A Probabilistic model for early prediction of abnormal clinical events using vital sign correlations in home-based monitoring

Abdur Rahim Mohammad Forkan
School of Science (Computer Science), RMIT University
and National ICT Australia (NICTA)
Melbourne, VIC 3000, Australia
Email: s3361322@student.rmit.edu.au

Ibrahim Khalil
School of Science (Computer Science), RMIT University
and National ICT Australia (NICTA)
Melbourne, VIC 3000, Australia
Email: ibrahim.khalil@rmit.euu.au

***Abstract*—Chronic diseases are major causes of deaths in Australia and throughout the world. This necessitates the need for a self-care, preventive, predictive and protective assisted living system where a patient can be monitored continuously using wearable and wireless sensors. In real-time home monitoring system, various biological signals of a patient are obtained continuously using a mobile device (smart phone or tablet) and sent to the cloud to discover patient-specific abnormalities. The objective of this work is to develop a probabilistic model that identifies the future clinical abnormalities of a patient using recent and past values of multiple vital signs (e.g. heart rate, blood pressure, respiratory rate). Chronic patients living alone in home die of various diseases for the lack of an efficient automated system having prior prediction ability in the irregularities of vital signs. In this paper, Hidden Markov Model (HMM) is adopted to predict different clinical onsets using the temporal behaviours of six biosignals. The HMM models are trained and evaluated using continuous monitoring data of more than 1000 patients collected from the MIMIC-II database of MIT physiobank archive. The best models are selected using expectation maximisation (EM) algorithm and used in personalized remote monitoring system to forecast the most probable forthcoming clinical states of a continuously monitored patient. The scalable power of cloud computing is utilized for fast learning of various clinical events from large samples. The results obtained from the innovative home-based monitoring application show a new approach of detecting clinical anomalies using multi-parameter trends.**

I. Introduction

Due to advances in body sensor , wireless sensor and smart phone technologies and availability of cloud resources at low cost it is now possible to develop personalized remote health-monitoring [1] applications with greater flexibility [2]. In modern healthcare monitoring, the data from different biological signals of a patient are obtained continuously using wearable sensors and analysed in the cloud [3] to identify patient-specific knowledge [4]. The distributed resources of the cloud simplify the knowledge build up process from large biomedical data using computationally intensive machine learning methods [5]. The accurate and early anticipation of health-related abnormalities [6] is an essential functionality of a remote monitoring system to support clinicians in diagnostic decision making [7]. Timely intervention of symptoms is also very important in preventing clinical deterioration before an emergency situation.

All over the world, chronic lifelong diseases are increasing with growth of population and already created intensified pressure on overall healthcare infrastructure. Recent reports indicate that hospitals are getting overcrowded and are having difficulties in treating the patient even in emergency situation due to increasing population. A recent news shows that, patients (including elderly people aged over 80) in Western Sydney hospitals waited more than two days in the emergency department and had to leave without treatment [8]. According to that report, there was up to 40 hours waiting time on average to be admitted to the hospital. The capacity in hospitals is inadequate to make place for all patients. The situation will become worse for the patients as well as for the healthcare providers unless a personalized, scalable, real-time, prognostic home-based monitoring system [3] is adopted. The establishment of continuous clinical event prediction system for disease diagnosis [9] will also reduce the healthcare cost. It provides flexibility for patients by allowing them to do regular activities while biomedical data are continuously collected and fully relying to remote monitoring system.

Wearable sensors (e.g. Shimmer) with wireless communication capability are available in market at a very low cost. These sensors are able to collect various vital signs such as heart rate, blood pressure, respiratory rate, O_2 saturation (SPO_2), pulse, body temperature and electrocardiogram (ECG) [10]. One sensor can monitor one or multiple biological attributes (e.g. pulse oximeter can monitor both pulse and SPO_2). The measured values of these vitals can be exploited to detect different clinical events (e.g. Tachycardia, Hypotension, Hypoxia or more than one of these at the same time). For example, continuous heart rate value above 100 bpm is a sign of abnormality and known as Tachycardia [11]. Vital parameters are also correlated in time and space. Such correlations are sometime patient-specific and help to predict future clinical conditions at early stage. The accurate prediction requires the right interpretation of individual's vital parameters using past and current data. This is also beneficial for monitoring systems

978-1-4673-8779-8/16/$31.00 ©2016 IEEE

in order to avoid unnecessary false alerts.

Traditional hospital-focused model of care neglects monitoring and treating diseases at home. A number of intelligent monitoring systems exist for clinical abnormalities prediction for the patients who are confined to hospital beds [12], [13], [14], but a very few attempts [15], [16], [17] have been made to develop such predictive system for home that could prevent and minimize health-related risk at early stage. The monitoring systems those exist suffer from high false alert rates and depend on manual observations of vital parameters by the medical experts after an anomaly is detected. Some of the developed models can forecast the changes in a specific vital parameter with good accuracy [18], [19], however, it remains a challenge to build these models to monitor and correlate multiple biosignals while maintaining the interpretability of the discovered knowledge because biomedical data can evolve and vary over the lifetime of the system.

In this paper we present a clinical event prediction system using Hidden Markov Model (HMM) [20] to classify clinical states for a given observation containing multiple vital signs well in advance of the actual occurrence of that state. For each targeted clinical event an HMM is build using the observed behaviours of multiple vital signs before that clinical event. Vital signs data are generated as a sequence and have temporal dependency. The progression of clinical events depend on past behaviours of multiple vitals. HMM is a suitable model for learning from such sequential data [15]. The benefit of using HMM over other machine learning models is its ability to use temporal trends in a continuous data for classification and by modelling the temporal data as a sequence of *state* changes. Therefore, the probability of the clinical episode is not only dependent on patient's present vital sign values, but a sequence of measurements from the past as well.

In the rest of this paper, Section II discusses the related works in this study and our contributions. Section III describes the model scenario and its concepts, section IV explains the technical approaches applied for the model. Section V shows the experimental results and comparisons. Finally, section VI concludes this paper.

