Markov Models for Anomaly Detection in Wireless Body Area Networks for Secure Health Monitoring

Osman Salem, Khalid Alsubhi, *Member, IEEE*, Ahmed Mehaoua, and Raouf Boutaba, *Fellow, IEEE*

***Abstract*— The use of Wireless Body Area Networks (WBANs) in healthcare for pervasive monitoring enhances the lives of patients and allows them to fulfill their daily life activities while being monitored. Various non-invasive sensors are placed on the skin to monitor several physiological attributes, and the measured data are transmitted wirelessly to a centralized processing unit to detect changes in the health of the monitored patient. However, the transferred data are vulnerable to various sources of interference, sensor faults, measurement faults, injection and alteration by malicious attackers, etc. In this article, we propose a change point detection model based on a Markov chain for centralized anomaly detection in WBANs. The model is derived from the Root Mean Square Error (RMSE) between the forecasted and measured values for whole attributes. The RMSE transforms the monitored attributes into a univariate times series which is divided into overlapping sliding window. The joint probability of the sequence of RMSE values in each sliding window is calculated to decide whether a change has occurred or not. When an effective change is detected over k consecutive windows, the number of deviated attributes is used to distinguish faulty measurements from a health emergency. We apply our proposed approach on real physiological data from the Physionet database and compare it with existing approaches. Our experimental results prove the effectiveness of our proposed approach, as it achieves high detection accuracy with a low false alarm rate (5.2%).**

***Index Terms*— Faulty measurements, forecasting, ARIMA, outlier, Markov chain, healthcare, WBANs.**

I. Introduction

WITH the increase of average lifetimes, the number of elderly people is exponentially increasing in Europe and currently creating an overload in the medical sector, encumbering hospitals with persons under monitoring and increasing the waiting times for surgical operations. To prevent such problems, researchers and doctors are investigating new solutions for remote and pervasive vital signs monitoring through the use of biomedical sensors. Several sensors are placed on the body of the patient to monitor various physiological attributes, and to transmit the measured values to a portable central unit which processes the received physiological data for the prediction or early detection of diseases.

The Internet of Medical Things (IoMT) is the collection of wireless medical devices which can capture, store, analyze and transmit data to healthcare IT systems. Various medical sensors are available in the market that are able to measure different physiological attributes [1] such as Heart Rate (HR), Oxygenation Ratio (SpO2), body temperature (T°), Pulse, Blood Pressure (BP), Respiration Rate (RR), Galvanic Skin Ratio (GSR), Electrocardiogram (ECG), Electroencephalogram (EEG), Electromyogram (EMG), etc. Such wearable vital-signs monitoring sensors have important impacts on public health, overloading in hospitals, and healthcare costs.

The values measured by biosensors are transmitted to a central Local Processing Unit (LPU) using different wireless technologies for real-time analysis and early diagnosis of diseases. The LPU is modernizing the involvement of doctors by providing pre-diagnostic in IoMT for decision-making support. Sensors have already proven their utility in different fields of medicine, such as EEG for the detection of epileptic seizures [2], [3], ECG for the early detection of cardiovascular disease [4], EMG for Human Machine Interface (HMI) [5], etc. The pervasive monitoring and local processing of data in the LPU for epileptic seizure detection allows raising alarms for family or healthcare professionals upon the detection of seizure onset when the patient cannot call for help, especially if they are living an independent life or are isolated and out of the sight of other persons. The heart attack detection system allows reperfusion at earlier stages and can prevent serious heart damage. The HMI helps amputees or disabled persons to control devices and to accomplish some daily life tasks using muscle contractions, body movements, or other physiological attributes.

While WBANs have numerous advantages, their disadvantages range from poor reliability to a high susceptibility to security attacks [6] after deployment. The wireless transmission of data between the sensors and the LPU makes them susceptible to various sources of interference and to attacks such as forgeries and modifications. Furthermore, sensors are prone to both hardware and software issues such as impaired components, sensor calibration, battery exhaustion, or dislocation. The data acquisition process is also subject to faulty measurements, faulty sensors, and improperly attached devices [7]. This leads to unexpected results, false alarms, wrong diagnoses, and unreliable monitoring systems.

Manuscript received November 12, 2019; revised April 12, 2020; accepted April 16, 2020. Date of publication September 29, 2020; date of current version January 15, 2021. This work was supported in part by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under Grant KEP-KEP-6-611-39. *(Corresponding author: Osman Salem.)*

Osman Salem and Ahmed Mehaoua are with LIPADE Laboratory, University of Paris Descartes, 75270 Paris, France (e-mail: osman.salem@parisdescartes.fr; ahmed.mehaoua@parisdescartes.fr).

Khalid Alsubhi is with the Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia (e-mail: kalsubhi@kau.edu.sa).

Raouf Boutaba is with the David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: rboutaba@uwaterloo.ca).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2020.3020602

To enhance the reliability of a monitoring system, automatic detection of an abnormal change is required. This change may be generated by intrusions, faults, or changes in physiological values. To distinguish intrusions or faulty measurements from physiological changes, a spatio-temporal correlation analysis is required, where changes in vital signs are reflected in many attributes and faulty or injected measurements are uncorrelated with other measurements.

In this article, we present a centralized change point detection approach for anomaly detection from data collected by biosensors. The proposed approach is based on a Markov Model (MM) derived from the root mean square of the errors between the forecasted and measured values for whole attributes. It is intended to work with an LPU to detect any abnormal deviation in the collected data and to reject identified faulty and injected measurements. The system raises an alarm for a healthcare professional after the detection of physiological correlated changes and the suppression of faulty or injected measurements.

Our proposed system uses forecasting to reduce the energy consumption from data transmission from sensors to the LPU, and transmission occurs only if the difference between the forecasted and measured values falls outside an acceptance range. Our approach also takes into account the limited processing power of the sensor, where the forecasting model for each physiological attribute is derived and updated on the LPU and transmitted to the associated sensor. It is computationally inexpensive, with low detection delay and high detection accuracy.

The contributions of this article compared to the state of the art is as follows:

- A reduced energy consumption where only suspected measurements (that heavily deviated from forecasted values) are transmitted to the LPU, instead of sending all measured values (normal and abnormal).
- A reduced processing where the parameters of the forecasting model are derived in the LPU as it has more resources than sensors.
- A pre-set number of states in MM to reduce computational complexity.
- A new and lightweight method to derive initial probabilities of MM.
- A lightweight approach to derive the transition probability matrix.
- A reduced false alarm rate through temporal and spatial correlation between monitored attributes.

The rest of this article is organized as follows. In section II, we review recent related work, while Section III presents the building blocks of our proposed change point detection system. In Section IV we present the experimental results from the application of our proposed system on real physiological data as well as the performance analysis results. Finally, section V concludes the paper and presents our future work.

II. Related Work

Several remote healthcare systems have been proposed in the literature, such as MEDiSN [8], CodeBlue [9], MoteCare [10], AlarmNet [11], careNet [12], Mobicare [13], Vital-Jacket [14], WSN4QoL [15], etc. These systems monitor one or several physiological attributes using sensors and aim to provide more pervasive and better healthcare services. However, all these systems focus on architecture and services, without considering faulty and missing measurements. Furthermore, the use of IoT in remote healthcare monitoring requires a robust middleware for effective interactions with the things [16].

Various schemes for anomaly detection have been proposed to detect changes in data collected by wireless sensors, and different Machine Learning (ML) algorithms for data classification have been applied to detect events by distinguishing between normal and abnormal records, such as Naïve Bayes (NB [17], [18]), MultiLayer Perceptron (MLP [18]), Bayesian Network (BN [19], [20]), Decision Tree (J48 [21]), Support Vector Machine (SVM [18], [22], [23]), etc. These algorithms generate a mathematical model from labelled training data and apply it to classify test instances as either normal or abnormal. The SVM is the most used due to its simple numerical comparison for data classification, and as it provides the optimal solution for specific contexts. Given the required computational complexity $O(n^3)$ to derive the classification model in the optimal algorithm (SVM), where n is the number of records in the training data, and the required labelled training data with balanced classes to derive an accurate classification model. These data are usually patient-depending and hard to build or are unavailable in real-life scenarios. Due to the required computation complexity, we will not use supervised ML algorithms in our proposed approach. The computation complexity to derive the decision model in our proposal is lighter than SVM and our approach does not require balanced labelled data.

Unsupervised ML algorithms are used to group similar measurements in a single cluster and to label as abnormal smaller size clusters [24], [25]. Some of the widely used unsupervised algorithms are: K-means, hierarchical clustering, Fuzzy C-means, and Gaussian Mixture Models (GMMs). One challenge facing the use of these clustering methods is that they assume that anomalous data are easily distinguished from normal data, which is unrealistic in physiological data, where normal values for a patient are abnormal for other.

To resolve the problems of supervised and unsupervised learning, Optimum Path Forest (OPF) was proposed and used in several applications, where it provides the probability instead of hard classification, or took confidence interval with class of training data. In its unsupervised version, it does not need to know the number of clusters and considers the sample's neighborhood of each to derive its class. An anomaly detection system based on unsupervised OPF classifier was introduced by Guimarães *et al.* in [26] to distinguish normal data from outliers in WSNs. Only normal samples are provided to identify clusters of normal data, and outliers are not required (and not used) in clusters derivation process, where in real scenario, one may have only normal data without information about anomalies. The comparison results showed that OFP outperforms SVM and Multivariate Gaussian Distribution (MGD) on the used GST and IBRL WSN dataset [26]. However, even if the number of clusters is not required in unsupervised OPF, several parameters must be configured and adjusted, such as the size of the neighborhood to define the clusters and the threshold used to decide if the new sample is abnormal or not.

The Fuzzy Optimum Path Forest (FOPF [27]) was proposed as an extension to OPF using fuzzy concepts. It learns the class of a sample in an unsupervised manner and incorporate the result in supervised training phase, which resolves the problem of unbalanced training data. In fact, samples located close to the cluster center have higher density of neighbors than those located at the border. Therefore, small membership values are assigned to border samples that have no importance for the training phase. However, the computational complexity of fuzzy OPF is $O(n^2)$, which is higher than MM used in our approach, and therefore consumes more energy.

