Application of TPRMine method for Identification of Temporal Changes on Patients with COPD: A Case Study in Telehealth

Catherine Inibhunu, Member IEEE
Ontario Tech University
Oshawa, Ontario, Canada

Carolyn McGregor AM, Senior Member IEEE
Ontario Tech University
Oshawa, Ontario, Canada
University of Technology Sydney
Ultimo, NSW, Australia

***Abstract*—Monitoring the vital status of an aging population especially those with chronic diseases can potentially reduce the multiple emergency room visits and hospitalizations if patients and care providers are provided with information that might help them make informed decisions on appropriate cause of actions. This process can be enabled by utilizing temporal abstraction and deriving temporal patterns in order to understand the underlying temporal relationships on vital status in data collected from patients participating in telehealth programs in combination with other data sets in order to get a complete patient flow. Such discovery can highlight when an elderly patient is at risk of an adverse event and this is information that can be utilized for provision of appropriate care for the patient.**

This paper demonstrates application of a method for deriving temporal patterns from patient's physiological data thereby quantifying the many states a patient can transition to before, during and after an adverse event. With this approach, it is possible to quantify patients with vital scores based on their physiology.

Keywords—Temporal Mining, Telehealth, Tele-Medicine, Temporal Analysis, Patient Scoring, Pattern Recognition

I. INTRODUCTION

The costs of lengthy hospital admissions and multiple emergency room visits (ER Visits) for patients with conditions such as heart failure (HF) and chronic obstructive pulmonary disease (COPD) can place a significant burden on healthcare systems Ward et al., [1]. A report generated from National Ambulatory Care Reporting System (NACRS) in Canada shows that in 2014-15 the leading medical conditions for which patients were admitted to the emergency departments (EDs) were COPD, HF and Pneumonia [2]. In 2012, it was noted in [3] that the direct costs associated with caring for patients with such conditions was estimated at over $50 billon in USA. In addition, Gorst et al. indicated that escalation of COPD and HF has a direct link with an aging population. In Canada, it is estimated that by 2036, 1 in 4 Canadians will be a senior aged 65 and over and 85% of that population will have some type of chronic condition [1]. To this respect, there is a need for an effective management of chronic conditions like COPD or HF which can be facilitated through remote patient monitoring services in telehealth [3]. Remote Patient Monitoring (RPM) is such an approach that is considered to have potential to improve the quality of life and continuity of care for patients diagnosed with cardiac conditions as well as lead to reduction in healthcare costs through a decrease of avoidable emergency room visits and hospital admissions [3]. Such a process can greatly benefit from effective collection and analysis of physiological data captured from elderly patients especially those with chronic conditions such as the HF and COPD. In particular, remote collection and analysis of patient's data could aid in effective decision making as a proactive and earlier warning of the need for intervention. As noted in our earlier works resulting from research on predictive analytics utilizing RPM data in [4] and [5], use of RPM in provision of care for patients was recognized by several healthcare organizations to have great potential.

Our research utilized predictive modeling to determine key factors that are significant determinants of hospitalization and multiple ER Visits [4]. The results showed that gender, past medical history and vital status are key factors to hospital admissions and ER Visits. Men were more likely to have hospital admissions compared to women and that the probability of presence of past medical history was statistically significant for seniors aged 65 and over. Additional models that included a flag to indicate the period before during or after an adverse event showed SpO_2 and pulse rate as significant predictors of an adverse event [5]. However, it was not clear how the time flag impacted the predictive model and further analysis was required to understand the patterns before, during and after an adverse event. As such, more research is needed to fully understand the temporal patterns among variables before, during and after a hospital admission or an ER visit.

This paper facilitates this process by investigating the use of temporal abstraction and pattern recognition models to fully quantify the temporal patterns exhibited in patients' data before, during and after an adverse event. We demonstrate application of a temporal pattern recognition and mining algorithm for discovery of temporal patterns on data captured from elderly patients who participated in an ethics approved RPM pilot project described in [4] [5]. Details on this application are provided in section II while results are presented in section III and the conclusion is presented in Section IV.

