An Early Respiratory Distress Detection Method with Markov Models

Hariharan Ravishankar[1], Aditya Saha[1], Gokul Swamy[2] and Sahika Genc[3]

Abstract— **A method for early detection of respiratory distress in hospitalized patients which is based on a multi-parametric analysis of respiration rate (RR) and pulse oximetry (SpO2) data trends to ascertain patterns of patient instability pertaining to respiratory distress is described. Current practices of triggering caregiver alerts are based on simple numeric threshold breaches of SpO2. The pathophysiological patterns of respiratory distress leading to in-hospital deaths are much more complex to be detected by numeric thresholds. Our pattern detection algorithm is based on a Markov model framework based on multi-parameter pathophysiological patterns of respiratory distress, and triggers in a timely manner and prior to the violation of SpO2 85-90% threshold, providing additional lead time to attempt to reverse the deteriorating state of the patient. We present the performance of the algorithm on MIMIC II dataset resulting in true positive rate of 92% and false positive rate of 6%.**

I. INTRODUCTION

National Registry of Cardiopulmonary Resuscitation (NRCPR), an American Heart Association (AHA)-sponsored, prospective, multisite, observational study of in-hospital resuscitation, between January 1, 2000, and June 30, 2002, 14720 cardiac arrests occurred [1]. By mid-2010, 183,749 cardiopulmonary resuscitation events were in the registry [2]. The three most common reasons for cardiac arrest in adults were (1) cardiac arrhythmia, (2) acute respiratory insufficiency, and (3) hypotension. Despite the fact that a primary arrhythmia was one of the precipitating events in nearly one half of adult cardiac arrests, ventricular fibrillation (VF) was the initial pulseless rhythm in only 16% of in-hospital cardiac arrest victims [1].

An abnormal respiratory rate has been shown to be an important predictor of serious events such as cardiac arrest and admission to an intensive care unit (ICU) [3], [4]. However, its ubiquitous use in a hospital environment still remains a challenge mainly because of poor measurement accuracy. On the other hand, pulse oximetry (SpO2) measurements are ubiquitous despite its limited use as an early warning indicator as a decline in SpO2 below key physiological levels most often indicates severe patient distress. In this work, we develop an algorithm for early detection of respiratory distress through simultaneous analysis of RR and SpO2. Increased monitoring can reduce adverse events, improve outcomes, reduce length of stay, and minimize legal liability.

[1] are with the General Electric, Global Research, Bangalore, India `Hariharan.Ravishankar,Aditya.Saha at ge.com`

[2] is with Shell Oil Company, Bengaluru, Karnataka, India `swamygok at gmail.com`

[3] is with the General Electric, Global Research, Niskayuna, NY USA `gencs at ge.com`

The premise for our work comes from a seminal paper by Lynn and Curry [4] where the authors highlighted the shortcomings of the existing respiratory alarm monitoring practices. The authors argue that the traditional threshold breach method of detecting instability was not scientifically derived/proven and such a method is incapable of detecting instability early enough leading to a number of unexplained hospital deaths (UHD). The authors postulate that pathophysiological events leading to these UHDs are complex, multi-parametric and often progress along three distinct patterns designated as Type I, Type II and Type III. Type I pattern of unexpected hospital death (UHD) reflecting a clinically evolving process associated with microcirculatory failure induced by such common conditions as congestive heart failure (CHF), sepsis, and pulmonary embolism is shown in Fig. 1.

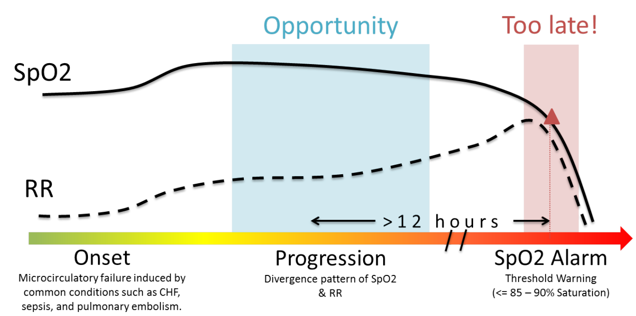

Fig. 1. Type I pattern of unexpected hospital death (UHD) reflecting a clinically evolving process associated with microcirculatory failure induced by such common conditions as CHF, sepsis, and pulmonary embolism to name a few.