On the other hand, statistical methods build a normal data profile in the training phase and flag deviations from a dynamically updated profile as anomalies. Several approaches, based on KullBack-Leibler distance [28], Mahalanobis Distance (MD [29]), entropy [30], etc., have been proposed and implemented. These techniques are faster and less complex than classification and clustering-based methods; however, distance based methods require access to all measured attributes for anomaly detection and run in a centralized manner on the LPU, which makes them costly in terms of energy consumption by sensors in the transmission of normal data to the LPU.

Several forecasting algorithms have been used [31], [32] in WSNs to extend system lifetimes by transmitting only deviated data. These methods build a prediction model for data measured by the sensors, and transmission only occurs when a measurement deviates from its predicted value. Various forecasting methods have implemented and tested, such as constant prediction, AutoRegression, Least Mean Square (LMS), Holt-Winters, Kalman Filter, Neural Network autoregressive predictor [33], and Auto Regressive Integrated Moving Average (ARIMA [34]).

Aderohunmu *et al.* in [35] compared the performance of 4 forecasting algorithms, i.e., constant prediction, LMS, Weighted Auto Regressive (WAR), and ARIMA in terms of prediction error, complexity (energy consumption) and data transmission reduction, and found that constant prediction slightly outperformed WAR, ARIMA, and LMS, where the data transmission rates were 21.5%, 21.7%, 21.7%, and 25.4% respectively. However, the forecasting accuracy of ARIMA was shown to be higher than the other algorithms, where the root mean square errors for the constant, WAR, ARIMA, and LMS were 0.163, 0.1415, 0.129, and 0.1689 respectively.

Haque *et al.* in [36] use Sequential Minimal Optimization (SMO) regression to predict the current value of the monitored attribute and calculates the Root Mean Square of Error (RMSE) or simply the deviation between predicted and measured values of the monitored attributes. When the number of monitored attributes is larger than a threshold, an alarm is raised. In their performance analysis, they compare their work with SVM, J48, MD. In the same spirit, recently, Smrithy *et al.* in [37] proposes an anomaly detection mechanism using dynamic sliding window in WBANs, where they use Weighted Moving Average (WMA) for prediction. When the deviation between measured and predicted values exceeds a threshold, an alarm is triggered. Both studies in [36] and in [37] use Multiple Intelligent Monitoring in Intensive Care (MIMIC) DB dataset 221 for performance analysis and for comparison with previous work, and we will conduct performance comparisons with these recent related approaches using the same dataset.

However, data collected by WBANs usually have low quality and poor reliability [7], [38], as they are affected by interference, errors, incorrect readings, faulty sensors, environmental noise, missing values, inconsistent readings, and injection and modifications by attackers. Different approaches for anomaly detection have been proposed and applied in WSNs to detect abnormal deviations. Existing solutions in the literature stem from different disciplines, and the reader may refer to the studies in [38]–[41] for a comprehensive review of outlier detection techniques used to distinguish events from errors.

Santos *et al.* in [1] presented a survey of proposals between 2015 and 2019 for the remote monitoring and diagnosis of CardioVascular Diseases (CVD) using Internet of Health Things (IoHT). Their survey was not limited to the detection of myocardial infarction from ECG signal, but also covers a variety of CVD, such as: Arrhythmia, Hypertension, Coronary, Heart Rate Variability (HRV), etc. Their aim is to enumerate the required steps in a reference model for online heart monitoring systems. They identify the key challenges (energy, signals, power, security and privacy) for such monitoring system and show how security issues have been addressed in existing projects.

Zhang *et al.* in [42] proposed a Hidden Markov Model (HMM) to detect faults in ECG data collected using WBAN. They used the Baum-Welch algorithm to derive the parameters of the HMM, which requires computational power that quickly consumes the battery of the processing device in WBAN, especially when the HMM model is updated every time interval T to adapt to normal daily changes. In fact, the HMM is a classifier that uses the measurements to derive the underlying two states of the Hidden Markov Model ($S_0 = Normal$ and $S_1 = Abnormal$) depending on the sequence of ECG measurements. Even if HMM has been often proposed for anomaly detection in measured ECG values, in a real-life deployment scenario, only normal ECG data are available from the user in the training phase, and the derived Markov model is distorted and incorrect. An HMM is more convenient when the data associated with all hidden states are available, e.g., in human activity recognition, body posture identification, etc.

A few anomaly detection approaches based on Markov Model (MM) have been proposed in the literature and applied on sensor data from different applications. Pang *et al.* in [43] proposed an approach based on probability prediction and MM to detect sensor failure, transmission errors and the major faults in spacecraft telemetry. Khan *et al.* in [44] proposed a simplified MM to detect abnormality, intrusion or change in a set of features extracted from the ECG data. In this article, we will use the same procedure in [43], [44] to derive the state space and transition probabilities in MM. We also implemented the method presented in [44] to conduct

TABLE I
SUMMARY OF ANOMALY DETECTIN TECHNIQUES

Type	Method	Technique	Processing	Cons
Supervised Classification	NB [17], [18]	Naïve Bayes	Central	Require balanced labelled training data and have non-linear complexity ($O(n^3)$) to derive the classification model from training data
	BN [19], [20]	Bayesian Network	Central	
	MLP [18]	MultiLayer Perceptron	Central	
	J48 [21]	Decision Tree	Local	
	SVM [18], [22], [23]	Support Vector Machine	Central	
Clustering and density based	k-means [24]	k-means	Central	Failed to distinguish abnormal from normal data in the same cluster
	FC-means [25]	Fuzzy C-means	Central	
	k-medoids [24]	k-medoids	Central	
	GMM [46]	Gauss. Mix. Mod.	Central	
Semi-supervised classifier	OFP [26]	Optimum Path Forest	Central	The size of the neighborhood and the decision threshold need to be configured
	FOPF [27]	Fuzzy OPF	Central	
Distance Based	ED [24]	Euclidean Distance	Central	Energy consumption (costly) by data transmission to the LPU
	MD [29]	Mahalanobis Distance	Central	
	KL [28]	Kullback-Leibler	Central	
Statistical methods	Entropy [30]	Entropy	Central	Normal data distribution
	z-score [24]	Standard score	Local	
Forecasting	KS [31]	Kolmogorov-Smirov	Central	Subject to swamping and masking. Swamping is when outliers are considered normal and masking is when outliers are considered normal
	DBP [24]	Derivative-Based	Central	
	NNA [33]	NN Autoregressive	Central	
	ARIMA [34], [35]	ARIMA	Central	
	WAR [35]	Weighted AR	Central	
	LMS [35]	LMS	Central	
	SMO [36]	Seq. Min. Opt.	Local	
	DSW [36]	Dyn. Slid. Window	Central	
Markov Model	HMM [42]	Hidden MM	Central	The derivation of transition matrix and initial probability are complex
	MM [43], [44]	Markov Model	Central	
	AECG [45]	MM for ECG	Central	

performance comparison with our approach when evaluating both on the same input ECG data set.

Dehabadi *et al.* in [45] enhanced the reliability of anomaly detection system using MM to distinguish hardware failure from patient health degradation. In spirit, their work is like ours, where the objective is to distinguish failures and transient fault from health emergency. They conduct their experiment on physiological data and derive reduced MM with 4 states (normal, physiological anomaly, transient fault, false alarm) to estimate the reliability of the proposed approach. Table I summarizes existing related work highlighting their techniques, disadvantages and adequacy for online processing.

A common problem in the majority of the existing anomaly detection approaches in medical WBANs is the ignorance of both spatial and temporal correlation between the monitored physiological attributes. Given the constrained resources, multiple univariate time series are used to detect a change point, and when deviations are detected in a time series, an alarm is triggered. The existing work focuses on temporal correlation without considering the spatial relationships among attributes. Gao *et al.* in [47] exploited the spatial correlation to detect anomalies using MM, and implemented their ICAD approach in TinyOS and conducted experiments on a testbed with 17 TelosB motes. However, the MM becomes intractable with the various values of collected measurements, and the computational complexity of its parameters increases exponentially with the number of states.

To resolve the problem of computational complexity of deriving the initial probability vector and the transition matrix in MM, a pre-set number of states is required in a Markov Model to keep the computational complexity low, and to make the model dynamic, an automatic updates of its parameters are required to adapt to normal daily changes in the values of the monitored physiological attributes. Furthermore, the spatio-temporal dependencies must be exploited to distinguish between errors from medical emergencies, where measurements tend to be correlated in time and space, and the errors are usually not correlated with other attributes. Our proposed technique does not use any distance or classification technique and does not require labelled training data, and uses instead the Tukey Box outliers detection technique to derive 5 states MM ables to identify abnormal sequence, and exploits

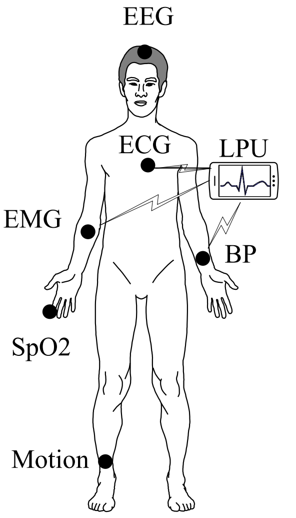

Fig. 1. Deployment scenario.

spatiotemporal correlation to reduce false alarms and achieves high detection accuracy.

III. Proposed Approach

We assume a real deployment scenario where many wireless sensors with restricted resources are placed on the patient's body (as shown in Figure 1) and are used to collect vital signs and transmit the collected data to the LPU. The LPU processes the received data for anomaly detection and may send the data with the associated label (normal or abnormal) to remote medical servers. Then, healthcare professionals or clinicians decide whether to react as for an emergency situation or to assess it as a faulty alarm.