II. METHODS

A. Overview of TPRMine

As part of this research a mathematical formulation of data streams was completed by utilizing the state model developed and detailed in [6]. Using this approach, a problem space is modelled as a finite state machine representation whereby for a given timeframe, a time series data segment transitions from one state to another based on some associated weights. Representation of time series data with weights is a form of dimension reduction strategy and in Markov chains this weight is normally based on some probability distribution as noted in [6]. For a given time segment the adoption of hidden markov modes (HMMs) generates the sets of probable states that a time series dataset can transition to and the inferred paths of such transitions. States themselves are unique and can be inferred from the data or predefined per data domain [7].

When processing data captured from elderly patients who have been diagnosed with COPD or HF, there are several questions that may arise: what are the most frequent transitions and paths leading to this transition in any period of time and, with multiple time series data sets, are there any hidden relationships among these state transitions?

To answer these questions, this paper adopts the Temporal Pattern Recognition and Mining (TPRMine) algorithm detailed in [8] to process data captured from elderly patients.

TPRMine enables a scaled approach to processing high frequency time series data streams with predefined data windows. As patient physiological data streams are generated approx. every second, processing of temporal windows over some length of time i.e. hours can derive hierarchies of information which would be instrumental to building a knowledge base about the underlying patients.

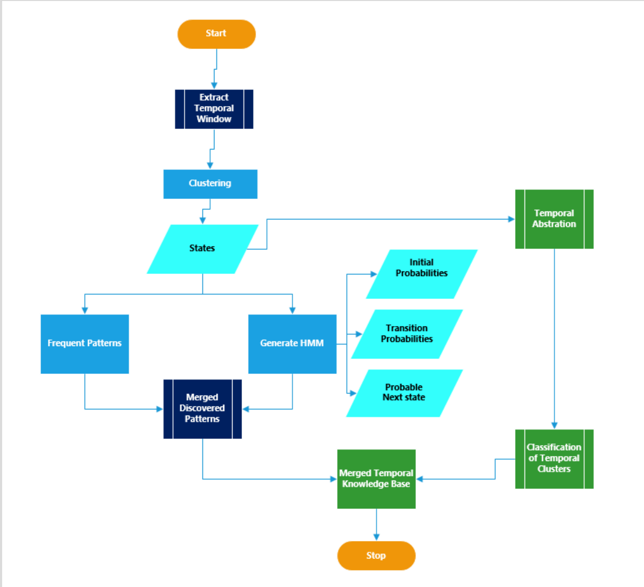
This research premise is that understanding the various factors contributing to lengthy hospitalization and multiple ER visits could aid in cost-effective management in the delivery of services leading to potential improvement on quality of life for elderly patients.

Utilizing the TPRMine algorithm is accomplished in a step wise manner as presented in Figure 1 where output from each step is provided as input to subsequent steps.

In the first step, temporal data windows were formulated by dividing the input data streams into unique groups such that each group is termed as a unique problem space with its own characteristic. This paper utilized preprocessed data comprised of averages of a patient weight, pulse, SpO_2, minimum and maximum blood pressure. Each of these data sets were captured remotely and due to the limitation of the data capture process, the data is sporadic and infrequent.

With TPRMine, 3 temporal data windows are derived based on 3 different events identified from this data. Day before, day of and day after an ER visit. Each of these three data sets are then supplied as input to the subsequent steps. The next step utilizes the data driven clustering detailed in [8] in order to derive the number of temporal states that a patient may transition to in each of the 3 data sets.

Figure 1: TPRMine Process Flow



The resulting states are then fed to HMM for determining the probability of transitioning from one state to the next and predicting the next probable state. Additionally, the clusters and the cluster centroids form the transaction sets for input to frequent pattern mining and association rules generation. Moreover, the clusters and cluster centroids form the information needed for deriving patient scores by integration of clinical context for abstracting physiological data as described next.

B. Temporal Abstraction with Clinical Context

The second step of the TPRMine algorithm generated data driven clusters. A quantification of the variability of those clusters was completed by using a clinical abstraction score thereby determining if a cluster is normal or abnormal. As clusters are the derived states, subsequently, this process identifies normal and abnormal states.

This process is completed by adopting the temporal abstraction criteria detailed in Baker, et al., [9] where by a physiological feature such as SpO_2 is categorized into 3 groups; normal, abnormal and danger. Baker et al. (2012) recommended this abstraction for use on data captured from patients aged 16 and over [9].