In this paper, we propose a multi-parametric, real-time analysis framework for detecting patterns of respiratory distress described in [4]. We restrict our analysis to RR and SpO2 only because RR and SpO2 are two noninvasive measurements that can be acquired in various care areas including but not limited to ICUs and General Care Floor (GCF). We develop models that assess the changes in RR and SpO2 values over time and detect if any of the three patterns of UHD are present. We define temporal abstractions from raw RR and SpO2 measurements, identify useful multi-parametric events using them and finally detect temporal patterns of interest using a Markov model based framework. We have conducted experiments on detecting Type I patterns from a set of records from MIMIC II database [5] of patients who suffered from respiratory complications and suffered mortality.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

II. Technical Approach

We approach the problem of respiratory distress detection as a multi-parametric temporal pattern recognition problem as shown in Fig. 2. We start with the one-minute-summary RR and SpO2 values and define various levels of temporal abstractions (TA) before the Markov model based Finite State machine (FSM) is called into action. These TAs in-turn morph the problem to an event-sequence detection problem which can be readily solved by the Markov-model based framework.

A. Trending

Detection of clinically actionable decisions from noisy, turbulent data is challenging. It is necessary to pre-process the one-minute summary values to achieve consistent, reliable and relevant alarms. Towards that goal, our pre-processing step includes de-noising, outlier rejection, and trending. To achieve outlier rejection, we use the classical Chauvenet's criterion [6]. Since the dynamic range and the fluctuations in values are different for RR and SpO2, we use different statistical parameters in the Chauvenet's criterion for RR and SpO2 time-series measurements.

Trend analysis in time-series data has attained considerable attention from the research community over the years. Simpler techniques include parametric regression where an appropriate curve fitting function of suitable order is computed in a least-squares framework. One drawback of such an approach is that it is almost impossible to model the entire time-series by a single function which would lead to large amounts of representation error. We use the locally weighted scatter plot smoothing (LOWESS) technique and its generalization LOESS [7] to achieve robust long-term and short-term trends. LOWESS method combines multiple regression models in a k-nearest neighbor-based beta model. One of the major advantages of this method is that it does not require the specification of global fitting function but only the *smoothing parameter* and the degree of the local polynomial.

B. Segment Classification and Event Detection

In this section, we describe two important temporal abstractions that enable us in detecting a multi-parametric temporal pattern of interest: 1) *Segment Labels* - stream level abstraction, describing the variation in RR and SpO2 independently, and 2) *events* - models the temporal relationships between multiple streams by assimilating the stream level labels.

Segmenting time series data into segments of similar characteristics is a well-studied problem in temporal data mining [8]. Methodologies towards efficient representation of temporal data has become an effective and a necessary step in many a task ranging from compression, clustering, classification and rule-mining of time series data. Some of the high-level representations include symbolic mappings, SAX and transform-based while the most frequently used technique has been piecewise linear representation (PLR) [9], [10], [11]. Intuitively PLR refers to approximating a time-series data of length n into K linear segments leading to an efficient representation for storage, transmission and more importantly for computational mining tasks. The authors of [8] discuss variants of PLR approach including an online variant which makes the representation readily suitable to real-time data.

An alternate representation paradigm involves transforming time-series data into an interval based representation. The streams are divided into frames and features and/or labels are extracted from these intervals of constant length. While the techniques described earlier are 'content based' and represent similar portions using a single label, the interval-based representation produces a label for every frame and similar portions might span over multiple segments. The frame-based representation has been extensively employed in classic temporal pattern recognition problems like speech recognition [12], handwriting recognition, etc. The biggest advantage from such a representation is that it results in a series of feature-vectors with specific time-stamps which can then be handled by Hidden Markov models [12] or dynamic time warping techniques [13] to perform the matching or detection tasks. There have been works in the medical domain too that use this interval-based representation [14].