However, data transmission by sensors is more costly with regard to energy consumption than local processing, and the use of a lightweight forecasting procedure limits the sensor transmission to measurements that deviate from forecasted values. In fact, during the training phase, the sensor transmits the collected data to the LPU (as shown in Figure 2), which will use these data to derive the forecasting models, as it has more resources than the sensors. Afterwards, the LPU transmits the derived model to the associated sensor. The sensor and LPU are both able to forecast the current measurement, and only when this measurement deviates from its forecasted value, the sensor will transmit the deviated measurement to the LPU, as illustrated in Figure 2. The LPU may also update the parameters of the forecasting model each specific time interval to further reduce data transmissions from the sensors by integrating the received deviated measurements that do not trigger a medical alarm, as they are not correlated with other measurements.

For clarification, when the difference between the measured and forecasted values is greater than the predefined percentage p of a forecasted value (e.g., $p = 10\%$ of the forecasted value), the measurement is transmitted to the LPU:

$$|x_{tj} - \hat{x}_{tj}| \geq p \times \hat{x}_{tj} \tag{1}$$

where x_{tj} represents the measured value for the j^{th} physiognomical attribute at time instant t, and $\hat{x}_{tj}$ is the forecasted value. It is important to note that even if 10% is an acceptable deviation range for some physiological attributes (HR, PULSE, BP, etc.), this range is considered large for respiration rates, temperature, etc. The value of p depends on the accepted variation range of the monitored attribute.

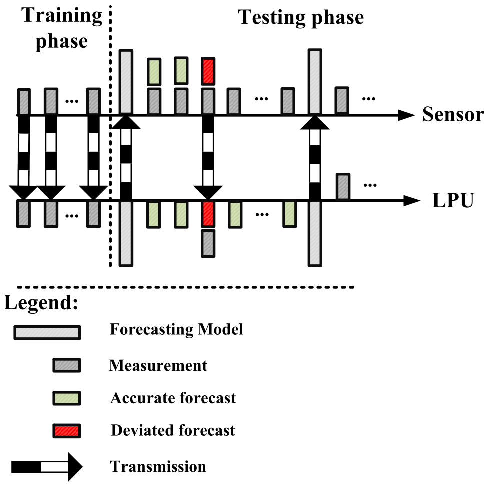

Fig. 2. Reduction of transmitted data.

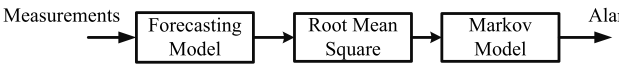

Fig. 3. Flow diagram of the proposed system.

A flow diagram of our proposed architecture is presented in Figure 3, and each block will be detailed in the following subsections.

The collected signal by a wireless sensor is subject to various sources of noises. Therefore, a preprocessing step is required to discard erroneous and replace missing measurement values. The missing measurements have been derived from the median of past W measurements. However, the most important preprocessing step is to remove erroneous measurements in the collected data, which are unavoidable. We discarded measurements above the 95% percentile and replaced them with the median. Afterwards, normalization is performed by subtracting the mean and dividing by the standard deviation of past W measurements. These scaled data are used as input for the forecasting model in the sensor.

A. Forecasting Model

To predict the current value of the physiological attribute, we used the Autoregressive Integrated Moving Average (ARIMA), which is a popular forecasting procedure in WSNs and is chosen due to its simplicity when developing and implementing the algorithm. As well, ARIMA is well-known for providing good forecasting accuracy of time series data with low computational cost. ARIMA(p,d,q) uses a residual time series of order d, with p and q the order of the AutoRegressive (AR) and Moving Average (MA) given in Equation 2:

$$\begin{aligned}\Delta^d x_{t,k} = c + \varphi_1 \Delta^d x_{t-1,k} + \ldots + \varphi_p \Delta^d x_{t-p,k} \\ + e_t + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}\end{aligned} \tag{2}$$

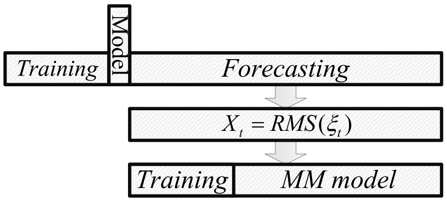

Fig. 4. Training phase for Markov Model.

where φ and θ are the parameters of AR and MA respectively. These parameters will be updated every time period for each physiological parameter in the LPU and then transmitted to the associated sensor.

B. Root Mean Square of Errors

To detect deviations between forecasted and measured values, we used the Root Mean Square of the forecasting Error (RMSE), which is defined as the square root of the mean square error:

$$X_t = RMSE(\xi_t) = \sqrt{\frac{\sum_{j=1}^{K} |x_{t,j} - \hat{x}_{t,j}|^2}{K}} \tag{3}$$

Let X_t represent the time series associated with the root mean square of errors at time instant t ($RMS(\xi_t)$ between the measured $x_{t,j}$ and forecasted values $\hat{x}_{t,j}$. X_t must be near zero for normal measurements, and becomes large when an attribute deviates from the predicted value. The use of X_t reduces the number of analyzed time series from k to a univariate time series, which is more convenient for resource-constrained devices in WBANs.

X_t is calculated on the LPU as it requires the measured values for whole monitored attributes. However, the forecasting model allows the transmission of deviated measurements only and the LPU does not have access to the current values of whole monitored attributes, where normal data are not transmitted by the associated sensor. The LPU has the forecasted values for the whole attributes and will only receive deviated measurements. As the un-transmitted data are within $\pm p\% \times \hat{x}_{t,j}$ of the forecasted value, the LPU replaces each measured value by its upper band $\hat{x}_{t,j} + 0.1 \times \hat{x}_{t,j}$ in order to detect deviations assuming the worst case scenario.

We note that X_t is calculated after the training phase and the derivation of the ARIMA forecasting model. Therefore, splitting the calculated values of X_t into 2 phases (training and testing) to derive the Markov model parameters consists of splitting the X_t times series after the derivation of the forecasting model as shown in Figure 4.

C. Markov Model

The RMSE X_t has a Markov property if it satisfies the memoryless condition:

$$\begin{aligned} P(X_t = s_t | X_{t-1} = s_{t-1}, \ldots, X_0 = s_0) \\ = P(X_t = s_t | X_{t-1} = s_{t-1}) \end{aligned} \tag{4}$$

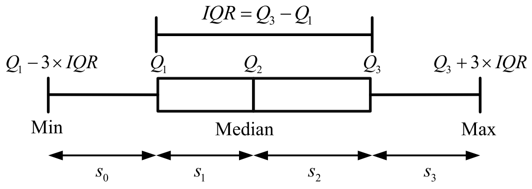

Fig. 5. States of Markov Model.

where the future is independent from the past and depends only on the present. The set of countable states $s = \{s_0, s_1, \ldots, s_n\}$ with the transitions between them is called a Markov chain, which is a model built from random variables X_t that evolve over time in such a manner that the future depend only on the current state, and is independent from past states.

However, as the number of values taken by the random variables X_t is infinite, the number of states will be infinite. To resolve this problem and reduce the computational complexity, we use the robust outlier detection technique Tukey box to fix the number of states in advance, as shown in Figure 5.

The Tukey box (also known as a boxplot or box and whisker diagram) is a way to display the distribution of data based on a five-number summary, i.e., minimum ($Q_1 - 3 \times IQR$), first quartile (Q_1), median(Q_2), third quartile (Q_3) and maximum ($Q_3 + 3 \times IQR$). The InterQuartile Range (IQR) is the distance between the first and third quartiles ($IQR = Q_3 - Q_1$), and is used to determine the minimum and maximum. The Tukey box defines two types of range, i.e., $1.5 \times IQR$ for suspected outliers, and $3 \times IQR$ for Outlier, and we used a range of $3 \times IQR$ to reduce false alarms by including more data.

Four states are defined in the Tukey box (shown in Figure 5) depending on the value of the root mean square of errors X_t:

$$X_t = \begin{cases} s_0 & if\ X_t \in [Min, Q_1] \\ s_1 & if\ X_t \in [Q_1, Q_2] \\ s_2 & if\ X_t \in [Q_2, Q_3] \\ s_3 & if\ X_t \in [Q_3, Max] \end{cases} \tag{5}$$

The use of Tukey box reduces the number of states in MM to 4 and significantly reduces the number of required parameters to 4 initial probabilities and 16 transition probabilities between states. Various types of MMs have been used for learning and for identifying anomalies in data based on Markov properties, where the probability of an event is determined by the previous event only. To consider more than one prior event, a n-order Markov model (n=4 in our approach) takes the previous n events into account making it more suitable for anomaly detection in the collected data. One requirement to derive the parameters of the MM is the availability of training data free from anomalies (without outliers) to prevent the masking problem, where abnormal data is considered normal by the MM.

