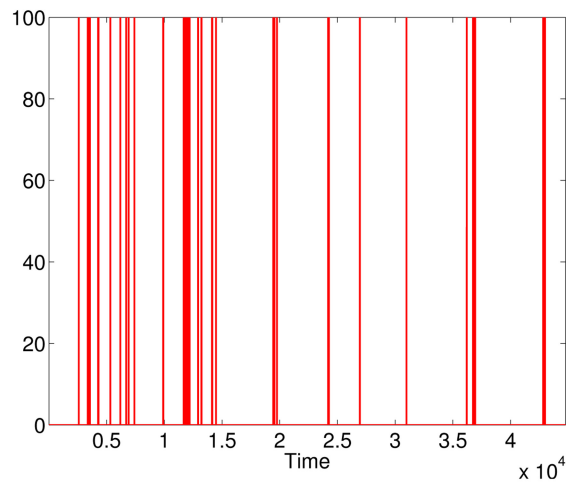


Fig. 12. Raised alarms by Mahalanobis distance.

k attributes. A visual inspection in the variation of monitored attributes in Fig. 14 confirms the accuracy and the utility of raised alarms, where alarms resulted from simultaneous changes in at least 2 attributes.

Furthermore, the false alarms triggered by inconsistent measurements are discarded when comparing Figs. 12 and 15. It is important to note the difference between the number of raised alarms by MD (shown in Fig. 11) and the number of alarms transmitted to the healthcare emergency team after the application of KDE and p-value (shown in Fig. 15).

To evaluate the performance of our proposed approach, we inject abnormal values at different time instants in the different attributes. We use the receiver operating characteristic (ROC) curve to analyze the impact of detection threshold (h) on the detection accuracy and the false alarm ratio. The ROC curve presented in Fig. 16 shows

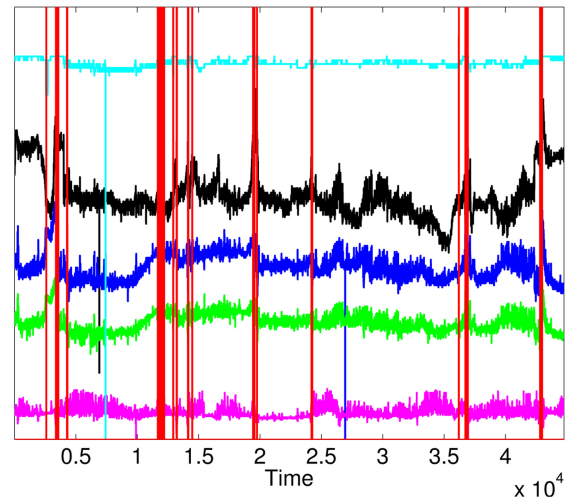


Fig. 14. Medical alarms.

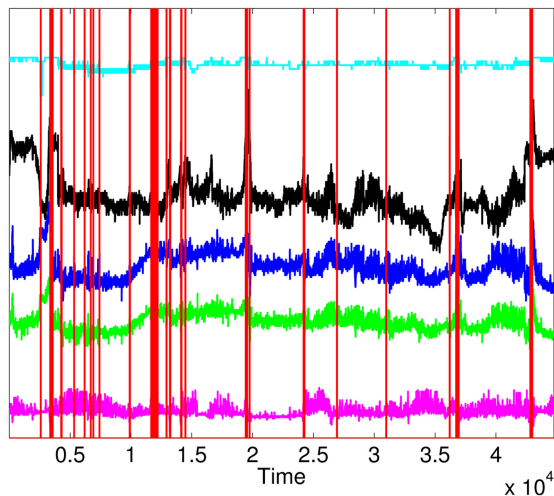


Fig. 13. Mahalanobis distance alarms & 5 attributes.

