Personalized Temporal Medical Alert System

Juan Pablo Suarez Coloma, Christine Verdier, Claudia Roncancio

Personalized Temporal Medical Alert System

Trend configuration and follow-up

Juan-Pablo Suarez-Coloma[1,2], Christine Verdier[1], Claudia Roncancio[1]
[1] Univ. Grenoble – Alpes, France
[2] Calystene S.A.
Firstname.Lastname@imag.fr

***Abstract*—The continuous increasing needs in telemedicine and healthcare, accentuate the need of well-adapted medical alert systems. Such alert systems may be used by a variety of patients and medical actors, and should allow monitoring a wide range of medical variables. This paper proposes Tempas, a personalized temporal alert system. It facilitates customized alert configuration by using linguistic trends. The trend detection algorithm is based on data normalization, time series segmentation, and segment classification. It improves state of the art by treating irregular and regular time series in an appropriate way, thanks to the introduction of an observation variable valid time. Alert detection is enriched with quality and applicability measures. They allow a personalized tuning of the system to help reducing false negatives and false positives alerts.**

Keywords—time series; trend; alert; fuzzy logic; quality metric; valid time; personalization

I. INTRODUCTION

Alert systems have been largely implemented in different domains as home, car, natural risks surveillance, or medical follow-up. In most cases, alerts are notified when a monitored variable value is out of a predefined range. In medical domain, alerts concern different users and a variety of interests. More automatic and popular alert systems concern drug prescription [1]. Alert systems detect drug interaction, contraindication, cross allergies, and other drug related events [2] [3] [4]. Such alert systems are often included in drugs databanks (in France, Claude Bernard's, Vidal's or Thériaque's). More expert alert systems are used in Intensive Care Units (ICUs) [5] [6]. These systems are based on the same model: vital parameters monitoring and alert notification when the measured value goes beyond predefined thresholds. Alerts are notified to physicians and nurses. Medical alert systems are mostly used in health professional environments (office or hospital).

Telemedicine tends to a largely and legal development for different and complementary reasons as reducing health costs and improving patients' quality of life. Transferring medical actions towards paramedical professionals or patients themselves become a reality. Context-aware medical alerts measuring trends becomes particularly important in the new private environment where there are no, or few, medical skills. Some already existing systems are dedicated to elderly surveillance (physiological or home sensors). These Alert systems are preconfigured with most often no actions except: on or off position. Systems are activated manually by elderly people themselves each time they feel in danger. A customer service is informed and decisions are taken.

This paper presents Tempas, a context-aware alert system based on linguistic trends. Tempas works with different kind of variables, and can be used in all environments (specialized or not) by every kind of users (expert or not). Trends are detected over regular and irregular time series extracted from patient observations. Linguistic trend values are stemmed from fuzzy logic. The *application index* measures the quality of the trend classification. It is used for alert filtering.

Tempas extends Pas [7], an alert system, based on fuzzy logic and linguistic values, connected to medical and environmental databases. The linguistic values are expressed in natural language such as "low", "normal", and "high". Tempas introduces time management in the system. Linguistic values are used to classify variable evolution such as "decreasing", "stable", and "increasing". Tempas has been implemented and integrated within an ERP solution called Futura, owned by Calystene S.A.

Section II presents a global view of Tempas and focuses on alert configuration. Trend definition and trend detection are explained in detail in section III and section IV, respectively. Section V presents related works on alert systems, time series, and trends. We present our conclusions and perspectives in section VI.

II. TEMPAS

In this section, we present an overview of Tempas and a brief summary of the alert configuration process. Tempas is not an expert system but a decision making help tool. Next sections explain how the user defines trends and how the algorithm detects alerts.

Tempas is a context-aware alert systems based on trend analysis and user personalization. The personalization allows medical actors to create their own alert system by defining:

- **The variables to monitor**, vital and non-vital parameters, environmental conditions, etc.
- **Specific valid value ranges** for the monitored variables. This information allows overriding default values if required
- **Relevant trends** leading to alerts during variable monitoring. For example monitoring increasing trends in body temperature to alert detection.
- **Target population**, patient or group of patients potentially concerned by the alert.