Therefore, a 4-state Markov model is derived independently from the values of the random variables X_t. To derive the transition probability between two states, we consider two consecutive values of X_t, X_0 and X_1, with value of X_0 in the range of state s_0 and value of X_1 in the range of state s_1. The Bayes theorem can then be used to derive the joint probability of such variables:

$$P(X_0 = s_0, X_1 = s_1) = P(X_0 = s_0).P(X_1 = s_1/X_0 = s_0) \tag{6}$$

With the sum of initial probability equal to 1:

$$\sum_{i=0}^{3} P(X_0 = s_i) = 1 \tag{7}$$

Similarly, for a sequence of $n + 1$ random variables $X_0, X_1, X_2, \ldots, X_n$, the joint probability using the Bayes theorem becomes:

$$\begin{aligned} & P\left(X_0 = s_0, \ldots, X_n = s_i\right) \\ & = P\left(X_0 = s_0\right).P(X_1 = s_1, \ldots, X_n = s_i/X_0 = s_0) \\ & = P\left(X_0 = s_0\right).P(X_1 = s_1/X_0 = s_0)\,.P(X_2 = s_2, \\ & \quad \ldots, X_n = s_i/X_1 = s_1) \\ & = P(X_0 = s_0).P(X_1 = s_1/X_0 = s_0).P(X_2 = s_2/X_1 = s_1). \\ & \quad \ldots P(X_n = s_i/X_{n-1} = s_k) \end{aligned} \tag{8}$$

We denote by Q the vector of the initial probability distribution vector, i.e., $Q = [q_0, q_1, q_2, q_3]$:

$$q_i = P(X_0 = s_i) \tag{9}$$

And denote by $P_{i,j}$ the transition probability from state s_i to state s_j, as given in Equation 10:

$$P_{i,j} = P(X_n = s_j/X_{n-1} = s_i) \tag{10}$$

The joint probability distribution in Equation 8 can be simplified using Equations 9 and 10:

$$\begin{aligned} & P(X_0 = s_0, \ldots, X_n = s_i) \\ & \quad = q_0.P(s_1/s_0).P(s_2/s_1) \ldots P(s_i/s_k) \\ & \quad = q_0.P_{0,1}.P_{2,1} \ldots P_{k,i} = q_0. \prod_{t=1}^{n} P_{t-1,t} \end{aligned} \tag{11}$$

The transition probability matrix (P) contains the transition probability $P_{i,j}$ from state s_i at time t to state s_j at time $t+1$. Given that our model contains 4 states, P is 4×4 matrix:

$$P = \begin{pmatrix} P_{0,0} & P_{0,1} & P_{0,2} & P_{0,3} \\ P_{1,0} & P_{1,1} & P_{1,2} & P_{1,3} \\ P_{2,0} & P_{2,1} & P_{2,2} & P_{2,3} \\ P_{3,0} & P_{3,1} & P_{3,2} & P_{3,3} \end{pmatrix} \tag{12}$$

The sum of probability of whole outgoing transitions from any state is equal to 1, i.e., the sum of each row in matrix P (Equation 12) is equal to 1:

$$\sum_{j=0}^{3} P_{i,j} = 1 \tag{13}$$

The associated transition diagram with P in Equation 12 is shown in Figure 6. To derive the initial probability distribution Q and the transition matrix P, the data in the training phase of the Markov Model are used to derive these parameters. It is important to note that data set used to derive the Markov model was different from the training data set used to derive the forecasting model. In the training phase of the Markov Model, the system may ask the user to move during the first few minutes in order to cover a larger range of data variations. The simplified Markov model can also be updated every specified time interval to cover changes in the monitored attributes.

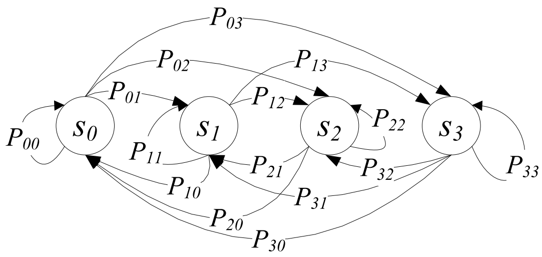

Fig. 6. Transition diagram of MM.

Let N denote the number of record received in the training phase. We start by replacing each X_t (RMSE value) by the associated state with respect to the location of its value in the predefined intervals of the Tukey box. Let N_i be the number of samples in state s_i, e.g., the number of occurrences of state s_i in the training set. The vector of initial probability Q is derived as follows:

$$q_i = \frac{N_i}{N} \tag{14}$$

We use $N_{i,j}$ to denote the number of state s_i followed by state s_j and T_i is the number of outgoing transitions from state i.

$$P_{i,j} = \frac{N_{i,j}}{T_i} \tag{15}$$

where $T_i = N_i$ except for the last state in the training data, which has one transition less than other ($T_i = N_i - 1$) states. The derived Markov model with parameters Q and P, calculated in Equations 14 and 15, is used to represent the temporal profiles of the normal physiological parameters with tolerance to some change given the burden of upper and lower limit of the whisker in the Tukey box.

To detect a change point associated with an intrusion, faulty measurements, or an emergency situation, the sequence of X_t in the testing phase is divided into a sequence of sliding windows (denoted by SW_i) of size N observations (e.g., $N = 5$ in our experiments). The values of X_t in each SW_i are replaced by the associated state, and the joint probability of states is calculated as given in Equation 11. A sequence of states (SW_i) with a high probability indicates similarity with the training data, and a sequence of states with a very low probability is abnormal. A threshold h must be used to distinguish normal from abnormal probabilities. If the joint probability is lower than the predefined threshold, a change point is detected and further investigation through spatio-temporal correlation is required to distinguish faulty/forged data from an emergency situation.

For clarification, we consider the following training window (or TRW in Equation 16) which contains sequence of states' occurrence derived from RMSE values (X_t):

$$\begin{aligned} TRW = (& s_0, s_3, s_2, s_1, s_3, s_3, s_2, s_0, s_1, s_2, s_2, s_0, s_2, s_1, s_3, \\ & s_0, s_1, s_3, s_1, s_1, s_2, s_2, s_0, s_1, s_3, s_3, s_1) \end{aligned} \tag{16}$$

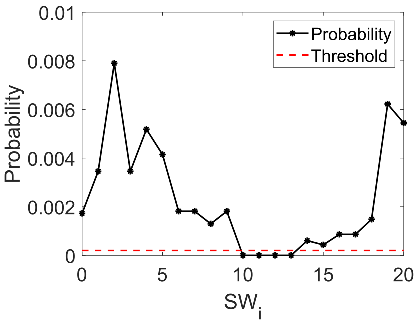

Fig. 7. Detection of abnormal sequence using threshold.

Q and P can be calculated from this training set, as given in Equations 14 and 15:

$$q = [5/27, 8/27, 7/27, 7/27] = [0.185, 0.296, 0.259, 0.259]$$

$$P = \begin{pmatrix} 0 & 3/5 & 1/5 & 1/5 \\ 0 & 1/7 & 2/7 & 4/7 \\ 3/7 & 2/7 & 2/7 & 0 \\ 1/7 & 2/7 & 2/7 & 2/7 \end{pmatrix}$$

Let us consider the following testing sequence of states (denoted by TW or testing window), which we divide into sliding window (SW) of 5 states ($N = 5$) to derive the joint probability and are given in Equation 16 and in Table II respectively.

$$TW = (s_0, s_2, s_1, s_3, s_3, s_1, s_3, s_2, s_0, s_3, s_2, s_2, s_0, s_2, s_3, s_0, s_4, s_3, s_1, s_1, s_2, s_2, s_0, s_1, s_3) \quad (17)$$

The sliding windows SW_{10}, SW_{11}, SW_{12} and SW_{13} have a joint probability equal to 0 and are abnormal whatever the value of threshold h. Figure 7 shows the probability of states' sequence in SW_i and the used threshold $h = 10^{-4}$.

The pseudo code of the MM is given in Algorithm 1, where deviations detected by the MM must be inspected through spatio-temporal correlation analysis. In fact, the physiological parameters are heavily correlated in time and space, and faulty measurements are spatially unrelated with other attributes, so we use this correlation to distinguish false alarms from emergency situations. A change in one physiological attribute induces variations in several other attributes, like an asphyxia which induces low rate of oxygen in the blood (low value of SpO2) and provokes an increase in the respiration rate. The anomaly detection phase in the LPU does not trigger a medical alarm before checking the spatial correlation between attributes. To reduce the number of false alarms, a medical alarm is triggered only when at least r ($r \geq 2$) deviated attributes are detected as shown in the building blocks of our complete system in Figure 8. On the other case, the LPU considers the measurement as faulty and will not raise any alarm.

Algorithm 1 Implementation of Markov Model

```
Collect Markov Model training data
Derive lower and upper bound of Tukey box
Replace $X_t$ by $s_i$
Calculate $N$, $N_i$, $N_{i,j}$ and $T_{i,j}$
Calculate $q_i = N_i/N$ and $P_{i,j} = N_{i,j}/T_i$
Set the size of $SW_i$ and threshold $h$
Replace new $X_t$ by state $s_i$
for each $SW_j \in Testing$ do
    $P(SW_j) = q_i \prod_{t=1}^{w} P_{t-1,t}$
    if $P(SW_j) \leq h$ then
        Raise an alarm for spatio-temporal analysis
    end if
end for
```