II. Related works and our contributions

The interest in analysing biomedical data has grown over the last decade. One particular focus is the analysis of correlations among multiple biosignals generated by wearable sensors for future abnormality prediction. Several studies have focused on finding correlations among different biological signals [21], [17], [14], [19], [22] such as ECG, blood pressure, heart rate, respiration and O_2 saturation. Some of these systems are capable of finding future abnormalities [17], [14] and irregularities in a specific vital parameter [23], [18], [24], [25]. These studies are mostly at theoretical level and still far behind to implement and apply them at application level. One of the practical example is BioSign device [26] that uses stochastic process to model the multiparameter data by fusing information of five vital signs. However, this technique does not have predictive capabilities and can only minimize the time of occurrence of critical clinical situation.

Machine learning techniques are widely adopted for designing classifier model to find clinical abnormalities and to detect symptoms of various chronic diseases [9]. Examples of such model include Hidden Markov Model (HMM) [15], [27], Support Vector Machine(SVM) [28], Neural Network [29], Topic model [30] etc. These systems can detect abnormalities and predict future behaviours in one or more vital signs.

The use of cloud computing platform in biomedical data analysis is also becoming a popular research area. Several models are developed for analysing and processing large amount health data such as ECG [31] in the cloud. The novelty of using cloud for context-aware monitoring and abnormality detection was described and developed in some recent works [2], [31] and our previous works [3], [15], [4].

All these contributions have motivated us to develop this predictive model for abnormal clinical event detection using pervasive computing that uses machine learning techniques in the cloud platform. The contributions of the paper are the following.

1) We develop a probabilistic prediction model that can determine the clinical nature of a patient using current and past data of multiple vital signs. Existing systems can detect anomalies in a single vital sign but ignore the interactions of multiple vital signs. Our developed system has the unique capability of using the correlations of many vital signs to detect anomalies, especially for home-based monitoring. We choose HMM to design our learning engine as this model perfectly suits for our problem formulation.
2) We use cloud computing frameworks for HMM learning and classification. We have tested the suitability and feasibility of the model in the cloud through experimental evaluations and showed that the system does not impose any additional penalty in terms of time and resources for learning from large data. Moreover, the utilization of cloud includes the flexibility of learning from many vital signs data of a very large number of patients with various clinical conditions. Therefore, our system has an exclusive ability of handling a large group of patients simultaneously using the cloud computing technology.
3) We develop a personalized monitoring system for ambient assisted living (AAL). In real-time classification the continuous vital sign data of a patient are sent to the cloud using mobile device and the probability of the occurrence of a clinical event in future time is continuously computed. In our system, multiple HMMs can be trained using various clinical cases. Moreover, the system is expandable in the sense that a new clinical case can be added with a new HMM training. The model is more adaptive, because HMM parameters can be updated with a new set of observations that maximize its probability. No remote monitoring system exists with such unique set of capabilities.

III. Scenarios and Concepts

We propose to build an integrated cloud-based assisted healthcare infrastructure that will add a great advantage to remote healthcare industry. This section describes the overall scenarios for the model, the terminologies and the concepts used to design the model.

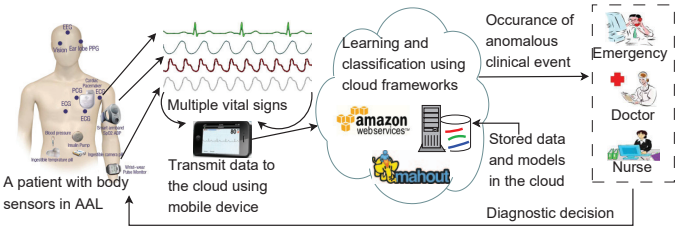

Fig. 1. The scenario of the developed system showing its components.

A. Model Scenario

The concept of the proposed system is visualized in Figure 1. We consider an ambient assisted living (AAL) environment where a patient live alone in home and his/her physiological conditions are continuously monitored by remote health monitoring stations. A number of wireless sensors are attached to patient's body that collect the information of different vital signs in form of biosignals, and then send the collected data to a portable device (e.g. smart phone). The portable device transmits these physiological data to the cloud in small batches for processing. The cloud, which has large storage and fast processing capability, run different algorithms with the help of various frameworks for data cleaning, segmentation, learning and knowledge discovery. Such system is described and validated in our previous works [3], [4]. The learned models of different clinical events are stored in cloud repositories for future classifications and predictions. The new incoming data from patients are then classified using the stored models. The suitable models are used for measuring future clinical states. If the system can predict the occurrence of an abnormal clinical event in near future, it sends notifications to the appropriate monitoring service. The monitoring person (e.g. doctor, nurse) can make diagnostic decision based on patient's clinical situation and send proper warning to the patient.

B. Vital sign and clinical event

In this study, we have considered four vital signs from six biosignals for defining a clinical event. These are listed in Table I. Here, we have used both MBP and SBP-DBP combination to ensure correct condition related to BP. Other vital signs are directly observed from biosignal trend values.

In this work, a clinical event means when one or more vital signs are not in expected range, not acted upon early enough and persist for a fixed period of time. The expected ranges of vital signs are patient-specific, although there are some general medical rules that define the normal and abnormal limit of different vital signs. For example, for a normal patient BP value should always be within normal range (SBP 80-120 and DBP 60-90 mmHg) according to the general medical rule. But for a hypertensive patient it is expected that most of the time his/her BP value is higher than the normal range. In this case, a clinical event occurs when this usual value (expected maximum BP value for that patient) rises to an abnormally high value. As in this particular work our focus is to describe the concepts and application areas of this model, so for simplification we have only used the general normal ranges of vital signs to define normality. In experimental evaluation only those clinical events are used where all 4 vital parameters go above or below general normal ranges.

TABLE I
The Vital Signs Used in This Study

Biosignal	acronym	unit	vital sign
Heart rate	HR	beast per minute (bpm)	HR
Systolic blood pressure	SBP	mmHg	BP
Diastolic blood pressure	DBP	mmHg	BP
Mean blood pressure	MBP	mmHg	BP
Respiratory rate	RR	breaths per minute	RR
blood oxyzen saturation	SPO_2	percentage	SPO_2

C. Training and Learning

The proposed model first learns about various clinical events from vital sign data of many patients. The behaviours and correlations of multiple vital signs preceding the targeted clinical events are used as features for model training. In next step, the models are trained in the cloud using large samples of normal and abnormal observations from patients with different clinical cases. The model learning process is shown in Figure 2 and briefly explained in section IV.