Within TPRMine, a ranking mechanism is adopted to quantify the risk of SpO_2 with three colors Red, Green and Yellow, each representing the abstracted group. In particular, normal is scored at risk 0 and quantified with green color, abnormal is represented in yellow and a risk score 1 and danger is represented with red which indicates risk score 2.

Using this classification, there are 3 derived risk scores that can be associated with a state. A similar scoring mechanism was detailed in McGregor et al. where researchers utilized 4 different data streams and assigned 4 levels based on some quantified threshold to each raw data streams separately [10].

In this paper, a score is assigned based on a combination of the elements in a cluster (state) and not to raw data streams.

As a state can contain multiple data streams, the scoring per state can simultaneously incorporate all the underlying context in the data. Note that this is a scalable approach where different rules can be integrated based on the physiological data streams utilized. Risk scores form the vital scores assigned to a patient such that if a patient happens to transition to a specific state, then the patient vital score is taken as the risk score in that particular state.

III. RESULTS

The three temporal windows are provided as input to TPRMine thereby generating data driven clusters associated with the data processed. The results are presented in Table 1 whereby there are a total of 9 states (s1-s9) generated from a day before and Day of an ER visit, while a day post ER has 8 states (s1-s8). The cluster means of each of the derived states per data set is also presented in Figure 2.

Figure 2: Data Driven Clusters Derived from 3 Temporal Periods Before During and After an Adverse Event

Cluster Means 1: 1 Day Before ER

	s1	s2	s3	s4	s5	s6	s7	s8	s9
ave_weight	87.4344	62.4564	70.637	103.917	93.8623	68.9085	61.7465	36.8663	91.2987
ave_pulse	85.51	92.7001	68.4551	70.2015	80.9251	70.174	85.9908	88.0365	69.0001
ave_spo2	92.4195	95.5889	93.2124	91.0428	93.6488	95.7322	92.6914	88.2651	94.9999
ave_bpmin	79.6296	76.3638	67.6896	75.6716	73.0444	66.139	72.0283	66.3006	138.997
ave_bpmax	130.558	128.985	126.873	139.664	135.754	107.75	135.111	108.419	70.0013

Cluster Means 2: Day of ER

	s1	s2	s3	s4	s5	s6	s7	s8	s9
ave_weight	61.6821	98.1506	41.3843	76.672	62.7554	93.8425	60.5143	81.5521	68.3202
ave_pulse	96.0065	67.8959	85.9093	76.2542	80.7403	80.4798	86.7387	86.8592	79.6403
ave_spo2	95.565	91.1241	89.2445	91.8775	94.7998	93.6293	92.8774	90.8169	92.5007
ave_bpmin	63.1167	81.5134	72.6984	77.7116	73.7128	72.997	67.6891	81.0877	73.3683
ave_bpmax	133.289	149.528	102.49	134.404	123.473	135.442	120.767	131.105	126.206

Cluster Means 3: 1 Day Post ER

	s1	s2	s3	s4	s5	s6	s7	s8
ave_weight	48.1715	64.8232	91.2758	93.6392	59.3537	99.5019	69.2012	65.2493
ave_pulse	91.0383	68.0116	69	78.3895	82.6387	78.562	92.4015	84.4968
ave_spo2	92.8198	90.9741	95.0013	93.502	91.8174	93.1963	94.0373	93.3892
ave_bpmin	64.8779	74.8744	138.946	75.4529	108.168	72.538	79.933	68.8962
ave_bpmax	120.138	133.325	70.0114	134.718	151.263	139.285	122.887	130.468

Next, the probability of transitioning from one state to the next in the sequence of before, during and after a hospital admission or an ER visit was determined. Figure 3 shows 3 periods of time before, during and after a visit to ER and the associated transition matrix.

After the abstraction, quantification of the states identified for each of the periods, before, during and after an ER Visit was performed. The results are shown in Figure 3 where TPRMine identifies that there are 9 states that a patient can transition to during the period before an ER visit. Using the temporal abstraction applied to the SpO2 data, states that are normal, abnormal or danger are easily identified. In the period 1, day before an ER Visit there are 3 normal states (green), 5 abnormal states (yellow) and 1 danger state (red) as presented in Figure 4.