TABLE II
PROBABILITIES OF STATES' SEQUENCE

Window	Joint probability
$SW_0 = \{s_0, s_2, s_1, s_3, s_3\}$	$P(SW_0) = q_0.P_{0,2}.P_{2,1}.P_{1,3}.P_{3,3}$ $= 0.001727$
$SW_1 = \{s_2, s_1, s_3, s_3, s_1\}$	$P(SW_1) = q_2.P_{2,1}.P_{1,3}.P_{3,3}.P_{3,1}$ $= 0.00345535$
$SW_2 = \{s_1, s_3, s_3, s_1, s_3\}$	$P(SW_2) = q_1.P_{1,3}.P_{3,3}.P_{3,1}.P_{1,3}$ $= 0.007897944$
$SW_3 = \{s_3, s_3, s_1, s_3, s_2\}$	$P(SW_3) = q_3.P_{3,3}.P_{3,1}.P_{1,3}.P_{3,2}$ $= 0.00345535$
$SW_4 = \{s_3, s_1, s_3, s_2, s_0\}$	$P(SW_4) = q_3.P_{3,1}.P_{1,3}.P_{3,2}.P_{2,0}$ $= 0.005183026$
$SW_5 = \{s_1, s_3, s_2, s_0, s_3\}$	$P(SW_5) = q_1.P_{1,3}.P_{3,2}.P_{2,0}.P_{0,3}$ $= 0.00414642$
$SW_6 = \{s_3, s_2, s_0, s_3, s_2\}$	$P(SW_6) = q_3.P_{3,2}.P_{2,0}.P_{0,3}.P_{3,2}$ $= 0.001814059$
$SW_7 = \{s_2, s_0, s_3, s_2, s_2\}$	$P(SW_7) = q_2.P_{2,0}.P_{0,3}.P_{3,2}.P_{2,2}$ $= 0.001814059$
$SW_8 = \{s_0, s_3, s_2, s_2, s_0\}$	$P(SW_8) = q_0.P_{0,3}.P_{3,2}.P_{2,2}.P_{2,0}$ $=0.001295756$
$SW_9 = \{s_3, s_2, s_2, s_0, s_2\}$	$P(SW_9) = q_3.P_{3,2}.P_{2,2}.P_{2,0}.P_{0,2}$ $=0.001814059$
$SW_{10} = \{s_2, s_2, s_0, s_2, s_3\}$	$P(SW_{10}) = q_2.P_{2,2}.P_{2,0}.P_{0,2}.P_{2,3}$ $=0$
$SW_{11} = \{s_2, s_0, s_2, s_3, s_0\}$	$P(SW_{11}) = q_2.P_{2,0}.P_{0,2}.P_{2,3}.P_{3,0}$ $= 0$
$SW_{12} = \{s_0, s_2, s_3, s_0, s_3\}$	$P(SW_{12}) = q_0.P_{0,2}.P_{2,3}.P_{3,0}.P_{0,3}$ $=0$
$SW_{13} = \{s_2, s_3, s_0, s_3, s_3\}$	$P(SW_{13}) = q_2.P_{2,3}.P_{3,0}.P_{0,3}.P_{3,3}$ $=0$
$SW_{14} = \{s_3, s_0, s_3, s_3, s_1\}$	$P(SW_{14}) =q_3.P_{3,0}.P_{0,3}.P_{3,3}.P_{3,1}$ $=0.0006046863$
$SW_{15} = \{s_0, s_3, s_3, s_1, s_1\}$	$P(SW_{15}) = q_0.P_{0,3}.P_{3,3}.P_{3,1}.P_{1,1}$ $=0.0004319188$
$SW_{16} = \{s_3, s_3, s_1, s_1, s_2\}$	$P(SW_{16}) = q_3.P_{3,3}.P_{3,1}.P_{1,1}.P_{1,2}$ $=0.0008638376$
$SW_{17} = \{s_3, s_1, s_1, s_2, s_2\}$	$P(SW_{17}) = q_3.P_{3,1}.P_{1,1}.P_{1,2}.P_{2,2}$ $=0.0008638376$
$SW_{18} = \{s_1, s_1, s_2, s_2, s_0\}$	$P(SW_{18}) = q_1.P_{1,1}.P_{1,2}.P_{2,2}.P_{2,0}$ $=0.001480864$
$SW_{19} = \{s_1, s_2, s_2, s_0, s_1\}$	$P(SW_{19}) = q_1.P_{1,2}.P_{2,2}.P_{2,0}.P_{0,1}$ $=0.006219631$
$SW_{20} = \{s_2, s_2, s_0, s_1, s_3\}$	$P(SW_{20}) = q_2.P_{2,2}.P_{2,0}.P_{0,1}.P_{1,3}$ $=0.005442177$

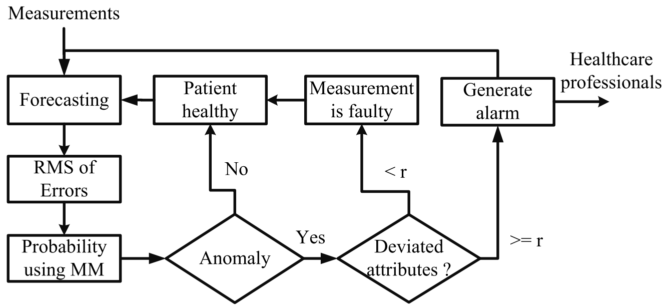

Fig. 8. Flow diagram of the proposed system.

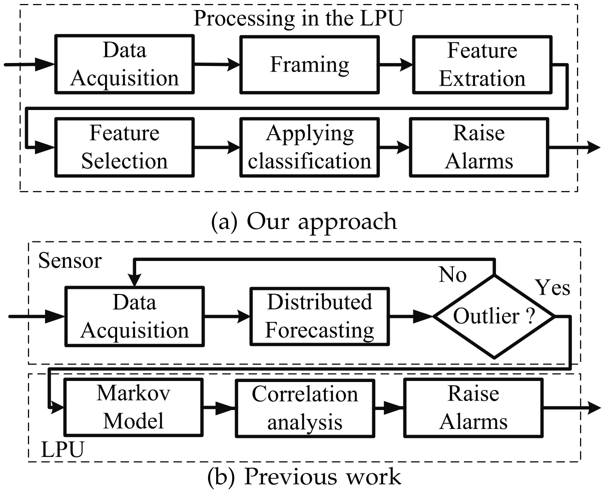

Fig. 9. Block diagrams.

The threshold h is set in our experiment to a small probability value (e.g. smaller than 10^{-4} in our experiments) to exclude the sequence of states containing anomalies (i.e. the probability of occurrence is smaller than 0.0001). The value of h must be small (zero or near to zero) to pinpoint abnormal sequence in measured values. The impact of the threshold on the performance is analyzed through the Receiver Operating Characteristic (ROC) curve in the next section.

The block diagram in Figure 9a illustrates the overall idea underlying our approach, where the processing is realized in the sensor and only in case of anomaly, the MM processing is triggered in the LPU. Our approach achieves spatiotemporal correlation between attributes before raising a medical alarm, in contrast to the traditional frameworks illustrated in Figure 9b, where the processing is centralized in the LPU to derive a classification model from training data. The latter are complex to update and greedy in terms of energy consumption for devices with restricted resources in WBAN.

IV. Experimental Results

In this section, we conduct experiments on real physiological data to analyze the detection accuracy of our proposed system. We use several annotated real physiological data sets from the MIMIC database from the Physionet [48] web site. We will present the variations of the physiological attributes for two patients (221 and 259), as well as the RMSE and the deviations detected by the MM. Afterwards, we will compare the performance of our approach with recent related work: first, with SMO [36] and WMA [37] approaches; second, with linear SVM, K-NN (K=3), J48, MD; and third, with dynamic size MM in [44] through the ROC. The implementation of supervised classifiers (SVM, K-NN and J48) was conducted using WEKA [49] API.

We develop the proposed approach from scratch in Python using PyOD (Python toolkit for Outlier Detection), Numpy and Pandas. The parameters used by our proposed approach in the distributed forecasting are automatically derived from the training data, and when the deviation (RMS) between forecasted and measured values is greater than p ($p = 10\%$ in our experiments), the measurement is transmitted to the LPU. A small value of $p = 1\%$ provokes the transmission of all measurements to the LPU and drains the energy of sensors, whereas a large value reduces the detection accuracy.

Similarly, the initial probability vector and the transition matrix for MM are also derived from training data in the LPU. The anomaly detection threshold h is set to a low value ($h = 10^{-4}$) in our experiments. Increasing the value of h will reduce both the false alarms and the detection accuracy and vice versa.

The first data sets from subject 259, contains 12 attributes: systolic Arterial BP (ABPsys), diastolic Arterial BP (ABPdias), mean Arterial BP (ABPmean), Cardiac Output (C.O.), mean Pulmonary Artery Pressure (PAPmean), Systolic PAP (PAPsys), Diastolic PAP (PAPdias), Heart Rate (HR), PULSE, Respiration (RESP), oxygenation ratio (SpO2) and T°. However, some attributes are missing from the data set (detached sensors) during interesting changes, and we focus only on 5 attributes: HR, PULSE, RESP, SpO2, and T°. The second data set, from subject 221, contains 7 attributes: ABPsys, ABPdias, ABPmean, HR, PULSE, RESP and SpO2. As ABPmean is derived from ABPsys and ABPdias, we focus only on five attributes: BPmean, HR, PULSE, RESP and SpO2. To simulate real measurements scenarios, a Raspberry PI 2 (Model B) is used as a transceiver to send the data to a tablet as a bulk of records containing 5 fields in each discrete time interval T. The measured values of each attribute are compared with the predicted values using ARIMA(7,1,1). This prediction model was chosen because it gives the best smoothed fit for the used physiological data sets. The duration of training phase to derive the parameters of the prediction model was fixed at 250 samples (5 sec at 50Hz). The duration of the training phase of MM in the LPU was set to 1 minute in our experiments, in a similar manner to existing applications for activity recognition in smartphones.

We start by showing the variations of part of the physiological parameters from subject 259. Figures 10 and 11 show the variations of HR and PULSE respectively. Both parameters are in beats per minute (bpm) and represents the same physiological attributes measured using different sensors at different locations. The variations of the respiration rate (respirations per minute (rpm) are shown in Figure 12. The SpO2 is the percentage of oxygen in the blood, and its normal value is within the interval 95% to 100%. A rate lower than 95% is representative of asphyxia, lack of oxygen,

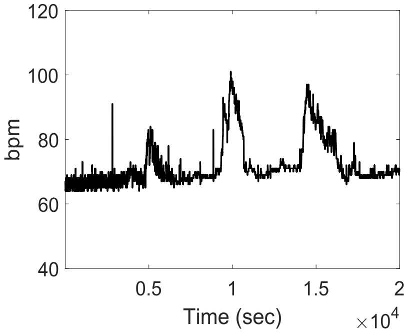

Fig. 10. Heart rate.

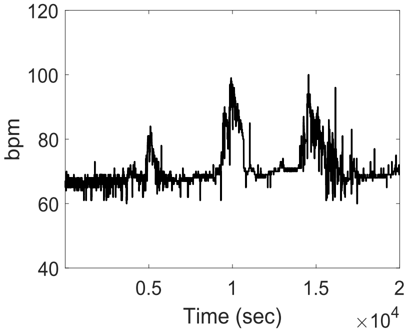

Fig. 11. PULSE.

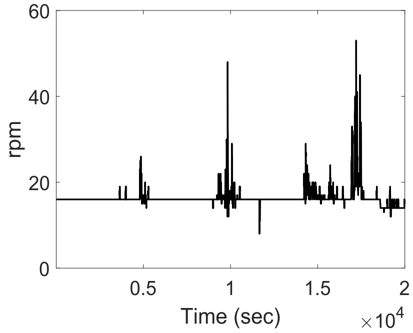

Fig. 12. Respiration rate.