D. Real-time and personalised monitoring

The ultimate target is to use the proposed method in real-time home-based health monitoring system to classify the streaming data of multiple vitals of a particular patient. The continuous data of a new unknown patient are sent to the classification engine inside the cloud and continuous classification results and predictions are obtained over time based on probabilistic match. Therefore, our developed model can be used as personalized home-based monitoring. The scenario is shown in Figure 3. This is also described in section IV.

IV. Technical Approach

To implement the model scenarios described above we have used Hidden Markov Model (HMM) for learning and classification. Cloud platforms, frameworks and tools are used to learn from large data and for fast processing. These techniques are described in the following sections.

A. Constructing HMM from multiple biosignals

Hidden Markov Model (HMM) is the simplest form of dynamic bayesian network (DBN) that has been widely used in modelling and classification time-series data. It is a very popular machine learning method in the area of bioinformatics such as human activity classification [27], speech recognition [20], gene prediction [32], DNA motif discovery [33] etc.

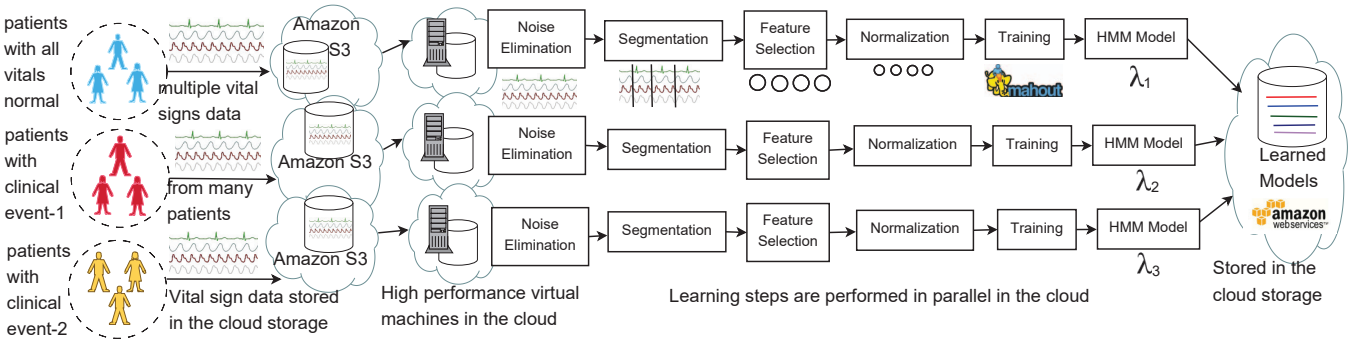

Fig. 2. The model learning process using multiple vital sign data from several group of patients (normal and with different clinical conditions). The learning phases are performed in the cloud. The output of this learning process is multiple HMMs and these learned models are stored in the cloud storage.

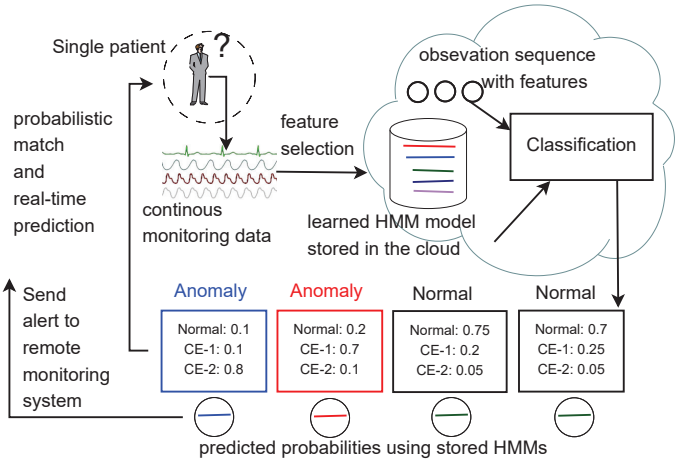

Fig. 3. Real-time clinical abnormality prediction for a patient with unknown clinical condition using multiple vital sign data. Here CE means a clinical event. The current observations are matched with stored HMMs and the probabilities of matching are continuously estimated.

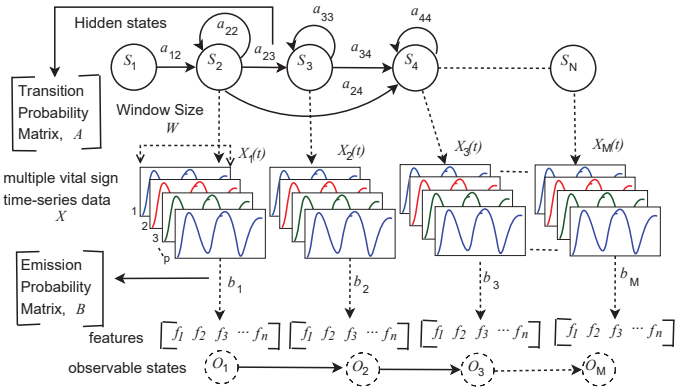

Fig. 4. The structure of Hidden Markov Model and its parameters for multiple vital sign time-series data

The structure of an HMM in our system using multiple vital sign time-series (e.g. HR, BP, RR, SPO_2) is shown in Figure 4. Let, the collection of discrete time-series data X of length T is sub-divided into M equally-sized window W. Each window has T/M samples, that is $W = \frac{T}{M}$. Therefore multiple vital sign time-series data can be considered as a sequence of M windows, $X_1(t), X_2(t), X_3(t), ...X_M(t)$. A feature vector $(f_1, f_2, f_3, ..., f_n)$ is constructed using the samples of each window. The features are generated from the statistical natures of the biosignals and their correlation coefficients. If x and y are two biosignals and n is number of samples, then the correlation coefficients r is defined by (1).