- **Alert notification parameters** as the users to notify and the notification method.
- **The expected alert quality.** This is provided by the *application index*, which expresses how much an alert concerns a patient.

An example of alert definition is the following: “Send an alert to all nurses if a decreasing body temperature trend is detected for patient Johnson”. The alert behavior is adapted to the whole context (user, patient, alert itself) e.g. use a range of [36°C – 42°C] for body temperature in all alerts, and [37°C – 40°C] in alerts created by Dr. Smith. Only the alerts with an *AI* superior to 0.8 are notified.

Personalization and context-awareness tend to reduce the false negatives and false positives, and to have always relevant alerts. Anyone can use Tempas. No high mathematics or informatics skills are required to create or edit trend alerts. The more the system is used, the highest its quality becomes.

III. TREND DEFINITION

We explain in this section how the user defines linguistic trends. Each linguistic value is stemmed from fuzzy logic. Each fuzzy set represents a linguistic value. Segments are classified in a trend from its slope (angle) and the fuzzy sets. In section IV we explain how to find segments in time series and how to proceed to alert detection.

Before be informed by alerts, users need to define how trends can be linguistically classified e.g. “decreasing”, “stable”, “increasing”. A range of angles is specified for each classification set. Tempas provides a fuzzy set generation algorithm to help the user to avoid the tedious task of trapezoidal sets definition. The algorithm requires two input values: the number of sets (the different linguistic values) and a classification tolerance *CT*. Generated fuzzy sets can be modified by users if needed. Fig. 1 shows the generated sets for different *CT* and five linguistic values.

The *Trapezoidal membership function* maps each element from the universe X to a value between 0 and 1. The membership degree (computed from trapezoidal membership function) covers the classification ambiguity. Fig. 2 shows the trapezoidal membership function that computes the *AI* of a segment. The segment *AI* is the same as the alert *AI*. Equation (1) computes the membership degree.

Segment classification is closely related to false positives and false negatives. Users may reduce false positives increasing the *AI* quality filtering, redefining the number of classification sets, or redefining the *CT*. In case of false negatives, users may decrease *AI*, redefine the number of trends, or redefine the *CT*.

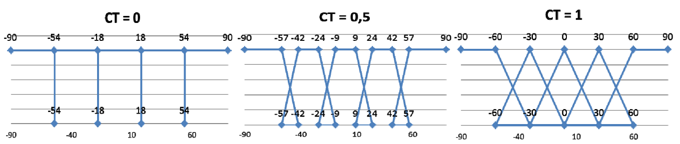


Fig. 1. Five generated fuzzy sets changing *CT*. Each set represents an ordered linguistic value among: “strongly decreasing”, “decreasing”, “stable”, “increasing”, “strongly increasing”.

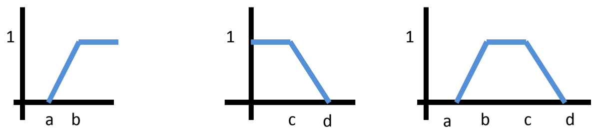


Fig. 2. Trapezoidal membership function

$$\mathrm{AI}(x)=\begin{cases}0, & x<a\\ \frac{x-a}{b-a}, & a\le x\le b\\ 1, & b\le x\le c\end{cases}\qquad \mathrm{AI}(x)=\begin{cases}0, & x>d\\ \frac{d-x}{d-c}, & c\le x\le d\\ 1, & x\le c\end{cases}\qquad \mathrm{AI}(x)=\begin{cases}0, & x<a \text{ or } x>d\\ \frac{x-a}{b-a}, & a\le x\le b\\ 1, & b\le x\le c\\ \frac{d-x}{d-c}, & c\le x\le d\end{cases} \tag{1}$$

IV. TREND DETECTION

Trend detection in Tempas consists in three steps explained in the next subsections. The normalization process transforms the observation information (measured value and timestamp) into a space between 0 and 1. Segmentation and fusion finds the most important k segments on regular and irregular time series. Finally, segment classification classifies each obtained segment as a linguistic trend.