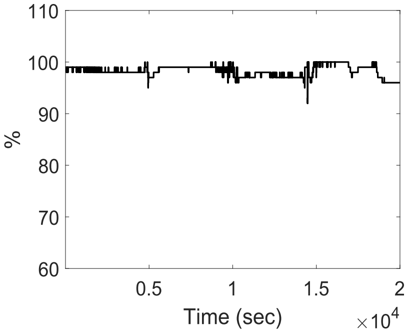

Fig. 13. Oxygenation ratio.

and heart disease. The variations in SpO2 are shown in Figure 13, and the variations in temperature are shown in Figure 14.

To visually identify the correlated deviations for many attributes, Figure 15 shows the variations of the 5 attributes,

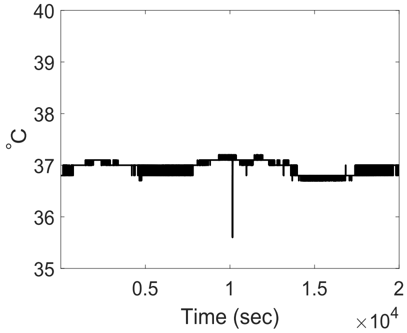

Fig. 14. Temperature.

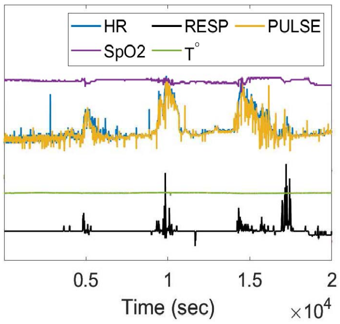

Fig. 15. 5 Attributes.

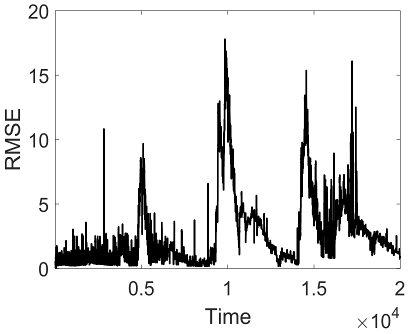

Fig. 16. RMSE.

where we can identify 3 zones of correlated changes around the time instant: $0.5.10^4$, 10^4 and $1.5.10^4$. The root mean square of error between the forecasted and measured values of whole attributes is derived on the LPU, and the variations are shown in Figure 16. The anomalies detected by the Markov Model from the RMSE time series are shown in Figure 17, where the deviations are flagged as anomalies. To discriminate faults from abnormal physiological changes, spatial correlation analysis is done through checking the number of received deviated data, where an alarm is raised only if the number of deviated attributes is greater than r. The alarms raised for healthcare professionals are shown in Figure 18, where the zones of deviation are accurately identified, and many anomalies flagged by the MM are discarded after the correlation analysis.

The variations of ABPmean, HR, PULSE, RESP, and SpO2 of subject 221 are shown in Figure 19, where we can

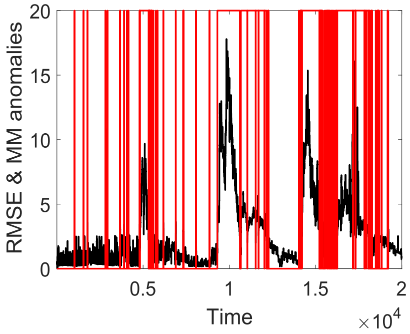

Fig. 17. Anomalies detected by the MM.

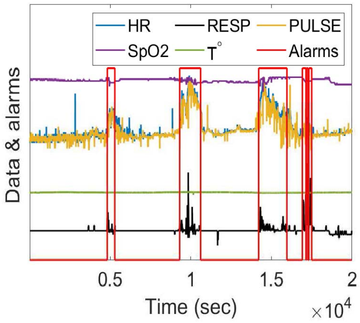

Fig. 18. Data with raised alarms.

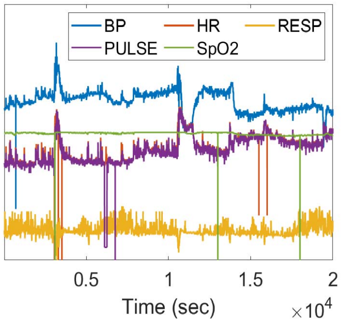

Fig. 19. Physiological data.

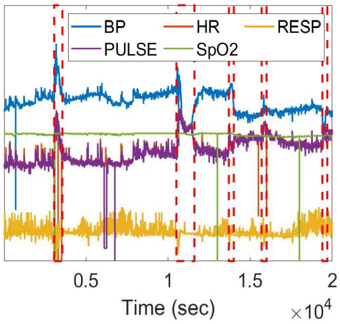

Fig. 20. Zones of change.

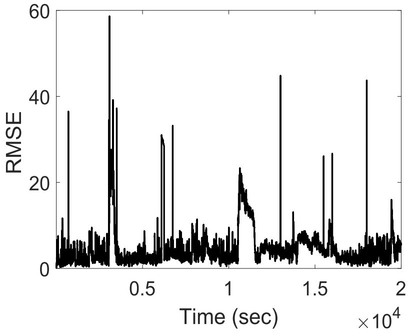

Fig. 21. RMSE.

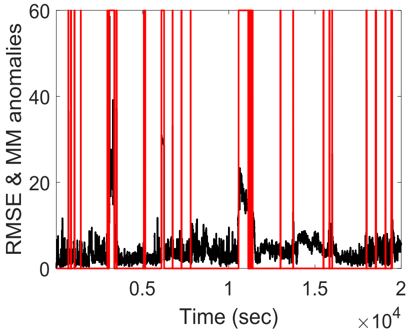

Fig. 22. Detected by MM.

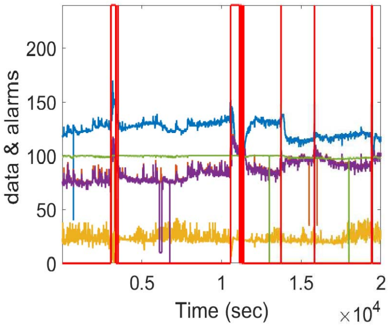

Fig. 23. Raised alarms

visually identify 5 zones of correlated changes as illustrated using the dashed rectangles in Figure 20. The ABPmean is measured in millimeters of mercury (mmHg). The variations in the RMS between the measured and forecasted values are shown in Figure 21, which transforms the problem into change detection in univariate time series.

The abnormal set of states (i.e., low probability) identified by the MM are shown in Figure 22, where the deviations are flagged as abnormal. However, as the physiological parameters are correlated, sensor data anomalies are determined by analyzing the number of deviated attributes (r). A small value of r will increase the true detection rate and at the same time will also increase the false alarm rate, and a large value of r will decrease the detection rate as well as the false alarm rate. Therefore, as true detection and false alarms are in the same direction, a tradeoff is required to set the value of r. In our experiments, we set the value of r to 2, i.e., when only one measurement is received by the LPU (one deviation), it is considered as a fault and no medical alarm is raised.

To conduct a performance analysis of our proposed approach, we use the ROC curve to study the impact of the

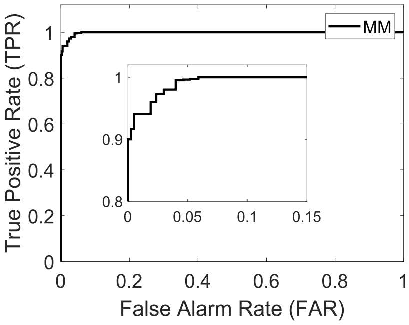

Fig. 24. ROC.

threshold on the True Positive Rate (TPR) and False Alarm Rate (FAR). The TPR and FAR are given in the following equations:

$$\text{TPR} = \frac{TP}{TP + FN} \tag{18}$$

$$\text{FAR} = \frac{FP}{FP + TN} \tag{19}$$

where TP is the number of True Positives, FP is the number of False Positives, FN is the number of False Negatives, and TN is the number of True Negatives. The ROC curve presented in Figure 24 shows the relationship between the TPR and FAR for our proposed approach using MM, which can achieve a TPR of 100% with 5.2% FAR on used dataset.

False positives are unavoidable in anomaly detection and have an impact on the detection delay and on the reliability of the monitoring system. As the FAR increases with detection accuracy (and vice versa), a tradeoff is needed. A large value of FAR makes the monitoring system unreliable and a low value implies a reduced detection accuracy and large detection delay. Our system has a FAR of 5.2%, which is relatively low and acceptable in real medical scenarios, where this rate is for example about 20% in breast cancer diagnosis (Annals of Family Medicine 2015).

Afterwards, we conduct performance comparison of our approach with recent related work for anomaly detection in remote healthcare monitoring systems. As we are using publicly available data set (MIMIC DB dataset 21) which has been used for testing proposed algorithms in several previous works [36], [37], we conduct performance comparison with these recent proposed approaches that used the same data set. The comparison is conducted in terms of TPR and FAR as shown by the ROC presented in Figure 25. The SMO approach presented in [36] has the lowest performance, where the detection accuracy reaches 100% with a FAR of 24% which is relatively high rate of false alarms. On the other hand, the WMA approach used in conjunction with dynamic size sliding window in [37] reaches 100% for a FAR of 17%.

Furthermore, Smrithy *et al.* in [37] compare the performance of their approach with our previous work, such as linear SVM, MD and J48. We continue the comparison with supervised machine learning such as the SVM, K-NN (k=3), decision tree (J48) and the distance-based method MD in terms of TPR and FAR. The ROC of each method is presented in Figure 26, where they reach a TPR of 100% with FAR of: 5.2%, 6.8%, 7.8%, 8% and 15% for our approach, SVM, K-NN, J48 and MD respectively. In Figure 26, we cannot distinguish the performance of K-NN and J48 as they achieve similar performances. However, the K-NN was slower than J48 but provided better performance. K-NN is heavy for constrained WBANs, since it requires high computational complexity and large amounts of memory to store the training data, in contrast to other classification methods which build a model and discard the training data after the model's creation.