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (1)$$

This results a sequence of M-length feature vectors which is the observation sequence ($O = O_1, O_2, O_3, ...O_M$) of HMM. The future observations are some probabilistic functions of past observations and so the observations have temporal dependency. If M observations can be described by N distinct states $S_1, S_2, S_3, ...S_N$ (as in Figure 4) then it forms a Markov process. These states are called the hidden states and are not directly visible but the observable outputs (O) depended on the states are visible. The change of one state to another depends on certain transition probabilities associated with that state. These state-transition probabilities form a matrix $A = \{a_{ij}\}$. The transition probability from state S_i at time t_{n-1} to the state S_j at time t_n is,

$$a_{ij} = P[X(t_n) = S_j | X(t_{n-1}) = S_i] \quad (2)$$

Each $a_{ij} \geq 0$ and $\sum_{j=1}^{M} a_{ij} = 1$. $(i = 1, 2, .., N; j = 1, 2, ..., N)$.

The observation O_j (generated from the features of multiple vital signs), when the system is in state S_i at time t, depends on observation probability distributions which form a matrix B=$\{b_j\}$. They are called the emission probabilities $\{b_j\}$. The initial state distribution is described by a matrix $\pi = \{\pi_i\}$, where π_i is the probability of the starting state being i. Using A, B and π it is possible to come up with a state transition sequence Q=$q_1, q_2, ...q_M$ for the given HMM, which can best explain a given observation sequence O=$O_1, O_2, ..., O_M$ that is formed using multiple vital signs. Thus, an HMM (λ), constructed from multiple biosignals is described by the notation in 3.

$$\lambda = (A, B, \pi) \tag{3}$$

The first step of constructing an HMM is to compute these model parameters (A, B, and π). These are described in the following section.

B. Segmentation and problem formulation

We consider p vital signs (as in Figure 4) of a patient are continuously monitored. The collected measurement at any given time t is $x_t = \{x_{t,1}, x_{t,2}, ..., x_{t,p}\}$. x_t is a row of data matrix X where each row indicates the observed values of each vital sign at t. The columns indicate the indices of vital signs, that is, each column j is the time-series data of j-th vital sign.

X_{T_i} is a batch of continuous data X. T is the time period of data collected from X and also called the batch size. Therefore, X can be considered as a series of continuous batches $\{X_{T_1}, X_{T_2}, X_{T_3}, ..., X_{T_i}, ...\}$. Each of the batch can be subdivided into M equally spaced window of size W as discussed in Figure 4.

To formulate this data for supervised learning a list of known clinical events are targeted. These clinical cases are used as prior knowledge and identified visually by clinical experts. In our case, these events are identified using liner search over all data because the whole bunch of data were available. We consider that, if a clinical event starts at time t and persist for a fixed period of time T_p then the batch data immediately before t are used for generating feature vector for supervised learning. This is because, the variations in the physiological data just before the actual clinical onset has high impact over that clinical condition. The batch data before t is called observation window and T_p is called prediction window. The whole process is described in Figure 5. If there is no clinical event (e. g. when all vital sign values are normal) then a random batch is picked as observation window.

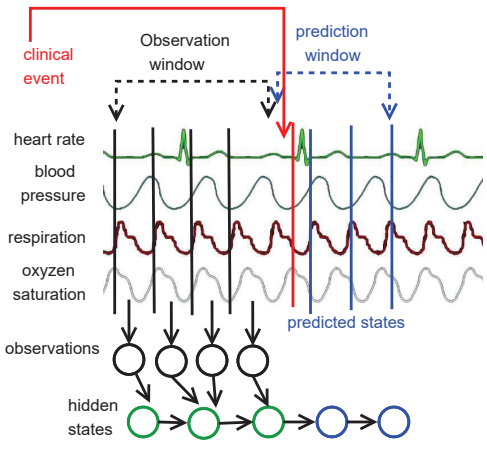

Fig. 5. The process of segmentation for constructing an HMM

A set of features f is calculated for each window W in observation window. Each f is an observable state O_i and thus each batch of observation window converted to a set of observations $O = O_1, O_2, ..., O_M$. An HMM is trained with large number of such observation sequences of multiple patients so that it can estimate the hidden states for a new observation. The overall process is depicted in Figure 5.

If a system wants to predict a clinical event from m number of known clinical cases then we need to generate m HMMs. i-th HMM represents the model for i-th clinical case (i=1, 2, ..., m). Therefore, i-th HMM is required to train with a large number of observations computed using the process described above. The procedures of training and classification are described in the next section.

C. HMM learning in the cloud

To construct an HMM λ, the main computation is performed in estimating model parameters A, B and π. With a random initial guess for model parameters the HMM is trained with a large number of observation samples from many patients for the targeted clinical condition. At first, noisy data are filtered using median-pass [34] and nearest-neighbour filter and then segmentation is performed (as described in previous section). Afterwards, statistical and correlation features are computed from the segmented samples and they are normalized to discrete values. In the training phase, the model parameters are adjusted to maximize the probability $P(O|\lambda)$. For training we have used BaumWelch algorithm [35], [36], the most popular algorithm used for HMM training. For all m clinical events m HMM model parameters are computed (as in Figure 2).

Training an HMM with a large amount of observation samples is computationally expensive. The BaumWelch algorithm uses expectation maximization (EM) [35] process for maximum probability estimation. The EM learning is slow if it is conducted in a local device. A real-time monitoring system demands quick learning. According to the proposed model, patients are geographically distributed and their continuous monitoring data are sent to the cloud.

Using cloud platforms such as Amazon Web Service [37] it is possible to create and configure high performance virtual machine (Amazon EC2) instances at low cost and store large amount of data (in Amazon S3 buckets [37]). In our approach for HMM model learning, large data samples from many patients are stored in Amazon S3 storage buckets. For supervised learning the data with similar clinical conditions are grouped together and stored in Amazon S3. When sufficient amount of samples for each target clinical event become available, the steps of computing features and HMM training are run on high performance virtual machines configured using Amazon EC2. Apache Mahout [38], a scalable machine learning library has procedures for HMM training and classification. For each m clinical events m HMMs are trained in parallel in m virtual machines (Amazon EC2 instances). Therefore, we get the estimated parameters of all the m HMM models within a very short time. The learned parameters of all HMMs are stored in Amazon S3. This scenario was described in Figure 2. The cloud platforms simplify HMM learning process and make computation faster.