1) Normalization

Data normalization translates data obtaining values between 0 and 1, allowing using segment slope as unique criteria for its classification. The algorithm uses segment angle, easy human eye understandable, for trend detection. The advantage is to avoid graphic perception problems. Graphically, the slope in Fig. 3 seems to be the same at left and right. Apparently, Unit changing does not affect the trend form. This happens when axis units are well chosen to draw both time series. In fact, the slope (defined as “vertical changing over horizontal changing”*)* changes if unit does. A temperature changing from 37.7 to 38.2 degrees Celsius in an 8 hours period can be presented as a changing from 99.86 to 100.76 degrees Fahrenheit in a 28800 seconds period. Slope is defined as 0.5/8 in the first case and 32.9/28800 in second. Slope is 0.1 in both cases if data are normalized. Tempas normalize data using variable and time ranges obtained from variable context and window length, respectively.

2) Segmentation and fusion

In this step, Tempas gets the observation set, segments the data, and merges until obtaining the *k* most significant segments. The *k* value is defined by the user. The segmentation algorithm uses a Piecewise Linear Representation [8] and data approximation by linear interpolation for two related reasons. It is easy to understand, thus, users can configure alerts by themselves.

The segmentation algorithm uses a bottom-up approach and a window based approach to deal with real time data streams [9]. A window corresponds to time series containing a

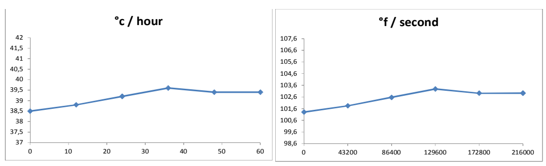


Fig. 3. Left graphic uses degrees Celsius for temperature and hours for time. Right one uses degrees Fahrenheit and seconds instead. Body temperature range is [37 °C - 42 °C] equivalent to [98.6 °F - 107.6 °F], time range is [0 h - 60 h] equivalent to [0 s - 216000 s].

patient observation set. The user specifies (1) the period length containing the desired observation set, and (2) the wished number of observations (*no*). These two parameters are used by the algorithm to create a hybrid window. A hybrid window has two measures, size and length. The window length is a temporal distance between the two extreme points of the observation set. The size is a quantitative measure defined by the number of points of the window. The window size and length do not always correspond to the number of observations and the period length specified by the user. A two-step algorithm uses the window to find the *k* most representative segments. Each segment is used to calculate one trend.

The first step is the segmentation. Tempas pass from *n* points to *n-1* or less segments. Tempas uses the observation variable valid time *OVVT* (how long the variable value is true for trend detection) [10]. OVVT is a segmentation parameter (see *Trendability*). This parameter let to handle regular and irregular time series differently. Segment fusion comes after time series segmentation.

Trendability is the ability that two consecutive points have to belong to the same trend. If the temporal distance between them is superior to the OVVT, they cannot belong to the same trend. Consequently, it is possible to find non-connected segments after segmentation. Fig. 4 shows two segmentations. One segmentation process (top-right) generates two non-connected segments. The second segmentation process (bottom-right) generates three connected segments.

An iterative algorithm finds the *k* most representative segments. Each cycle merges two connected segments into one. The decision is based on the smallest *merging cost*. *Merging cost* is a metric for merging two connected segments. It represents the Manhattan distance between the extreme points of two segments. If there is more than one smallest *merging cost*, the rightmost one is chosen. It is nearest the current time. The algorithm stops when the *k* segments have been found, or when the smallest *merging cost* goes over a user defined value. Fig. 5 shows the iterative fusion process until getting two segments.