In our current implementation, we use frame-based representation.A PLR technique would represent the RR and SpO2 trends by an independent set of linear approximations which would have varying start/end time and length depending on the changes on the individual streams. To ascertain an event of clinical significance (e.g., simultaneous increase in RR and decrease in SpO2), we have to take into account the labels as well as the varying time-stamps of these segments which can be tricky. Second, the parameters of the PLR such as the length and number of segments, have to be managed intelligently to exploit the full potential of the "content based" representation. Finally, a frame-based representation can be easily integrated into a Markov model based framework. Additional constraints on the time properties of the pattern can be incorporated into the learning framework using this frame length as a quantum of time unit.

C. Markov models - Finite state machines

The problem of respiratory distress detection can now be considered as an event-sequence detection problem. We have a string of events coming from time frames, and the task is to raise an alarm if a sequence of events corresponding to Type I pattern is present. There are multiple techniques that have been explored for sequence detection in time-series mining and also in areas including communications, speech recognition amongst many others. Some of the techniques include regular expression based [**?**], Finite State machines based (FSM) , Hidden Markov Models (HMM) [12], etc. These techniques differ in complexity as well as the statistical flexibility in modeling they can offer.

Markov models have been extremely popular in medical decision making [15]. As medical decisions are often sequential and uncertain, these models are an appropriate

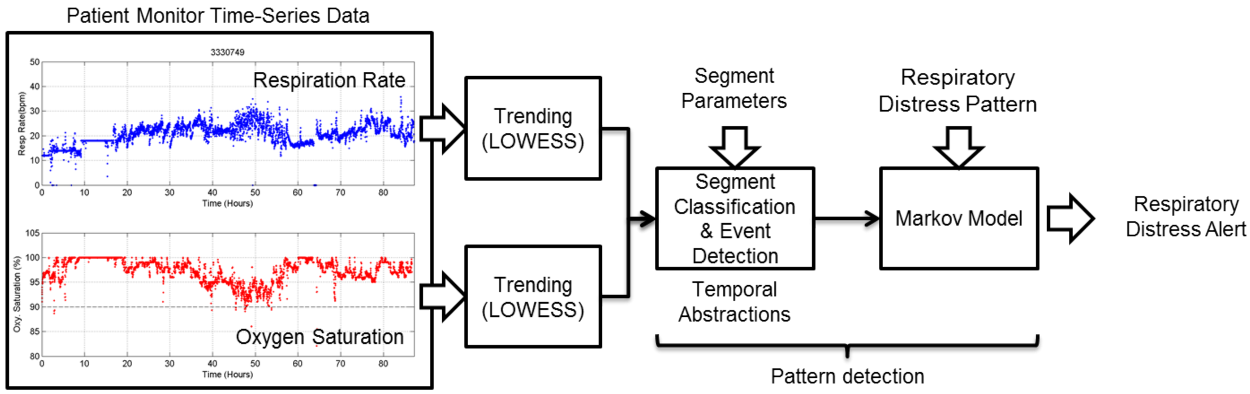

Fig. 2. Markov Model based framework for respiratory distress detection.

technique in solving for stochastic and dynamic changes. Markov models assume that the patient is always in one of a finite number of discrete health states. The events observed over time prompt the transition from one state to another and finally an important clinical condition can be inferred once the terminal state of this Markov chain is reached. In addition, the flexibility of modeling repetitive events by simply resetting the state machine to an initial state make them ideal in solving problems in health care like alarm detection.

III. Results on MIMIC II

As mentioned before, we have used records from the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) II database to test 1) the hypothesis that the patterns highlighted in [4] are indeed observed in patients with respiratory distress, and test 2) the sensitivity and specificity of the respiratory pattern detection algorithm on two patient groups, one group composed of patients with respiratory distress and the other group with no sign of respiratory distress.

MIMIC II database is a publicly and freely available database containing records of a diverse and very large population of ICU patients. There are essentially two basic types of data in the MIMIC II database; patients' clinical data and bedside monitor waveforms along with associated derived parameters (numerics). The clinical database contains patients' laboratory results, admission and death records, discharge summaries, ICD-9 codes, and nurse-verified down-sampled trends. The waveform database contains signals recorded by the bedside monitors such as electrocardiograms (ECG) and arterial blood pressure (ABP) waveforms. The derived parameters from these waveforms such as heart rate and systolic blood pressure are contained in the numerics database. In our study, we utilize ICD9 codes in the clinical database and one-minute summary values in the numerics database.