D. Classification and real-time probabilistic prediction

In patient classification phase, a new observation O is classified to class i according the maximum value of $logP(O|\lambda_i)$ where i=1,2,...m computed by EM algorithm [35]. In real-time monitoring, data of a patient with unknown clinical condition are continuously sent to the cloud (Amazon EC2 instances in our case) where segmentations and feature computations are performed. The computed features are then sent to the HMM classifiers (stored in the cloud) and continuous probability estimations of all HMMs are obtained. This process was described in Figure 3. The clinical condition of next prediction window is selected based on the maximum value of $logP(O|\lambda_i)$. Alerts are sent to the remote monitoring systems when any abnormal clinical event is detected using probabilistic estimations. This is how, the developed HMMs are used to classify and predict continuous monitoring data.

V. Experimental results and discussions

To validate the applicability of the proposed model described above, a prototype for clinical event detection and prediction was developed using real medical data and cloud frameworks. In this section, the results obtained from the experimental evaluations of the prototype are described.

A. Data preparation

To create and test the learning models that can predict clinical events, vital sign data from MIMIC-II numeric dataset [39] of MIT physiobank archive [40] were used. This dataset was preferred because it fulfilled the criteria for model development and evaluations. Moreover, there is no public dataset available which contain vital sign data from a large number of patients. The second version of MIMIC-II numeric dataset contains records of various physiological signals of 3704 adult ICU patients. The records, those contain at least 24 hours numerical trend data of six biosignals (HR, Systolic BP, Diastolic BP, Mean BP, RR and SPO_2) were used. Most of the data are sampled per minute. Some were sampled per second and those are converted to per minute sampling by taking the mean value in a minute. The data with consecutive missing values over long period were eliminated and noisy data were filtered. Finally, only 1023 records met our criteria which contain clean continuous monitoring data of six biosignals for more than 24 hours.

B. Experimental setup

Four vital signs are used for anomaly detection using six biosignals. The threshold values of different clinical types are listed in Table II. The values other than these are considered as normal according to the general medical concept [41], [42]. Also a list of 5 clinical events that targeted for the experiments are shown in Table III. The duration of an episode is considered for 30 minutes using similar concept defined for detecting acute hypotension episode in physionet CinC challenge 2009 [43].

In our prototype, we only consider the events where irregularities occur in multiple vital signs at the same time. The abnormality in multiple vital signs at the same time cause serious clinical emergency such as heart failure, myocardial infraction and respiratory failure [41]. As the goal is to detect clinical abnormalities due to the changes in multiple vital signs, so we considered only the events where all four vital signs deviate from normality. Such abnormality can occur also in home settings. The clinical events are selected based on the availability of adequate amount of samples in the selected dataset so that we have enough samples to perform training and testing. In some cases, multiple samples are collected from the same records to increase the sample size. The amount of samples are chosen carefully so that each clinical event has nearly similar distribution. The sample size is also shown in Table III.

TABLE II
General threshold values of various clinical condition

Clinical condition	**Reason**	**Threshold values**
Hypertension	high BP	(SBP $\geq$ 120 and DBP $\geq$ 80) or MBP $\geq$ 105
Hypotension	low BP	(SBP $\leq$90 and DBP $\leq$ 60) or MBP $\leq$ 70
Tachycardia	High HR	HR $\geq$ 100
Bradycardia	Low HR	HR $\leq$ 60
Tachypena	High RR	RR $\geq$ 17
Bradypena	low RR	RR $\leq$ 12
Hypoxia	low SPO_2	SPO_2 $\leq$ 93%

TABLE III
The occurrence of clinical event targeted for the evaluation and their sample size

Clinical event	**Acronym**	**Class label**	**Number of samples**
All values are in normal range	NNNN	0	700
Tachycardia, Hypotension, Tachypena and Hypoxia at the same time	THTH	1	500
Bradycardia, Hypotension, Tachypena and Hypoxia at the same time	BHTH	2	340
Tachycardia, Hypertension, Tachypena and Hypoxia at the same time	TTTH	3	470
Tachycardia, Hypotension, Bradypena and Hypoxia at the same time	THBH	4	410

C. Modelling HMMs

According to the experimental setup, our classifier consists of 5 HMMs (λ_{NNNN}, λ_{THTH}, λ_{BHTH}, λ_{TTTH}, λ_{THBH}). For each case, 70% samples were used for training and the rest for testing (e.g. for NNNN 490 samples for training and 210 samples for testing). Since our main goal is the prediction of abnormal clinical events in advance, we want to model each λ_i using the sequence of data just before the clinical episode. Therefore, the features of each sample of each λ_i are built using 60 minutes data preceding the targeted clinical event.

The 60 minutes data of six signals are divided into six equally-sized windows of 10 samples each. We have computed 5 statistical measures (mean, standard deviation, median, number of increasing and decreasing trends) [14], [44] and pairwise correlations of 4 vital signs, resulting 36 features

in total. Therefore, each 60 minutes sample is converted to a sequence of six 36-dimensional feature vector. To simplify the computation of HMMs, each feature is discretized and normalized to nominal value between 0 to 9. Afterwards, the 5 HMMs are trained in Apache Mahout [45] using the normalized features of the training samples as described. The number of hidden states for each model is chosen experimentally to have the maximum flexibility of modelling, but to avoid overfitting. This was described in one of our previous work [15].

Here the HMM-based classification system serves multiple purposes. It can tell which one of possible m clinical event a patient will be in prediction window T_p using temporal data immediately before that clinical event. Once it knows the possible clinical event it can estimate the most likely future states for a given observation. This helps to identify future abnormalities. Moreover, this can be used in real-time monitoring by estimating maximum likelihood probability of all m clinical events in every minute. No other machine learning model can serve such multiple purpose using temporal trend of multiple vital sign data.