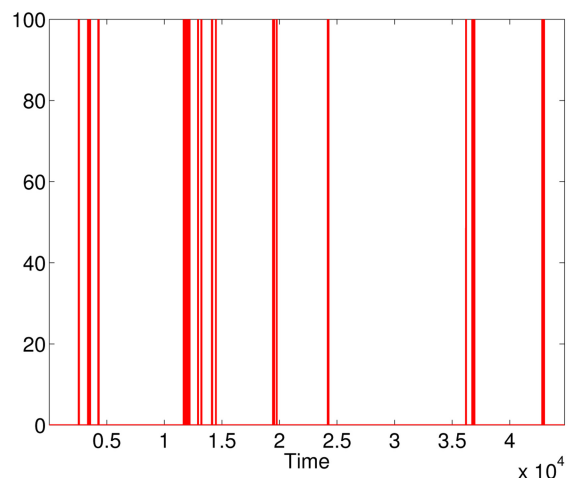


Fig. 15. Raised alarms

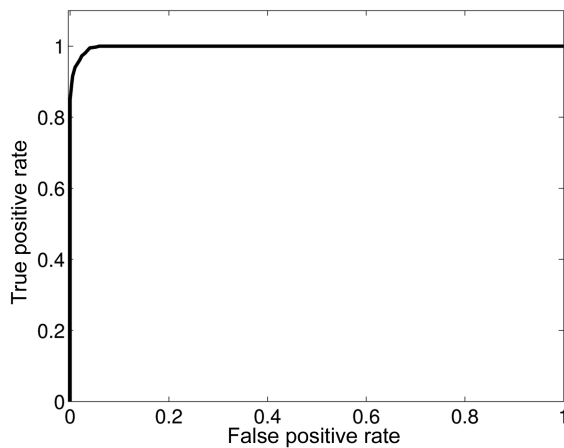


Fig. 16. The receiver operating characteristic.

the relationship between the detection rate (DR; Eq. (13)) and the false alarm rate (FAR; Eq. (14)).

$$\text{Detection rate} = \frac{TP}{TP+FN} \quad (13)$$

where TP is the number of true positives and FP the number of false positives. The false positive rate is defined as:

$$\text{False positive rate} = \frac{FP}{FP+TN} \quad (14)$$

As existing anomalies are not enough to realize this analysis, we synthetically injected 100 anomalies at known time instants in the used dataset. A good detection mechanism should achieve a high detection ratio with the lowest false alarm rate. Fig. 16 shows that our proposed framework can achieve a DR of 100% with an FAR of 5.5%.

VI. CONCLUSION

In this paper, we proposed an unsupervised approach for anomaly detection in medical WSNs, where faulty measurements and injected data could threaten the life of the monitored patient. The proposed approach is based on the MD and a KDE to detect abnormal measurements and to distinguish faulty measurement from a clinical emergency, through the use of spatial and temporal correlation between monitored attributes. The system keeps its relevancy over time by updating the statistical parameters and obtaining more precise evaluation of the normal state of the patient. The proposed approach is suitable for online detection and isolation of faulty or injected measurements with low computational complexity and storage requirement.

We have evaluated the proposed approach using real and synthetic medical datasets. Our experimental results

show the effectiveness of our proposed approach in reducing the number of false alarms triggered by faulty measurements (or maliciously injected data) in medical WSNs.

Most of the time, collected measurements are normal. The reduction of exchanged data between wireless sensors and sink node will be studied in future work. Our next task will be oriented toward distributed detection of an anomaly in sensors to reduce the wasted energy by the transmission of faulty measurements.

ACKNOWLEDGMENTS

This research was supported by the Korea Science and Engineering Foundation, under the World Class University (WCU) program.