Figure 3: Transition Matrices on 3 Periods Before, During and After an Adverse Event

Transition Matrix 1: 1 Day Before ER

	s1	s2	s3	s4	s5	s6	s7	s8	s9
s1	94.5%	0.7%	0.0%	0.3%	0.3%	0.7%	1.7%	1.7%	0.0%
s2	0.5%	93.9%	0.3%	0.5%	2.1%	1.1%	0.5%	0.8%	0.3%
s3	0.5%	0.5%	94.4%	0.0%	0.9%	0.9%	0.5%	0.9%	1.4%
s4	0.7%	0.7%	0.5%	95.8%	0.5%	0.2%	0.7%	0.2%	0.7%
s5	0.3%	2.0%	0.3%	0.7%	92.3%	0.7%	2.7%	0.7%	0.3%
s6	0.3%	1.2%	0.3%	1.2%	0.3%	94.1%	1.2%	0.6%	0.6%
s7	0.9%	0.2%	0.7%	0.9%	1.1%	0.9%	93.0%	0.9%	1.4%
s8	1.1%	1.4%	0.5%	0.0%	0.3%	0.5%	1.1%	94.3%	0.8%
s9	0.0%	0.5%	0.5%	1.1%	0.8%	0.5%	1.1%	0.5%	94.8%

Transition Matrix 2: Day of ER

	s1	s2	s3	s4	s5	s6	s7	s8	s9
s1	94.9%	0.0%	0.3%	0.9%	1.6%	1.3%	0.3%	0.6%	0.0%
s2	1.1%	93.6%	0.4%	0.4%	0.0%	1.4%	1.1%	1.1%	1.1%
s3	0.2%	0.6%	95.1%	0.4%	0.9%	0.4%	0.2%	0.9%	1.3%
s4	0.7%	0.3%	1.0%	92.2%	2.0%	1.0%	1.4%	0.7%	0.7%
s5	0.5%	0.3%	2.5%	1.4%	94.2%	0.5%	0.3%	0.0%	0.3%
s6	0.3%	1.6%	0.6%	0.3%	0.0%	93.9%	1.3%	0.3%	1.6%
s7	0.7%	0.0%	0.4%	2.6%	0.7%	0.4%	93.7%	0.0%	1.5%
s8	0.5%	1.0%	1.0%	0.5%	0.3%	0.3%	0.5%	95.9%	0.0%
s9	0.7%	0.9%	0.7%	0.5%	0.7%	0.2%	0.2%	0.9%	95.0%

Transition Matrix 3: Day of ER: 1 Day After ER

	s1	s2	s3	s4	s5	s6	s7	s8
s1	93.5%	1.3%	1.7%	0.9%	0.4%	0.4%	0.9%	0.9%
s2	0.4%	96.4%	0.0%	0.0%	0.7%	0.7%	0.4%	1.4%
s3	1.5%	0.7%	96.0%	0.4%	0.7%	0.4%	0.4%	0.0%
s4	2.2%	1.1%	0.0%	91.4%	1.1%	4.3%	0.0%	0.0%
s5	2.0%	0.0%	1.3%	1.3%	91.9%	1.3%	1.3%	0.7%
s6	0.4%	0.4%	1.2%	0.4%	1.2%	94.5%	0.8%	1.2%
s7	0.7%	0.7%	0.7%	0.7%	0.7%	2.0%	94.6%	0.0%
s8	2.4%	1.2%	0.6%	0.6%	0.6%	0.6%	0.0%	94.0%

Figure 4: Temporal Abstraction on Derived States

Cluster Means 1: 1 Day Before ER

	s1	s2	s3	s4	s5	s6	s7	s8	s9
ave_weight	87.4344	62.4564	70.637	103.917	93.8623	68.9085	61.7465	36.8663	91.2987
ave_pulse	85.51	92.7001	68.4551	70.2015	80.9251	70.174	85.9908	88.0365	69.0001
ave_spo2	92.4195	95.5889	93.2124	91.0428	93.6488	95.7322	92.6914	88.2651	94.9999
ave_bpmin	79.6296	76.3638	67.6896	75.6716	73.0444	66.139	72.0283	66.3006	138.997
ave_bpmax	130.558	128.985	126.873	139.664	135.754	107.75	135.111	108.419	70.0013