D. Clinical event classification

To classify an unknown test observation sequence O into one of the 5 clinical classes (as in Table III), $\log P(O|\lambda)$ is calculated using forward-backward procedure [46] and compared which HMM describes the sequence better. That is, an unknown observation is assigned to i-th class, such that, $i = argmax_i[p|\lambda_i]$, $i \in$ {NNNN, THTH, BHTH, TTTH, THBH}

The confusion matrix from the classification result is shown in Table IV.

TABLE IV
THE CONFUSION MATRIX OBSERVED AFTER PERFORMING THE CLASSIFICATION USING 5 HMMS. HERE 97.8% ACCURACY IS OBTAINED

	NNNN	THTH	BHTH	TTTH	THBH
NNNN	206	1	2	1	0
THTH	3	146	0	1	0
BHTH	1	0	101	0	0
TTTH	4	0	0	137	0
THBH	2	1	0	0	120

From this result, we see that if HMMs are properly trained with a large number of samples then they can isolate the normal and abnormal clinical events with a very good accuracy (which is 97.8% in our evaluation). Therefore, the developed HMMs can identify different clinical events using incoming data of unknown observations.

E. Prediction in continuous monitoring

The next experiment is to use the develop HMMs to classify clinical situations using continuous data of multiple vital signs of a particular patient so that it can be used in real-time monitoring. 6 patients are picked for the experiment. One of the patient has 2 clinical events (THTH, THBH) in whole the data including the normal (NNNN) episodes. One patient is with fully normal data and the rest 4 patients have combination of normal and abnormal data (each 4 patient is from one of the 4 clinical case used in the evaluation).

The objective here is to send the data continuously as a batch of 60 minutes, convert it to observation sequence, feed it to HMM classification engine, and keep getting continuous classification result for next 30 minutes window. In continuous data classification, 5 probability ($\log P(O|\lambda)$) values (for five HMMs) are obtained for a 60-minute observation and decision is made for the next 30-minute segment by taking the maximum probability value (log likelihood). The value can be in one of the classes (in our case from 0 to 4) based on the probability calculation (shown in Figure 3).

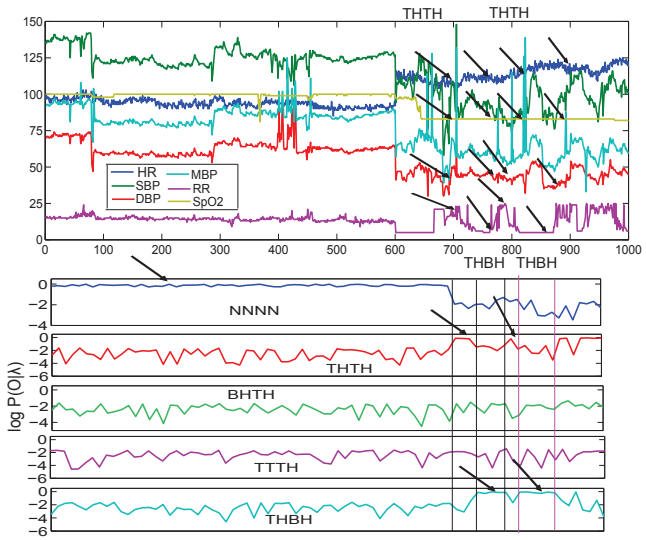

Fig. 6. The upper plot shows the continuous monitoring values of 6 biosignals. The lower plot depicts $logP(O|\lambda)$ values obtained for 5 HMMs for patient a40943n over 16 hours (1000 minutes) period. The changes in multiple vital signs and corresponding probabilities are indicated by arrows.

Figure 6 shows the result obtained for such an experiment using more than 16 hours (1000 minutes) data of patient a40943n. Three types of clinical events (NNNN,THTH,THTB) were observed for this patient. The abnormalities (e.g. change in multiple vital signs here) occur after 600 minutes. In Figure 6, the upper plot shows the vital signs values and the lower plot depicts the calculated log likelihood values by 5 HMMs. We can see, there is not much long term deviation in multiple vitals for first 600 minutes and so most of case $logP(O|\lambda)$ values of NNNN is much higher than the others. We see two THTH episodes (one after 700 minutes and another near 900 minutes) and two THBH episodes in the plot (one after 800 minutes and another after 900 minutes) and corresponding predicted log likelihood values go high. Similar behaviours were observed for other patients data. This proves that, our model can predict the probability of abnormal clinical event well ahead of time to avoid potential clinically dangerous situations.

F. Evaluation using cloud frameworks

To evaluate the performance of the system in the cloud the HMM training time for case NNNN is measured in 6 types of Amazon EC2 instance as well as on a local machine.

The observation sample size was 100 (each sample was 36-dimensional feature vector). The HMM implementation of Apache mahout framework is used for evaluation in the cloud. The obtained result is shown in Table V.

TABLE V
HMM TRAINING TIME FOR CASE NNNN IN DIFFERENT AMAZON EC2 INSTANCES

Type	**No of vCPUs**	**Memory (GB)**	**Training time (min)**
t2.micro	1	1	129
t2.small	1	2	102
t2.medium	2	4	67
m3.large	2	7.5	41
m3.xlarge	4	15	30
Local machine	1	8	86

Form the above results we can see that the model training time improves with the increment of the resources in a virtual instance. The training time on a local machine with 8GB RAM took 86 minutes. The lowest training time obtained is 30 minutes for a m3.xlarge virtual machine with 4 CPUs and 8GB RAM. That is, training time decreased almost one third on a high performance virtual machine. Trivially, the more powerful virtual machine instance will reduce more training time. Therefore, this proves that the cloud implementation for model training improve the model performance.

Moreover, to increase parallelism 5 HMMs are trained in 5 different m3.large virtual instances at the same time and we got the model parameters of all HMMs in short time. In sequential approach on a local machine or on a hospital server it requires to train one model at a time which takes very large time to estimate all the model parameters.