REFERENCES

1. P. Kumar and H. J. Lee, "Security issues in healthcare applications using wireless medical sensor networks: a survey," *Sensors*, vol. 12, no. 1, pp. 55-91, 2012.
2. J. Ko, C. Lu, M. B. Srivastava, J. A. Stankovic, A. Terzis, and M. Welsh, "Wireless sensor networks for healthcare," *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1947-1960, 2010.
3. O. Chipara, C. Lu, T. C. Bailey, and G. C. Roman, "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*, Zurich, Switzerland, 2010, pp. 155-168.
4. J. Ko, J. H. Lim, Y. Chen, R. Musvaloiu-E, A. Terzis, G. M. Masson, T. Gao, W. Destler, L. Selavo, and R. P. Dutton, "MEDiSN: medical emergency detection in sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 10, no. 1, article no. 11, 2010.
5. T. Yilmaz, R. Foster, and Y. Hao, "Detecting vital signs with wearable wireless sensors," *Sensors*, vol. 10, no. 12, pp. 10837-10862, 2010.
6. J. P. Marshall, "Continuous quality improvement in colonoscopy," in *Colonoscopy: Principles and Practice*, Malden, MA: Blackwell, pp. 89-101, 2003.
7. S. Adams, "6000 die a year due to poor patient checks," <http://www.telegraph.co.uk/health/healthnews/9430403/6000-die-a-year-due-to-poor-patient-checks.html>.
8. M. Won, S. M. George, and R. Stoleru, "Towards robustness and energy efficiency of cut detection in wireless sensor networks," *Ad Hoc Networks*, vol. 9, no. 3, pp. 249-264, 2011.
9. H. Wang, H. Fang, L. Xing, and M. Chen, "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*, Kyoto, Japan, 2011, pp. 1-5.
10. Y. Zhang, N. A. S. Hamm, N. Meratnia, A. Stein, M. van de Voort, and P. J. M. Havinga, "Statistics-based outlier detec-