Cluster Means 2: Day of ER

	s1	s2	s3	s4	s5	s6	s7	s8	s9
ave_weight	61.6821	98.1506	41.3843	76.672	62.7554	93.8425	60.5143	81.5521	68.3202
ave_pulse	96.0065	67.8959	85.9093	76.2542	80.7403	80.4798	86.7387	86.8592	79.6403
ave_spo2	95.565	91.1241	89.2445	91.8775	94.7998	93.6293	92.8774	90.8169	92.5007
ave_bpmin	63.1167	81.5134	72.6984	77.7116	73.7128	72.997	67.6891	81.0877	73.3683
ave_bpmax	133.289	149.528	102.49	134.404	123.473	135.442	120.767	131.105	126.206

Cluster Means 3: 1 Day Post ER

	s1	s2	s3	s4	s5	s6	s7	s8
ave_weight	48.1715	64.8232	91.2758	93.6392	59.3537	99.5019	69.2012	65.2493
ave_pulse	91.0383	68.0116	69	78.3895	82.6387	78.562	92.4015	84.4968
ave_spo2	92.8198	90.9741	95.0013	93.502	91.8174	93.1963	94.0373	93.3892
ave_bpmin	64.8779	74.8744	138.946	75.4529	108.168	72.538	79.933	68.8962
ave_bpmax	120.138	133.325	70.0114	134.718	151.263	139.285	122.887	130.468

To demonstrate the results of TPRMine algorithm on the actual patients within the three periods, 3 patients are randomly selected. Figure 5 shows the temporal transition of the three patients in the day before the ER visit. Figure 6 presents the same patients during the day when ER Visit happened and Figure 7 shows the same 3 patients the day after the occurrence of an adverse event. Within each of the unique time periods, frequent patterns were then generated forming a knowledge base for retrospective analysis.

Figure 5: State Transitions by 3 Patients on Day before ER Visit. The Y-Axis is states in the period, X-Axis is the unit time of each transition and the red line is the state characterized as a danger state.

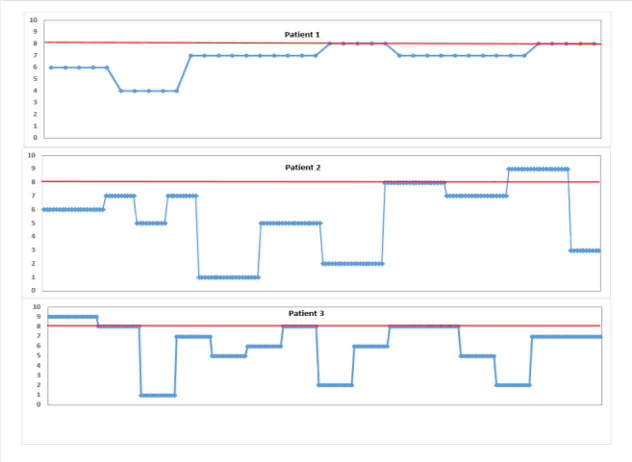


Figure 6: State Transitions by 3 Patients on Day of ER Visit.

The Y-Axis represents the states in the period and the X-Axis represents the unit time of each transition. The red line indicates the state that is characterized as a danger state.