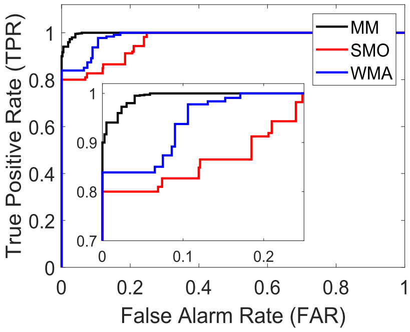

Fig. 25. Comparison with SMO & WMA.

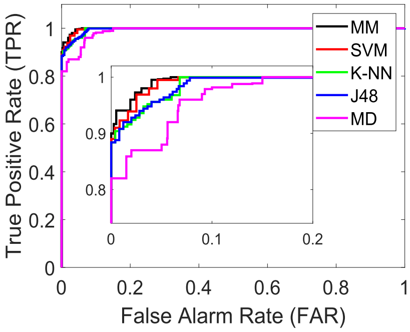

Fig. 26. Comparison with SVM, K-NN, J48 & MD.

The processing of our model can be divided into two steps: the first step is the forecasting, which is distributed to each sensor and is lightweight in terms of computation; the second step is the use of MM, which is centralized in the LPU to have access to all the collected data. At first sight, the use of MM for decision appear to be greedy in processing and in energy consumption and may hence seem heavy for WBANs. However, in our model, we exploit the Tukey box to fix the number of states in MM to 4 states, which reduces the parameters to 20: 4 values for the initial probability vector and a transition matrix of 16 (4×4) elements. Therefore, to derive the decision model, the complexity of our system is constant $O(1)$ when the complexity of the SVM model is $O(n^3)$ where n is the number of records in the training data. The update of the SVM model will drain the battery of any WBAN device in contrast to our 4-state MM. On the other hand, during the testing phase, the complexity of our model is $O(1)$ (similar to SVM), which makes it suitable for WBANs and for devices with constrained resources.

An anomaly detection system for ECG signal was proposed in [44] without fixing the number of states. The authors

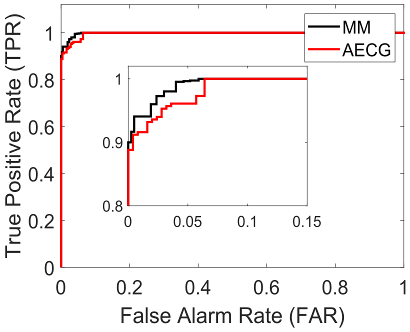

Fig. 27. Comparison with MM for abnormal ECG.

apply MM on ECG data, which is composed of 5 periodic waves: P, Q, R, S and T. In fact, the periodicity in ECG signal will limit the number of states in MM to n in their application (with $n \gg 4$). However, as we are processing physiological measurements without periodicity in our experiment, the values of measurements are real with continuous changes and the number of states in their proposed approach increases significantly (compared to our approach, where we fixed the number of states to 4). Moreover, deriving the state transition probabilities and initial probability vector requires more processing power and memory space.

Figure 27 shows the comparison result of our proposed approach with that of [44] (denoted by abnormal ECG or AECG) on the ECG signal from physionet. Their approach achieves a good detection accuracy with low false alarm rate (6.4%), but it is more complex and slower as the number of states is dynamic and larger than the 4 states fixed in our approach. The complexity resides in deriving the initial probability vector for n states and the transition probability matrix. For the same detection accuracy of 100%, our approach has a FAR of 5.2% whereas their approach incurs a slightly higher FAR of 6.4%. It is worth noting that we develop their model from scratch and as described in the paper without preprocessing of raw data or additional block.

With the increase of aging population in many countries, the costs of elderly healthcare will continue growing and will cripple the healthcare infrastructure with required new beds in hospitals for patients under long term monitoring. The use of remote monitoring systems will significantly reduce the number of occupied bed by person under monitoring and it will allow patients to be monitored while continuing their daily life activities using sensors and a smartphone.

A reliable remote monitoring system can also be used in elderly houses to reduce the number of healthcare professionals. Our proposed method would be suitable for assisting professionals and for evaluating the need for intervention with the required drug or medication. The cost of sensor devices is minor in context of such health applications. The industry is working today to manufacture active sensors, that not only capture and transmit data for remote processing, but also capable to process data in order to detect abnormal situations, such as devices for detecting epilepsy in [50], [51]. The fast growing number of people using such devices motivated by the added value of automated techniques for assisting doctors in making decisions will definitely create an opportunity for the industry to leverage economies of scale. In this context, preprocessing techniques used during data acquisition to improve the performance of the detection system by controlling (or reducing) the effects of noise will evidently enhance the reliability and consequently the widespread deployment of such devices.

V. Conclusion

In this article, we proposed an approach based on a Markov Model for the detection of anomalies in WBANs. The proposed system uses forecasting to reduces the energy consumption due to normal data transmission, and only measurements deviating from the forecasted values are transmitted to the LPU. This last derives a univariate time series using the root mean square of error between the forecasted and received values. The Markov Model is derived from the RMSE and used to check if a sequence of states is normal or not. To distinguish faults from abnormal physiological changes, the number of deviated attributes is checked before raising an alarm for healthcare professionals when at least r attributes are received by the LPU.

We applied our approach on real physiological data from a publicly available repository, and we conducted several experiments on data from different subjects for performance analysis. Our results show that the proposed system is able to achieve 100% detection accuracy with a low FAR of 5.2%. We further compared the performance of our approach with existing MM model for anomaly detection in ECG and with three supervised machine learning algorithms: SVM, K-NN and J48 and with MD. We found that our system slightly outperforms the abnormal ECG system, and achieves better accuracy than SVM, K-NN, J48 and MD.

In future work, we intend to investigate other forecasting algorithms by implementing more lightweight predictive models without significant differences in accuracy.

Acknowledgment

The authors acknowledge with thanks DSR for technical and financial support.

References

[1] M. A. G. Santos, R. Munoz, R. Olivares, P. P. R. Filho, J. D. Ser, and V. H. C. D. Albuquerque, “Online heart monitoring systems on the Internet of Health Things environments: A survey, a reference model and an outlook,” *Inf. Fusion*, vol. 53, pp. 222–239, Jan. 2020.

[2] C. Jory, R. Shankar, D. Coker, B. McLean, J. Hanna, and C. Newman, “Safe and sound? A systematic literature review of seizure detection methods for personal use,” *Seizure*, vol. 36, pp. 4–15, Mar. 2016.

[3] K. Abualsaud, A. Mohamed, T. Khattab, E. Yaacoub, M. Hasna, and M. Guizani, “Classification for imperfect EEG epileptic seizure in IoT applications: A comparative study,” in *Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2018, pp. 364–369.

[4] C. Wang *et al.*, “A low power cardiovascular healthcare system with cross-layer optimization from sensing patch to cloud platform,” *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 2, pp. 314–329, Apr. 2019.

[5] M. Simao, N. Mendes, O. Gibaru, and P. Neto, “A review on electromyography decoding and pattern recognition for human-machine interaction,” *IEEE Access*, vol. 7, pp. 39564–39582, 2019.

[6] T. W. Tseng, C. T. Wu, and F. Lai, “Threat analysis for wearable health devices and environment monitoring Internet of Things integration system,” *IEEE Access*, vol. 7, pp. 144983–144994, 2019.

[7] R. R. Swain, T. Dash, and P. M. Khilar, “A complete diagnosis of faulty sensor modules in a wireless sensor network,” *Ad Hoc Netw.*, vol. 93, Oct. 2019, Art. no. 101924.

[8] J. Ko *et al.*, “MEDiSN: Medical emergency detection in sensor networks,” *ACM Trans. Embed. Comput. Syst.*, vol. 10, no. 1, pp. 1–29, Aug. 2010.

[9] D. Malan, T. Fulford-jones, M. Welsh, and S. Moulton, “Codeblue: An ad hoc sensor network infrastructure for emergency medical care,” in *Proc. Workshop Appl. Mobile Embedded Syst. (MobiSys)*, 2004, pp. 12–14.

[10] K. F. Navarro, E. Lawrence, and B. Lim, “Medical MoteCare: A distributed personal healthcare monitoring system,” in *Proc. Int. Conf. eHealth, Telemedicine, Social Med.*, Feb. 2009, pp. 25–30.

[11] A. Wood *et al.*, “ALARM-NET: Wireless sensor networks for assisted-living and residential monitoring,” Dept. Comput. Sci., Univ. Virginia, Charlottesville, VA, USA, Tech. Rep. 2, 2006.

[12] S. Jiang *et al.*, “CareNet: An integrated wireless sensor networking environment for remote healthcare,” in *Proc. 3rd Int. ICST Conf. Body Area Netw.*, 2008, pp. 1–3.

[13] M. P. Rajasekaran, S. Radhakrishnan, and P. Subbaraj, “Sensor grid applications in patient monitoring,” *Future Gener. Comput. Syst.*, vol. 26, no. 4, pp. 569–575, Apr. 2010.

[14] 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 *Proc. 4th Int. ICST Conf. Pervas. Comput. Technol. Healthcare*, 2010, pp. 1–2.

[15] S. Tennina *et al.*, “WSN4QoL: WSNs for remote patient monitoring in e-health applications,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[16] M. A. A. da Cruz, J. J. P. C. Rodrigues, J. Al-Muhtadi, V. V. Korotaev, and V. H. C. de Albuquerque, “A reference model for Internet of Things middleware,” *IEEE Internet Things J.*, vol. 5, no. 2, pp. 871–883, Apr. 2018.

[17] B. C. P. Lau, E. W. M. Ma, and T. W. S. Chow, “Probabilistic fault detector for wireless sensor network,” *Expert Syst. Appl.*, vol. 41, no. 8, pp. 3703–3711, Jun. 2014.

[18] D. Praveen Kumar, T. Amgoth, and C. S. R. Annavarapu, “Machine learning algorithms for wireless sensor networks: A survey,” *Inf. Fusion*, vol. 49, pp. 1–25, Sep. 2019.