G. Comparison with other models

Although there are differences in the way of experimental setup for our model, the dataset, platform we have used and the application area we have targeted in compare with other similar models, it is still possible to draw some performance and feature-wise comparisons. Therefore, we compared our model with the 3 existing works. The work in [28] describes a hospital-focused probabilistic model using Gaussian mixture model (GMM) and one-class support vector machine (SVM) to identify patient deterioration using 4 vital signs (HR, BP, RR, SPO_2) that we have also used in this study. Others have used fewer biosignals for predictions. This comparison is presented in Table VI.

From the comparisons we can conclude that, our model has more impressive features and results. It is built on large samples and capable of predicting many clinical events more accurately.

TABLE VI
COMPARISON OF OUR MODEL WITH THE RESULTS AND FEATURES OBTAINED IN SIMILAR WORKS

	Our model	Work in [28]	Work in [18]	Work in [14]
Number of biosignals	6	6	1	2
Vital signs	HR, BP, RR, SPO_2	HR, BP, RR, SPO_2	BP	HR, BP
Clinical event	any	any	only Acute hypotension	Hemodynamic instability
Number of normal samples	700	1370	30	571
Number of abnormal samples	1720	130	30	116
Accuracy	97.7%	max 95% (GMM) & 96% (SVM)	94%	ROC max 0.86
Early prediction capability	yes	yes	yes	yes
Continuous monitoring support	yes	yes	yes	yes
Cloud implementation	yes	no	no	no

VI. CONCLUSION AND FUTURE WORK

In this work, we have presented a prediction model for early detection of anomalous clinical episodes caused by multiple vital signs. Hidden Markov Model is used for learning and classification of various clinical events from the behaviours of multiple vital signs. This model is also suitable for continuous monitoring of a patient's health using the information learned from the changes in many physiological parameters. The advantages of cloud computing are utilized for scalable and faster learning and real-time monitoring. This model is easily deployable in home-based remote monitoring system and capable of handling vital sign data from many patients. The proposed work is a valuable contribution to the healthcare industries as they are struggling to meet the demand of increasing population and chronic illness. This systemic prediction model can reduce healthcare cost and chronic disease related deaths. We have showed through prototyping and experimental evaluations that the model is able to forecast clinical episodes and can generate real-time probabilistic estimation of anomalies with very good accuracies. Thus, the model also simplifies the jobs of healthcare professionals by sending them early warnings and by reducing false predictions.

As a part of ongoing work, we are interested to utilize this model to predict clinical episodes related to cardiovascular disease and diabetes. We plan to implement the scenarios with big data from a large number of vital signs of many patients.

ACKNOWLEDGMENT

The authors wish to acknowledge the support of National ICT Australia (NICTA) and RMIT University for funding the research work presented in this paper.

REFERENCES

[1] K. A. Sidek, , I. Khalil, and H. F. Jelinek, "Ecg biometric with abnormal cardiac conditions in remote monitoring system," *IEEE Transactions on Systems Man and Cybernetics*, vol. 44, no. 11, pp. 1498–1509, 2014.

[2] D. Hoang and L. Chen, "Mobile cloud for assistive healthcare (mocash)," in *IEEE Asia-Pacific Services Computing Conference (APSCC), 2010*. Ieee, 2010, pp. 325–332.

[3] A. Forkan, I. Khalil, and Z. Tari, "Cocamaal: A cloud-oriented context-aware middleware in ambient assisted living," *Future Generation Computer Systems*, vol. 35, pp. 114–127, 2014.

[4] A. Forkan, I. Khalil, A. Ibaida, and Z. Tari, "Bdcam: Big data for context-aware monitoring-a personalized knowledge discovery framework for assisted healthcare," *IEEE Transactions on Cloud Computing*, 2015.

[5] I. H. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.

[6] D. Riboni, C. Bettini, G. Civitarese, Z. H. Janjua, and R. Helaoui, "Finegrained recognition of abnormal behaviors for early detection of mild cognitive impairment," in *Proc. of IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2015, pp. 149–154.

[7] F. Sufi, Q. Fang, I. Khalil, and A. Mahmoud, "Novel methods of faster cardiovascular diagnosis in wireless telecardiology," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 4, pp. 537–552, 2009.

[8] J. Robertson. (2015, January) Nsw health: Two-day wait for hospital bed the norm. [Online]. Available: http://www.smh.com.au/national/nsw-health-twoday-wait-for-hospital-bed-the-norm-20150113-12lbzs.html

[9] F. Sufi and I. Khalil, "Diagnosis of cardiovascular abnormalities from compressed ecg: A data mining-based approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 1, pp. 33–39, 2011.

[10] V. S. Tseng, L.-C. Chen, C.-H. Lee, J.-S. Wu, and Y.-C. Hsu, "Development of a vital sign data mining system for chronic patient monitoring," in *International Conference on Complex, Intelligent and Software Intensive Systems, 2008. CISIS 2008.* IEEE, 2008, pp. 649–654.

[11] A. Ibaida and I. Khalil, "Distinguishing between ventricular tachycardia and ventricular fibrillation from compressed ecg signal in wireless body sensor networks," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2010*. IEEE, 2010, pp. 2013–2016.

[12] C. R. Leite, G. Sizilio, A. Neto, R. Valentim, and A. Guerreiro, "A fuzzy model for processing and monitoring vital signs in icu patients," *BioMedical Engineering Online (Online)*, vol. 10, p. 68, 2011.

[13] J. C. Ho, C. H. Lee, and J. Ghosh, "Imputation-enhanced prediction of septic shock in icu patients," in *Proceedings of the ACM SIGKDD Workshop on Health Informatics (HI-KDD12)*, 2012.

[14] H. Cao, L. Eshelman, N. Chbat, L. Nielsen, B. Gross, and M. Saeed, "Predicting icu hemodynamic instability using continuous multiparameter trends," in *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008. EMBS 2008*. IEEE, 2008, pp. 3803–3806.

[15] A. R. M. Forkan, I. Khalil, Z. Tari, S. Foufou, and A. Bouras, "A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living," *Pattern Recognition*, vol. 48, no. 3, pp. 628–641, 2015.

[16] A. Copetti, J. C. Leite, O. Loques, and M. F. Neves, "A decision-making mechanism for context inference in pervasive healthcare environments," *Decision Support Systems*, vol. 55, no. 2, pp. 528–537, 2013.