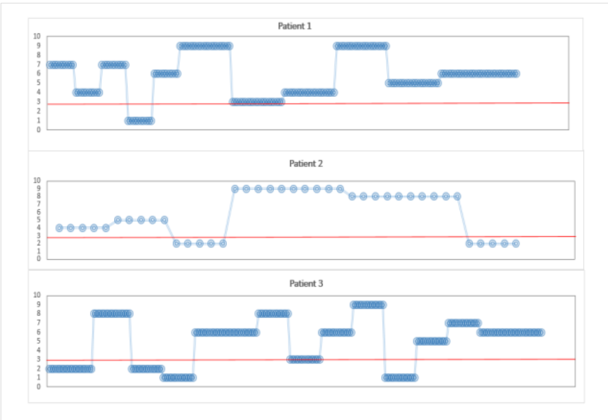
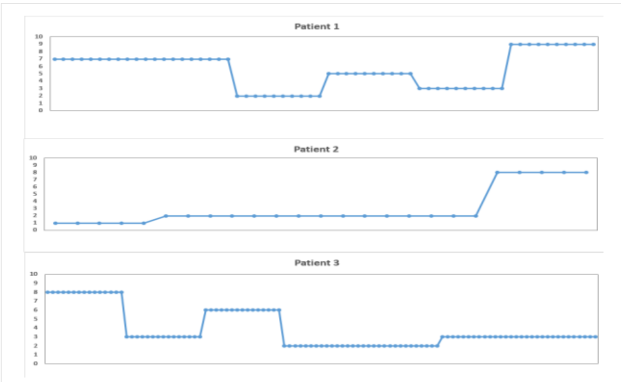


Figure 7: State Transitions by 3 Patients on Period Post ER Visit. The Y-Axis represents the states in the period and the X-Axis represents the unit time of each transition. The patients do not transition to a danger state in the Period Post ER Visit.



IV. CONCLUSION

This paper demonstrates that clinical application of TPRMine to data captured from elderly patients receiving care using a RPM program allows discovery of difference in patterns leading to and after an adverse event such as the visit to an ER. Temporal windows are formulated as input to TPRMine which then generates knowledge about 3 unique periods of time. The results demonstrate the potential to (a) quantify the number of states a patient diagnosed with COPD or HF can transition to before, during and after an adverse event, (b) identify the probability of transitioning from one state to the next and the many paths possible based on the HMM principles and (c) frequent patterns generated within each period of time can inform retrospective analysis and future knowledge discovery.

V. REFERENCES

[1] B. W. Ward, J. S. Schiller and R. A. Goodman, "Multiple chronic conditions among US adults: a 2012 update," *Prev Chronic Dis.,* vol. 11, no. doi: 10.5888/pcd11.130389., 2014.

[2] CIHI,2015.[Online].Available:https://secure.cihi.ca/free_products/NACRS_ED_QuickStats_Infosheet_2014-15_ENweb.pdf.

[3] D. Gotz, F. Wang and W. Perer, "A Methodology for Interactive Mining and Visual Analysis of Clinical Event Patterns Using Electronic Health Record Data," *Journal of Biomedical Informatics,* 2014.

[4] C. Inibhunu, A. Schauer, O. Redwood, P. Clifford and C. McGregor, "The impact of Gender, Medical History and Vital Status on Emergency Visits and Hospital Admissions: A Remote Patient Monitoring Case Study," in *1st IEEE Conference in Life Sciences*, Sidney, 2017.

[5] C. Inibhunu, A. Schauer, O. Redwood, P. Clifford and C. McGregor, "Predicting Hospital Admissions and Emergency Room Visits using Remote Home Monitoring Data," in *1st IEEE conference in Life Sciences*, Sidney, 2017.

[6] C. Inibhunu and C. McGregor, "Fusing Dimension Reduction and Classification for Mining Interesting Frequent Patterns in Patients Data," in *Machine Learning and Data Mining in Pattern Recognition. MLDM*, New York, 2018.

[7] C. Inibhunu and C. McGregor, "State Based Hidden Markov Models for Temporal Pattern Discovery in Critical Care," in *Life Sciences Conference (LSC)*, Montreal, Canada, 2018.

[8] C. Inibhunu, "A Temporal Pattern Discovery Method and its Application to Physiological Data: A Case Study," *Doctoral dissertation,* 2020.

[9] T. Baker, C. O. Schell, E. Lugazia, J. Blixt, M. Mulungu, M. Castegren, J. Eriksen and D. Konrad, "Vital Signs Directed Therapy: Improving Care in an Intensive Care Unit in a Low-Income Country," *PLoS ONE 10(12),* 10 2012.

[10] C. McGregor, A. James, M. Eklund, D. Sow, M. Eblinge and M. Blount, "OReal-time Multidimensional Temporal Analysis of Complex High Volume PhysiologicalData Streams in the Neonatal Intensive Care Unit," *Stud Health Technol Inform,* pp. 362-366, 2013.