[19] H. Zhang, Q. Zhang, J. Liu, and H. Guo, “Fault detection and repairing for intelligent connected vehicles based on dynamic Bayesian network model,” *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2431–2440, Aug. 2018.

[20] H. Zhang, J. Liu, and A.-C. Pang, “A Bayesian network model for data losses and faults in medical body sensor networks,” *Comput. Netw.*, vol. 143, pp. 166–175, Oct. 2018.

[21] M. M. Nezhad and M. Eshghi, “Sensor single and multiple anomaly detection in wireless sensor networks for healthcare,” in *Proc. 27th Iranian Conf. Electr. Eng. (ICEE)*, Apr. 2019, pp. 1751–1755.

[22] O. Aziz *et al.*, “Validation of accuracy of SVM-based fall detection system using real-world fall and non-fall datasets,” *PLoS ONE*, vol. 12, no. 7, pp. 1–11, Jul. 2017.

[23] A. Chriki, H. Touati, and H. Snoussi, “SVM-based indoor localization in wireless sensor networks,” in *Proc. 13th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Jun. 2017, pp. 1144–1149.

[24] B. Ahmad, W. Jian, Z. A. Ali, S. Tanvir, and M. S. A. Khan, “Hybrid anomaly detection by using clustering for wireless sensor network,” *Wireless Pers. Commun.*, vol. 106, no. 4, pp. 1841–1853, Jun. 2019.

[25] H. Qu, L. Lei, X. Tang, and P. Wang, “A lightweight intrusion detection method based on fuzzy clustering algorithm for wireless sensor networks,” *Adv. Fuzzy Syst.*, vol. 2018, Jun. 2018, Art. no. 4071851.

[26] R. R. Guimaraes *et al.*, “Intelligent network security monitoring based on optimum-path forest clustering,” *IEEE Netw.*, vol. 33, no. 2, pp. 126–131, Mar. 2019.

[27] R. W. R. De Souza, J. V. C. De Oliveira, L. A. Passos, W. Ding, J. P. Papa, and V. Albuquerque, “A novel approach for optimum-path forest classification using fuzzy logic,” *IEEE Trans. Fuzzy Syst.*, early access, Oct. 28, 2019, doi: 10.1109/TFUZZ.2019.2949771.

[28] M. Xie, J. Hu, S. Guo, and A. Y. Zomaya, “Distributed segment-based anomaly detection with Kullback–Leibler divergence in wireless sensor networks,” *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 101–110, Jan. 2017.

[29] C. Titouna, F. Titouna, and A. A. A. Ari, “Outlier detection algorithm based on mahalanobis distance for wireless sensor networks,” in *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, Jan. 2019, pp. 1–6.

[30] H. Kumarage, I. Khalil, and Z. Tari, “Granular evaluation of anomalies in wireless sensor networks using dynamic data partitioning with an entropy criteria,” *IEEE Trans. Comput.*, vol. 64, no. 9, pp. 2573–2585, Sep. 2015.

[31] H. Harb, C. A. Jaoude, and A. Makhoul, “An energy-efficient data prediction and processing approach for the Internet of Things and sensing based applications,” *Peer-to-Peer Netw. Appl.*, vol. 13, no. 3, pp. 780–795, May 2020.

[32] U. Raza, A. Camerra, A. L. Murphy, T. Palpanas, and G. P. Picco, “Practical data prediction for real-world wireless sensor networks,” *IEEE Trans. Knowl. Data Eng.*, vol. 27, no. 8, pp. 2231–2244, Aug. 2015.

[33] H. H. Bosman, G. Iacca, A. Tejada, H. J. Wörtche, and A. Liotta, “Spatial anomaly detection in sensor networks using neighborhood information,” *Inf. Fusion*, vol. 33, pp. 41–56, Jan. 2017.

[34] Q. Yu, L. Jibin, and L. Jiang, “An improved ARIMA-based traffic anomaly detection algorithm for wireless sensor networks,” *Int. J. Distrib. Sensor Netw.*, vol. 12, no. 1, Jan. 2016, Art. no. 9653230.

[35] F. A. Aderohunmu, G. Paci, D. Brunelli, J. D. Deng, L. Benini, and M. Purvis, “An application-specific forecasting algorithm for extending WSN lifetime,” in *Proc. IEEE Int. Conf. Distrib. Comput. Sensor Syst.*, May 2013, pp. 374–381.

[36] S. Haque, M. Rahman, and S. Aziz, “Sensor anomaly detection in wireless sensor networks for healthcare,” *Sensors*, vol. 15, no. 4, pp. 8764–8786, Apr. 2015.

[37] G. S. Smrithy, R. Balakrishnan, and N. Sivakumar, “Anomaly detection using dynamic sliding window in wireless body area networks,” in *Data Science and Big Data Analytics*, D. K. Mishra, X.-S. Yang, and A. Unal, Eds. Singapore: Springer, 2019, pp. 99–108.

[38] D. S. Shukla, A. C. Pandey, and A. Kulhari, “Outlier detection: A survey on techniques of WSNs involving event and error based outliers,” in *Proc. Innov. Appl. Comput. Intell. Power, Energy Controls with their impact Humanity (CIPECH)*, Nov. 2014, pp. 113–116.

[39] A. Ayadi, O. Ghorbel, A. M. Obeid, and M. Abid, “Outlier detection approaches for wireless sensor networks: A survey,” *Comput. Netw.*, vol. 129, pp. 319–333, Dec. 2017.

[40] Y. Zhang, N. Meratnia, and P. Havinga, “Outlier detection techniques for wireless sensor networks: A survey,” *IEEE Commun. Surveys Tuts.*, vol. 12, no. 2, pp. 159–170, 2nd Quart., 2010.

[41] C. OReilly, A. Gluhak, M. A. Imran, and S. Rajasegarar, “Anomaly detection in wireless sensor networks in a non-stationary environment,” *IEEE Commun. Surveys Tuts.*, vol. 16, no. 3, pp. 1413–1432, 3rd Quart., 2014.

[42] H. Zhang and J. Liu, “Fault diagnosing ECG in body sensor networks based on hidden Markov model,” in *Proc. 10th Int. Conf. Mobile Ad-hoc Sensor Netw.*, Dec. 2014, pp. 123–129.

[43] J. Pang, D. Liu, Y. Peng, and X. Peng, “Collective anomalies detection for sensing series of spacecraft telemetry with the fusion of probability prediction and Markov chain model,” *Sensors*, vol. 19, no. 3, p. 722, Feb. 2019.

[44] F. A. Khan, N. A. H. Haldar, A. Ali, M. Iftikhar, T. A. Zia, and A. Y. Zomaya, “A continuous change detection mechanism to identify anomalies in ECG signals for WBAN-based healthcare environments,” *IEEE Access*, vol. 5, pp. 13531–13544, 2017.

[45] M. S. Z. Dehabadi and M. Jahed, “Reliability modeling of anomaly detection algorithms for wireless body area networks,” in *Proc. Iranian Conf. Electr. Eng. (ICEE)*, May 2017, pp. 70–75.

[46] N. Ding, H. Ma, H. Gao, Y. Ma, and G. Tan, “Real-time anomaly detection based on long short-term memory and Gaussian mixture model,” *Comput. Electr. Eng.*, vol. 79, Oct. 2019, Art. no. 106458.

[47] Y. Gao, C. Chen, J. Bu, W. Dong, and D. He, “ICAD: Indirect correlation based anomaly detection in dynamic WSNs,” in *Proc. IEEE Wireless Commun. Netw. Conf.*, Mar. 2011, pp. 647–652.

[48] *Physionet*. Accessed: Nov. 2019. [Online]. Available: http://www.physionet.org/physiobank/database/mimicdb

[49] A. K. Ramotra, A. Mahajan, R. Kumar, and V. Mansotra, “Comparative analysis of data mining classification techniques for prediction of heart disease using the Weka and SPSS modeler tools,” in *Smart Trends in Computing and Communications*. Singapore: Springer, 2020, pp. 89–97.

[50] *Epi-Care Free*. Danishcare Techology. Accessed: 2020. [Online]. Available: http://danishcare.dk/

[51] *Epileply Detect*. Accessed: 2020. [Online]. Available: http://www.epdetect.com/

Osman Salem received the M.Sc. and Ph.D. degrees in computer science from Paul Sabatier University, Toulouse, France, in 2002 and 2006, respectively, and the Habilitation à Diriger des Recherches degree from the University of Paris Descartes, Paris, France, in 2016.

From 2006 to 2008, he was with the Department of Computer Science, Telecom Bretagne, Brest, France, as a Post-Doctoral Research Fellow. Since September 2008, he has been an Associate Professor with the University of Paris Descartes. His research interest includes security and anomaly detection in medical wireless sensor networks.

Khalid Alsubhi (Member, IEEE) received the B.Sc. degree in computer science from King Abdulaziz University (KAU) in 2003 and the M.Math. and Ph.D. degrees in computer science from the University of Waterloo, Canada, in 2009 and 2016, respectively.

He is currently an Associate Professor of computer science with KAU. His research interests include network security and management, cloud computing, and security and privacy of healthcare applications.

Ahmed Mehaoua received the M.Sc. and Ph.D. degrees in computer science from the University of Paris, France, in 1993 and 1997, respectively.

He is currently a Full Professor of computer communication with the Faculty of Mathematics and Computer Science, University of Paris Descartes, Paris, France. He is also the Head of the Department of Multimedia Networking and Security at the LIPADE, a governmental computer science research center in Paris. His research interests include wireless healthcare systems and networks and applications.

Raouf Boutaba (Fellow, IEEE) received the M.Sc. and Ph.D. degrees in computer science from Sorbonne University, Paris, in 1990 and 1994, respectively. He is currently a University Chair Professor with the David R. Cheriton School of Computer Science, University of Waterloo, Canada, and holds an INRIA International Chair in Nancy, France. His research interests include network management and resource virtualization. He is a fellow of the Engineering Institute of Canada, The Canadian Academy of Engineering, and The Royal Society of Canada.