3) Segment classification

Segment classification is the final step of trend detection. Final Segments are classified using the fuzzy sets generated by the user. Each segment classification returns a linguistic trend value and an *AI*. Detected trends with an *AI* over the threshold are notified to users e.g. a segment is classified as "stable" and "increasing" with a membership degree of 0.4 and 0.6, respectively. If threshold defined by the user is inferior or equal to 0.6, then, the system sends an increasing trend notification.

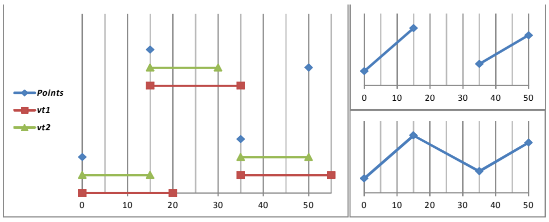


Fig. 4. At left, a 4 points time series with a valid time vt1 (square markers) and a valid time vt2 (triangle markers). At the top right, the segmentation uses vt1 and produces 2 non-connected segments. At the bottom right, the segmentation uses a valid time vt2 producing 3 connected segments.

V. Related Work

Medical alert systems are often preconfigured alerts. The main problem is that users lose interest in these types of alert systems because: they produce many false positives, false negatives, not well targeted alerts, or useless alerts [11][12][13]. For this reason, alert systems are not well accepted in clinical information systems [1].

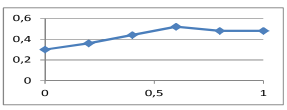
Time-series are multidisciplinary and produces huge amounts of data. Many works use dimensionality reduction based on signal treatment [14][15][16][17]. Charbonnier and Gentil [5] define three thresholds to detect a trend in a signal. Their values are chosen from normal behavior of monitored variables. Their algorithm uses these thresholds to classify online-detected segments as one of seven temporal shapes. Fuzzy logic on time-series helps to get human understandable results according to context situations [18] [19][20]. These works find only one trend on the time series from predefined fuzzy sets. In some cases, fuzzy sets are domain independent. Time warping techniques are used to compare time-series with different time length with the purpose of pattern matching [21]. Existing works are domain-dependent. To the best of our knowledge, no other works propose trend and variable information normalization

Most online segmentation algorithms use a sliding window approach. The algorithm tries to merge the new segment with the current segment. In case of positive merging, the current segment grows, else, the new segment becomes the current [5] [8]. Other approaches extend from two to three consecutive segments merging [22]. Time-series segmentation is achieved by bottom-up techniques [6]. SWAB is a generic algorithm for time-series segmentation [23]. SWAB mix a sliding window approach with a bottom-up approach for online segmentation. Systems express trends using rules over time-series segments e.g. "oxygen interval slope > 0.4 Kilopascal per second".

Papadimitriou, Sun et Faloutsos introduce SPIRIT [24]. SPIRIT monitors multiple streams at same time to found k

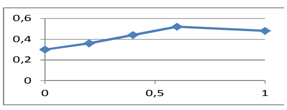
Time	Value
0	0,3
0,2	0,36
0,4	0,44
0,6	0,52
0,8	0,48
1	0,48

Segment	Cost
[0 - 0,4]	0,1796
[0,2 - 0,6]	0,1856
[0,4 - 0,8]	0,1616
[0,6 - 1]	0,1616



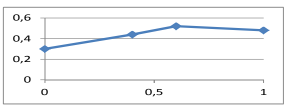
Time	Value
0	0,3
0,2	0,36
0,4	0,44
0,6	0,52
1	0,48

Segment	Cost
[0 - 0,4]	0,1796
[0,2 - 0,6]	0,1856
[0,4 - 1]	0,3616



Time	Value
0	0,3
0,4	0,44
0,6	0,52
1	0,48

Segment	Cost
[0 - 0,6]	0,4084
[0,4 - 1]	0,3616



Time	Value
0	0,3
0,4	0,44
1	0,48

Segment	Cost
[0 - 1]	1,0324

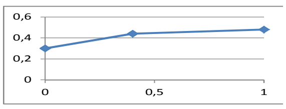


Fig. 5. Iteration process. A shadowed row contains the selected two consecutive segments to be merged. At the top left, the merging cost for the segments [0.4, 0.8] and [0.6, 1] is 0.1616 and is the smallest. The system chooses the rightmost one ([0.6, 1]).

hidden variables. Each hidden variable corresponds to a summarization of a group of correlated streams. Hidden variables are used to found (forecasting) trends using multiple variables. Forecasting trends are commonly used for anomaly detection [25] [26]. TrendX use trend templates to express expected behaviors (trend) of specific disorders [27]. Normal or abnormal behaviors are used for diagnostics. A trend template contains a temporal pattern in multiple variables.