[17] O. Salem, Y. Liu, A. Mehaoua, and R. Boutaba, "Online anomaly detection in wireless body area networks for reliable healthcare monitoring," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 5, pp. 1541–1551, 2014.

[18] T. Rocha, S. Paredes, P. Carvalho, J. Henriques, and M. Harris, "Wavelet based time series forecast with application to acute hypotensive episodes prediction," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2010*. IEEE, 2010, pp. 2403–2406.

[19] A. Aboukhalil, L. Nielsen, M. Saeed, R. G. Mark, and G. D. Clifford, "Reducing false alarm rates for critical arrhythmias using the arterial blood pressure waveform," *Journal of biomedical informatics*, vol. 41, no. 3, pp. 442–451, 2008.

[20] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[21] G. Manley, M. M. Knudson, D. Morabito, S. Damron, V. Erickson, and L. Pitts, "Hypotension, hypoxia, and head injury: frequency, duration, and consequences," *Archives of Surgery*, vol. 136, no. 10, pp. 1118–1123, 2001.

[22] X. He, R. A. Goubran, and X. P. Liu, "Evaluation of the correlation between blood pressure and pulse transit time," in *IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA), 2013*. IEEE, 2013, pp. 17–20.

[23] Q. Li, R. G. Mark, G. D. Clifford *et al.*, "Artificial arterial blood pressure artifact models and an evaluation of a robust blood pressure and heart rate estimator," *Biomedical engineering online*, vol. 8, no. 1, p. 13, 2009.

[24] C. M. Ennett, K. Lee, L. J. Eshelman, B. Gross, L. Nielsen, J. J. Frassica, and M. Saeed, "Predicting respiratory instability in the icu," in *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008. EMBS 2008.* IEEE, 2008, pp. 2848–2851.

[25] R. Ravichandran, E. Saba11, K.-Y. Chen, M. Goel, S. Gupta, and S. N. Patel, "Wibreathe: Estimating respiration rate using wireless signals in natural settings in the home," pp. 131–139, 2015.

[26] L. Tarassenko, A. Hann, A. Patterson, E. Braithwaite, K. Davidson, V. Barber, and D. Young, "Biosign: Multi-parameter monitoring for early warning of patient deterioration," in *The 3rd IEE International Seminar on Medical Applications of Signal Processing, 2005*. IET, 2005.

[27] F. J. Ordóñez, P. de Toledo, and A. Sanchis, "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors," *Sensors*, vol. 13, no. 5, pp. 5460–5477, 2013.

[28] L. Clifton, D. A. Clifton, P. J. Watkinson, and L. Tarassenko, "Identification of patient deterioration in vital-sign data using one-class support vector machines," in *Federated Conference on Computer Science and Information Systems (FedCSIS), 2011*. IEEE, 2011, pp. 125–131.

[29] J. Henriques and T. Rocha, "Prediction of acute hypotensive episodes using neural network multi-models," in *Computers in Cardiology, 2009*. IEEE, 2009, pp. 549–552.

[30] S. Saria, D. Koller, and A. Penn, "Learning individual and population level traits from clinical temporal data," in *Proc. Neural Information Processing Systems (NIPS), Predictive Models in Personalized Medicine workshop*. Citeseer, 2010.

[31] X. Wang, Q. Gui, B. Liu, Z. Jin, and Y. Chen, "Enabling smart personalized healthcare: a hybrid mobile-cloud approach for ecg telemonitoring," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 18, pp. 739–745, 2014.

[32] G. Rätsch, S. Sonnenburg, J. Srinivasan, H. Witte, K.-R. Müller, R.-J. Sommer, and B. Schölkopf, "Improving the caenorhabditis elegans genome annotation using machine learning," *PLoS Computational Biology*, vol. 3, no. 2, p. e20, 2007.

[33] K.-C. Wong, T.-M. Chan, C. Peng, Y. Li, and Z. Zhang, "Dna motif elucidation using belief propagation," *Nucleic acids research*, vol. 41, no. 16, pp. e153–e153, 2013.

[34] H. Cao, D. E. Lake, J. E. Ferguson, C. A. Chisholm, M. P. Griffin, J. R. Moorman *et al.*, "Toward quantitative fetal heart rate monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 1, pp. 111–118, 2006.

[35] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *The annals of mathematical statistics*, vol. 41, no. 1, pp. 164–171, 1970.

[36] L. R. Welch, "Hidden markov models and the baum-welch algorithm," *IEEE Information Theory Society Newsletter*, vol. 53, no. 4, pp. 10–13, 2003.

[37] Amazon web service. [Online]. Available: http://aws.amazon.com/

[38] Apache mahout. [Online]. Available: http://mahout.apache.org/

[39] M. Saeed, M. Villarroel, A. T. Reisner, G. Clifford, L.-W. Lehman, G. Moody, T. Heldt, T. H. Kyaw, B. Moody, and R. G. Mark, "Multiparameter intelligent monitoring in intensive care ii (mimic-ii): a public-access intensive care unit database," *Critical care medicine*, vol. 39, no. 5, p. 952, 2011.

[40] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[41] L. W. . Wilkins, *ECG Interpretation Made Incredibly Easy (5th Edition)*. Lippincott Williams & Wilkins, 2011.

[42] B. P. McGrath *et al.*, "Ambulatory blood pressure monitoring." *The Medical Journal of Australia*, vol. 176, no. 12, pp. 588–592, 2002.

[43] (2009) Predicting acute hypotensive episodes, a challenge from physionet and computers in cardiology 2009. [Online]. Available: http://physionet.org/challenge/2009/

[44] A. Janghorbani, A. Arasteh, and M. H. Moradi, "Prediction of acute hypotension episodes using logistic regression model and support vector machine: A comparative study," in *19th Iranian Conference on Electrical Engineering (ICEE), 2011*. IEEE, 2011, pp. 1–4.

[45] Hidden markov models using apache mahout. [Online]. Available: https://mahout.apache.org/users/classification/hidden-markov-models.html

[46] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys (CSUR)*, vol. 46, no. 3, p. 33, 2014.