- tion for wireless sensor networks,” *International Journal of Geographical Information Science*, vol. 26, no. 8, pp. 1373-1392, 2012.
11. P. K. Sahoo, “Efficient security mechanisms for mHealth applications using wireless body sensor networks,” *Sensors*, vol. 12, no. 9, pp. 12606-12633, 2012.
 12. M. Moshtaghi, C. Leckie, S. Karunasekera, J. C. Bezdek, S. Rajasegarar, and M. Palaniswami, “Incremental elliptical boundary estimation for anomaly detection in wireless sensor networks,” in *Proceedings of the 11th IEEE International Conference on Data Mining*, Vancouver, Canada, 2011, pp. 467-476.
 13. V. S. K. Samparathi and H. K. Verma, “Outlier detection of data in wireless sensor networks using kernel density estimation,” *International Journal of Computer Applications*, vol. 5, no. 7, pp. 28-32, 2010.
 14. T. R. Burchfield and S. Venkatesan, “Accelerometer-based human abnormal movement detection in wireless sensor networks,” in *Proceedings of the 1st ACM SIGMOBILE International Workshop on Systems and Networking Support for Healthcare and Assisted Living Environments*, San Juan, Puerto Rico, 2007, pp. 67-69.
 15. J. Chen, K. Kwong, D. Chang, J. Luk, and R. Bajcsy, “Wearable sensors for reliable fall detection,” in *Proceedings of the 27th Annual Conference of the Engineering in Medicine and Biology*, Shanghai, China, 2005, pp. 3551-3554.
 16. D. Malan, F. J. Thaddeus, M. Welsh, and S. Moulton, “CodeBlue: an ad hoc sensor network infrastructure for emergency medical care,” in *Proceeding on MobiSys 2004 Workshop on Applications of Mobile Embedded Systems*, Boston, MA, 2004, pp. 12-14.
 17. Harvard Sensor Networks Lab, “CodeBlue: wireless sensors for medical care,” <http://fiji.eecs.harvard.edu/CodeBlue>.
 18. K. Montgomery, C. Mundt, G. Thonier, A. Tellier, U. Udoh, V. Barker, R. Ricks, L. Giovangrandi, P. Davies, Y. Cagle, J. Swain, J. Hines, and G. Kovacs, “Lifeguard: a personal physiological monitor for extreme environments,” in *Proceedings of the 26th IEEE Annual International Conference on Engineering in Medicine and Biology Society*, San Francisco, CA, 2004, pp. 2192-2195.
 19. A. Wood, G. Virone, T. Doan, Q. Cao, L. Selavo, Y. Wu, L. Fang, Z. He, S. Lin, and J. Stankovic, “ALARM-NET: wireless sensor networks for assisted-living and residential monitoring,” Department of Computer Science, University of Virginia, Technical Report, 2006.
 20. K. F. Navarro, E. Lawrence, and B. Lim, “Medical MoteCare: a distributed personal healthcare monitoring system,” in *Proceedings of the International Conference on eHealth, Telemedicine, and Social Medicine*, Cancun, Mexico, 2009, pp. 25-30.
 21. J. P. S. Cunha, B. Cunha, A. S. Pereira, W. Xavier, N. Ferreira, and L. Meireles, “Vital-Jacket: a wearable wireless vital signs monitor for patients’ mobility in cardiology and sports,” in *Proceedings of the 4th International Conference on Pervasive Computing Technologies for Healthcare*, Munich, Germany, 2010.
 22. K. Grgic, D. Zagar, and V. Krizanovic, “Medical applications of wireless sensor networks: current status and future directions,” *Medicinski Glasnik*, vol. 9, no. 1, pp. 23-31, 2012.
 23. H. Alemdar and C. Ersoy, “Wireless sensor networks for healthcare: a survey,” *Computer Networks*, vol. 54, no. 15, pp. 2688-2710, 2010.
 24. M. Xie, S. Han, B. Tian, and S. Parvin, “Anomaly detection in wireless sensor networks: a survey,” *Journal of Network and Computer Applications*, vol. 34, no. 4, pp. 1302-1325, 2011.
 25. Y. Zhang, N. Meratnia, and P. Havinga, “Outlier detection techniques for wireless sensor networks: a survey,” *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 159-170, 2010.
 26. C. M. Bishop, *Pattern Recognition and Machine Learning*, New York, NY: Springer, 2006.
 27. X. Cheng, J. Xu, J. Pei, and J. Liu, “Hierarchical distributed data classification in wireless sensor networks,” *Computer Communications*, vol. 33, no. 12, pp. 1404-1413, 2010.
 28. N. Shahid, I. H. Naqvi, and S. B. Qaisar, “Quarter-sphere SVM: attribute and spatio-temporal correlations based outlier & event detection in wireless sensor networks,” in *Proceedings of the IEEE Wireless Communications and Networking Conference*, Shanghai, China, 2012, pp. 2048-2053.
 29. S. Xu, C. Hu, L. Wang, and G. Zhang, “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*, Shanghai, China, 2012, pp. 1-4.
 30. Y. Zhang, N. Meratnia, and P. Havinga, “Adaptive and online one-class support vector machine-based outlier detection techniques for wireless sensor networks,” in *Proceedings of the 23rd IEEE International Conference on Advanced Information Networking and Applications Workshops/Symposia*, Bradford, UK, 2009, pp. 990-995.
 31. Y. Li, Y. Wang, and G. He, “Clustering-based distributed support vector machine in wireless sensor networks,” *Journal of Information & Computational Science*, vol. 9, no. 4, pp. 1083-1096, 2012.
 32. S. Siripanadorn, W. Hattagam, and N. Teaumroong, “Anomaly detection in wireless sensor networks using self-organizing map and wavelets,” *International Journal of Communications*, vol. 4, no. 3, pp. 74-83, 2010.
 33. P. A. Forero, A. Cano, and G. B. Giannakis, “Distributed clustering using wireless sensor networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 707-724, 2011.
 34. K. Vu and R. Zheng, “Geometric algorithms for target localization and tracking under location uncertainties in wireless sensor networks,” in *Proceedings of the IEEE INFOCOM*, Orlando, FL, 2012, pp. 1835-1843.
 35. S. Theodoridis, A. Pikrakis, K. Koutroumbas, and D. Cavouras, *Introduction to Pattern Recognition: A Matlab Approach*, London, UK: Academic Press, 2010.
 36. F. Huang, Z. Jiang, S. Zhang, and S. Gao, “Reliability evaluation of wireless sensor networks using logistic regression,” in *Proceedings of the International Conference on Communications and Mobile Computing*, Shenzhen, China, 2010, pp. 334-338.