VI. Conclusion and Perspectives

This paper presented Tempas, a temporal alert system allowing customized variable monitoring. It facilitates alert configuration by using linguistic trends. The system is generic – working with several kind of variables – and can be instantiated in any application domain. Tempas has been implemented in a real ERP solution. Preliminary experiments have been made. A non-expert user was able to create an alert to monitor rising body temperature. He adjusted the application index and the classification tolerance following his common sense. An expert user validated the results. More experiments will be realized in the near future. Automatic validation is not advised given that quality is a subjective value in alert systems.

This work introduced a quality index expressing how much an alert concerns a patient. Future work will improve quality information by introducing a trust index metric. This trust index will reflect how much the user can trust the alert. We will investigate the introduction of complex alerts combining trends and simple events (as medicine taking) and simultaneous trends.

Acknowledgment

Many thanks to Mr. Babouchkine and Mr. Dufour (Calystene) for sharing their experiences in health information systems. This work is supported by the ANR project Innoserv.

References

[1] S. Phansalkar, J. Edworthy, E. Hellier, D. L. Seger, A. Schedlbauer, A. J. Avery and D. W. Bates, "A review of human factors principles for the design and implementation of medication safety alerts in clinical information systems," in *Journal of the American Medical Informatics Association:JAMIA*, vol. 17, no. 5, 2010, pp. 493-501.

[2] J. S. Luo, "Electronic Prescribing Systems with Computer Decision Support," *Primary Psychiatry*, 2006, pp. 19-21.

[3] L. Taylor and R. Tamblyn, "Reasons for physician non-adherence to electronic drug alerts," in *Studies in health technology and informatics*, vol. 107, no. 2, 2004, pp. 1101-1105.

[4] P. Kilbridge, E. Welebob and D. Classen, "Development of the Leapfrog methodology for evaluating hospital implemented inpatient computerized physician order entry systems," in *Quality & safety in health care*, vol. 15, no. 2, 2006, pp. 81-84.

[5] S. Charbonnier and S. Gentil, "A trend-based alarm system to improve patient monitoring in intensive care units," in *Control Engineering Practice*, vol. 15, no. 9, 2007, pp. 1039-1050.

[6] J. Hunter and N. McIntosh, "Knowledge-Based Event Detection in Complex Time Series Data," in *Proceedings of the Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, 1999, pp. 271-280.

[7] W. Manzi de Arantes Junior and C. Verdier, "Defining quality-measurable medical alerts from incomplete data through fuzzy linguistic variables and modifiers," in *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 4, 2010, pp. 916-922.

[8] E. Keogh, S. Chu, D. Hart and M. Pazzani, "Segmenting Time Series: A Survey and Novel Approach," in *Data mining in Time Series Databases*, 1993, pp. 1-22.

[9] L. Petit, C. Labbe and C. Roncancio, "An algebric window model for data stream management," in *Ninth ACM International Workshop on Data Engineering for Wireless and Mobile Access*, Indianapolis, 2010

[10] C. S. Jensen, C. E. Dyreson, M. H. Bohlen, J. Clifford, R. Elmasri, S. K. Gadia, F. Grandi, P. J. Hayes, S. Jajodia, W. Kafer, N. Kline, N. A. Lorentzos, Y. G. Mitsopoulos, A. Montanari and D. Nonen, "The Consensus Glossary of Temporal Database Concepts - February," in *Temporal Databases*, Dagstuhl, 1997, pp. 367-405.