In order to form the one patient group with respiratory distress and one control group with no respiratory distress, we first utilized the ICD9 codes and then manually annotated the matching numerical records for types of patterns described in [4]. We now describe the process of forming these two groups in detail. The MIMIC II clinical database as of 2013 contained 26654 and matching numerical database contained 2431 records. We downloaded 1117 records (46%) randomly out of 2431 for preliminary analysis. 1096 out of 1117 contained non-empty ICD9 files. There are six groups of ICD9 codes associated with respiratory condition only. 696 out of 1096 patients have one or more of these groups of respiratory conditions. 404 out of 696 (or 35% of 1096 patients we started with) contained ICD9 code associated with acute respiratory failure.

404 clinical records resulted in 807 numerical records because one patient may have multiple recordings because they were readmitted and their identification numbers were matched accordingly. 234 out 807 records were down-selected for further analysis because they contained long enough measurements of RR and SpO2 needed per discussion in [4]. After manual inspection of 234 records, we noted that the Type I pattern shown in Fig. 1 was the most common pattern. 186 and 11 out of 234 time-series numerical records contained Type I and III patterns, respectively. 37 records were eliminated from the initial analysis due to very noisy respiration rate measurements or lack of variation in respiratory rate, potentially due to mechanical ventilation.

696 clinical records with no respiratory distress resulted in 904 numerical records. In order to form balanced patient groups, we randomly picked 157 records with long enough measurements of RR and SpO2. After manual inspection of all the records, we observed that 31 records out of 157 had other respiratory ICD9 codes and contained Type I patterns. We eliminated 32 records due to noisy respiration data. 94 our of 157 records contained no Type I patterns. 48 out of 94 records contained no respiratory condition which we used for our preliminary analysis. Later, we tested our algorithm on the entire 157 records as well.

For our preliminary analysis, we wished to have balanced patient groups. Thus, we randomly selected 50 patient records out of 186 to match 48 patient records remaining which contained no respiratory related ICD9 codes. The results of the performance testing on the preliminary data of two groups are shown in Table I. Using our current implementation, we have achieved a True Positive Rate (TPR) of 92% and False Positive Rate (FPR) of 6%. Later, we extended the performance testing to the entire 157 records with no acute respiratory distress but may contain other respiratory conditions in one ore more of the six ICD9 respiratory disease groups which resulted in TPR of 88% and FPR of 12%.

TABLE I
THE PERFORMANCE OF THE MULTI-PARAMETER RESPIRATORY PATTERN DETECTION ALGORITHM.

Actual versus predicted	Positive - Actual	Negative - Actual
Positive	46	3
Negative	4	45
Total	50	48

The alarms produced by our algorithm provide significant lead times ranging from hours to days as shown in Fig. 3. The highlighted that in majority of these records (around 80%), threshold breaches were not met (resulting in longer lead times depending on the length of the recording) or were met relatively late, leaving very little time to save the patient from dying. In other words, existing systems would not alarm in these cases whereas our algorithm is able to detect respiratory instability.

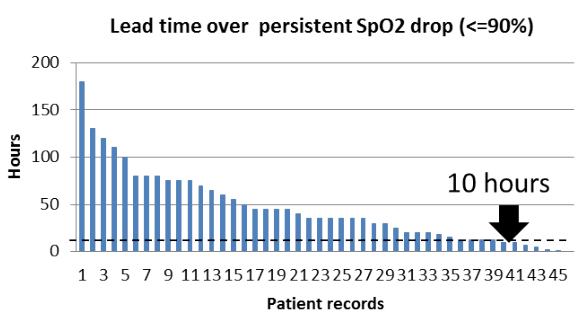

Fig. 3. The alarms produced by our algorithm provide significant lead times ranging from hours to days compared to current threshold logic based on SpO2 measurements.

In order to understand the contribution of RR and SpO2 trends resulting in respiratory distress pattern, we counted the number of times the Markov state model transitioned with one or more of the parameters. The results are shown in Table II in percentages. According to this table, RR plays a significant role in the onset of pattern while SpO2 plays a significant role leading to alarm.