37. X. Yang, A. Dinh, and L. Chen, "Implementation of a wearable real-time system for physical activity recognition based on Naive Bayes classifier," in *Proceedings of the International Conference on Bioinformatics and Biomedical Technology*, Chengdu, China, 2010, pp. 101-105.
38. J. Choi, B. Ahmed, and R. Gutierrez-Osuna, "Development and evaluation of an ambulatory stress monitor based on wearable sensors," *IEEE Transaction and Information Technology in Biomedicine*, vol. 16, no. 2, pp. 279-286, 2012.
39. S. Rajasegarar, C. Leckie, J. C. Bezdek, and M. Palaniswami, "Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 518-533, 2010.
40. Y. Xiaozhen, X. Hong, and W. Tong, "A multiple linear regression data predicting method using correlation analysis for wireless sensor networks," in *Cross Strait Quad-Regional Radio Science and Wireless Technology Conference*, Harbin, China, 2011, pp. 960-963.
41. A. B. Sharma, L. Golubchik, and R. Govindan, "Sensor faults: detection methods and prevalence in real-world datasets," *ACM Transactions on Sensor Networks*, vol. 6, no. 3, article no. 23, 2010.
42. F. Liu, X. Cheng, and D. Chen, "Insider attacker detection in wireless sensor networks," in *Proceedings of the 26th IEEE International Conference on Computer Communications*, Anchorage, AK, 2007, pp. 1937-1945.
43. S. Kumar, T. W. S. Chow, and M. G. Pecht, "Approach to fault identification for electronic products using Mahalanobis distance," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 8, pp. 2055-2064, 2010.
44. Y. C. Chen and J. C. Juang, "Outlier-detection-based indoor localization system for wireless sensor networks," *International Journal of Navigation and Observation*, vol. 2012, article no. 961785, 2012.
45. R. A. Maronna and R. H. Zamar, "Robust estimates of location and dispersion for high-dimensional datasets," *Technometrics*, vol. 44, no. 4, pp. 307-317, 2002.
46. P. J. Rousseeuw and K. Van Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212-223, 1999.
47. M. Huberta, P. J. Rousseeuw, and T. Verdonck, "A deterministic algorithm for robust location and scatter," *Journal of Computational and Graphical Statistics*, vol. 21, no. 3, pp. 618-637, 2012.
48. PhysioNet, "PhysioBank ATM," <http://www.physionet.org/cgi-bin/atm/ATM>.



Osman Salem

Osman Salem graduated in computer network engineering in 2001. He received his Ph.D. from the Paul Sabatier University in 2006. He was a researcher in the department of computer science (2006-2007) at Telecom Bretagne. Since September 2008, he is Associate Professor at University of Paris Descartes. His research activities are on anomaly detection in medical wireless sensor networks.



Yaning Liu

Yaning Liu received the Ph.D. degree from Beijing University of Posts and Telecommunications in 2010, and from Telecom Bretagne in France in 2011. She was a post-doctoral researcher at INRIA, France (2011-2012), and has joined JCP-Consult R&D in Rennes in 2013. Her research interests include content centric networking, energy efficient network, peer-to-peer streaming systems and network measurement.



Ahmed Mehaoua

Ahmed Mehaoua is a Full Professor of Computer Communication in the Faculty of Mathematics and Computer Science at the University of Paris Descartes, France. He is currently a visiting professor at the Division of IT convergence Engineering, POSTECH University, Korea, where he is conducting research on Wireless Healthcare Systems and Applications. He is also the Head of the Multimedia Networking and Security Department at the LIPADE, a governmental computer science research center in Paris, France. He is the Vice President of the Scientific Council and the Director of the post-graduate studies on Computer Science of the Faculty of Mathematics and Computer Science, University of Paris Descartes. Prior to joining University of Paris, he was an associate professor at the University of Versailles, France, where he founded and led the Multimedia and Wireless Networking Laboratory. From 1998 to 2011, he was with the University of Cambridge, UK, as a post-doctoral research fellow in the Center for Communications Systems Research (CCSR), a University of Cambridge and Hitachi Telecom Europe joint R&D center. From 1995 to 1998, he was with the Computer Science Research Institute of Montreal, Canada, and was involved in various R&D projects on telecommunication System Management, and Quality of Service Management in broadband networks with Ericsson & Bell Canada. His research interests include security and resource Management in Wireless Sensor Networks for medical and health applications, and Quality of Service (QoS) optimization for audiovisual communications. His research has been supported by government and industry funded projects such as the European Commission, French Ministry of Education and Research (ANR), CNRS, THALES, Thomson, Siemens, Ericsson and Orange Labs., among others. He published more than 120 papers on refereed journals and conferences and he served on various program committees of conferences in computers and communications (Globecom, ICC, HealthCom, VTC, PIMRC, LCN, WCNC, ...).