[11] J. Iskio, G. Uperman , B. Lumenfeld , E. Ecklet, D. Ates and T. Andhi, "Improving Acceptance of Computerized Prescribing Alerts in Ambulatory Care," in *Journal of the American Medical Informatics Association:JAMIA*, vol. 13, no. 1, 2006, pp. 5-11.

[12] S. Weingart , M. Toth , D. Sands and M. Aronson , "Decisions to override computerized drug alerts in primary care," in *Archives of Internal Medicine*, vol. 163, no. 21, 2003, pp. 2625-2631.

[13] H. van der Sijs, J. Aarts , A. Vulto and M. Berg, "Review Paper: Overriding of Drug Safety Alerts in Computerized Physician Order Entry," in *Journal of the American Medical Informatics Association:JAMIA*, vol. 13, no. 2, 2006, pp. 138-147.

[14] R. Agrawal, G. Psaila, E. L. Wimmers and M. Za, "Querying Shapes of Histories," in *VLDB'95, Proceedings of 21th International Conference on Very Large Data Bases*, Zurich, 1995.

[15] K.-p. Chan and A. W.-C. Fu, "Efficient Time Series Matching by Wavelets," in *Proceedings of the 15th International Conference on Data*, Sydney, 1999.

[16] R. Agrawal, C. Faloutsos and A. Swami, "Efficient similarity search in sequence databases," in *Proceedings of the 4th Conference on Foundations of Data Organization and Algorithms*, Chicago, 1993.

[17] K. Eamonn, K. Chakrabarti, M. Pazzani and S. Mehrotra, "Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases," in *Knowledge and Information Systems*, vol. 3, no. 3, 2001, pp. 263-286.

[18] C.-H. Chen , T.-P. Hong and V. Tseng, "Mining Linguistic Trends from Time Series," in *Data Mining: Foundations and Practice*, vol. 118, 2008, pp. 49-60.

[19] D. F. Sittig , K.-H. Cheung and L. Berman, "Fuzzy classification of heart rate trends and artifacts," in *Fifth Annual IEEE Symposium on Computer-Based Medical Systems*, Durham, 1992.

[20] A. Udechukwu, K. Barker and R. Alhajj, "Discovering all frequent trends in time series," in *Proceedings of the winter international synposium on Information and communication technologies*, 2004, pp. 1-6.

[21] M. Müller, "Dynamic Time Warping," in *Information Retrieval for Music and Motion*, 2007, pp. 69-84.

[22] A. alatian and J. Hunter, "Deriving Trends in Historical and Real-Time Continuously Sampled Medical Data," in *Journal of Intelligent Information Systems*, vol. 13, no. 1-2, 1999, pp. 47-71.

[23] E. . J. Keogh , S. Chu, D. Hart and M. J. Pazzani, "An Online Algorithm for Segmenting Time Series," in *Proceedings of the 2001 IEEE International Conference on Data Mining (ICDM 2001)*, San Jose, 2001.

[24] S. Papadimitriou, J. Sun and C. Faloutsos, "Streaming Pattern Discovery in Multiple Time-Series," in *Proceedings of the 31st International Conference on Very Large Data Bases*, Trondheim, Norway, 2005.

[25] K. Kalpakis, D. Gada and V. Puttagunta, "Distance Measures for Effective Clustering of ARIMA Time-Series," in *Proceedings of the 2001 IEEE International Conference on Data Mining (ICDM 2001)*, San Jose, California, USA, 2001.

[26] N. Liu, S. Nong, J. Yan, B. Zhang, Z. Chen and Y. Li, "Similarity of Temporal Query Logs Based on ARIMA Model," in *Proceedings of the 6th IEEE International Conference on Data Mining (ICDM 2006)*, Hong Kong, China, 2006.

[27] I. J. Haimowitz and I. S. Kohane, "Managing temporal worlds for medical trend diagnosis," in *Artificial Intelligence in Medicine*, vol. 8, no. 3, 1996, pp. 199-321.