TABLE II
THE SIGNIFICANCE OF RR AND SPO2 MEASUREMENTS IN TRIGGERING MARKOV STATE CHANGES LEADING TO TYPE I PATTERN DETECTION.

State Transitions	Both	RR only	SpO2 only
From Start to Onset		57%	43%
From Onset to Alarm	55%	25%	20%
From Reaction to Alarm	30%	15%	55%

IV. CONCLUSIONS

The need for relevant, robust and dependable alarm systems in patient care has been identified as one of the key challenges to be addressed in our current healthcare climate. Alarm fatigue and subsequent patient mortality due to the existing threshold-based alarm systems have been well-documented in the recent past. In this paper, we proposed algorithms towards smarter alarm systems pertaining to respiratory distress. We perform a multi-parametric analysis of patients' physiological parameters over time and ascertain the potential to develop respiratory distress. It has been highlighted in [4] that respiratory distress patterns evolve as a condition over time and current alarm systems are not equipped to deal with them. We have proposed a framework to analyse these patterns and have achieved promising results on 100 records from MIMIC II database for detecting Type I patterns. Our future research includes efforts on optimizing the system parameters for sufficient lead times and alarm detection rate using inputs from clinical experts. We plan to expand this framework to detect other patterns of interest and also extend this effort to other clinical conditions.

REFERENCES

[1] M. Peberdy, W. Kaye, J. Ornato, G. Larkin, V. Nadkarni, M. Mancini, R. Berg, G. Nichol, and T. Lane-Trultt, "Cardiopulmonary resuscitation of adults in the hospital: a report of 14720 cardiac arrests from the national registry of cardiopulmonary resuscitation," *Resuscitation*, vol. 58, no. 3, p. 297308, 2003.

[2] (2010, November) American heart association joins hospital quality programs to improve cardiac and respiratory outcomes. American Heart Association. [Online]. Available: http://newsroom.heart.org/news/1168

[3] C. M.A., B. R., H. K., J. Chen, S. Finfer, and A. Flabouris, "Respiratory rate: the neglected vital sign," *Med J. Aust.*, vol. 188, no. 11, pp. 657–9, 2008.

[4] C. J. Lynn L.A., "Patterns of unexpected in-hospital deaths: a root cause analysis," *Patient Safety in Surgery*, vol. 5, no. 3, 2011.

[5] R. A. e. a. Saeed M., Villarroel M., "Multiparameter intelligent monitoring in intensive care ii (mimic-ii): a public-access intensive care unit database," *Crit. Care Med.*, vol. 39, no. 5, pp. 952–60, 2011.

[6] W. Chauvenet, *A Manual of Spherical and Practical Astronomy*, 5th ed. Dover, N.Y., 1960, vol. V, no. II.

[7] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.

[8] E. Keogh, S. Chu, D. Hart, and M. Pazzani, *Data Mining in Time Series Databases*. World Scientific, Singapore, 2004, ch. Segmenting time series: a survey and novel approach.

[9] R. Agrawal, K. I. Lin, H. S. Sawhney, and K. Shim, "Fast similarity search in the presence of noise, scaling, and translation in times-series databases." in *21th International Conference on Very Large Data Bases*, 1995, pp. 490–50.

[10] R. Agrawal, G. Psaila, E. L. Wimmers, and M. Zait, "Querying shapes of histories." in *21st International Conference on Very Large Databases*, 1995.

[11] K. Chan and W. Fu, "Efficient time series matching by wavelets," in *15th IEEE International Conference on Data Engineering*, 1999.

[12] L. R. Rabiner, "Tutorial on hidden markov model and selected applications in speech recognition," in *IEEE XXX*, vol. 77, no. 2, 1989, pp. 257–285.

[13] D. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *AAAI-94 workshop on knowledge discovery in databases*, 1994, pp. 229–248.

[14] A. R. Post and J. H. Harrison, "Temporal data mining," *Clinics in Laboratory Medicine*, vol. 28, no. 1, pp. 83–100, 2008.

[15] F. Sonnenberg and J. Beck, "Markov models in medical decision making: a practical guide." *Med. Decis. Making*, vol. 13, pp. 322–338, 1993.