Schreier G, Hayn D, Hörbst A, Ammenwerth E, editors. Proceedings of the eHealth2012. 2012 Mai 10-11; Vienna, Austria. OCG; 2012.

ALARM MANAGEMENT IN PATIENT HEALTH STATUS MONITORING

Vukovic M[1], Drobics M[1], Kreiner K[2], Hayn D[2], Schreier G[2]

Abstract

Based on the current research, false alarm generation is a problem with patient health monitoring. Occurrence of high rates of irrelevant alarms might lead to ineffective therapy. The existing approaches dealing with the alarms are statistical or artificial intelligence. The presented results compare fixed manually adjustable vs. dynamic automatically adjusted alarm thresholds, using the statistical approach on test data. The developed automated alarm management approaches are recommended in case of reduced medical staff effectiveness due to the large number of false alarms.

Keywords – Alarm management, Dynamic threshold adjustments, Patient health monitoring, Telemonitoring, Decision support systems

1. Introduction

Clinical alarms in health monitoring represent warnings to the caregivers of immediate or potential adverse patient status. However, sometimes the alarms may actually foster the occurrence of adverse events, which is the reason why ECRI institute lists alarm hazard as the top one health technology hazard for 2012 [1]. Problems with clinical alarms have existed since the advent of monitoring and therapy device use in healthcare and were first reported in the 1974 issue of Health Devices [2]. Patient monitoring alarm shortcomings have been the topic of numerous studies and analysis in the literature. Publications have shown the existence of limitations of current alarm systems [3]. The most reported negative side-effect is the large number of nuisance alarms. A paper on adverse events in low-risk patients with chest pain in emergency department, reported that 99.4% of the alarms were false, not resulting in a change of patient treatment management [4]. Another study on intensive care unit monitoring showed that over 90% of the alarms were false or clinically insignificant [5]. Our previous study resulted in similar findings: 87% of alarms were false [6]. The annoying alarms result from the lack of the systems reliability and accuracy and rarely from an adverse patient condition. Some of the consequences of false or nuisance alarms include interference with patient care resulting in reduced effectiveness of the nursing staff. The large number of false alarms demands substantial caregiver's time, patience, full attention, fast reactions and commitment which are not always easy to achieve. All such limitations and poor decision support could lead to dangerous undermining of the true alarms by the clinicians [7].

Review of the existing studies revealed a differentiation between statistical and artificial intelligence approaches [5]. The identified approaches were mostly used in intensive care

[1] Safety and Security Department, AIT Austrian Institute of Technology GmbH, Vienna
[2] Safety and Security Department, AIT Austrian Institute of Technology GmbH, Graz

monitoring (30 studies), less in general monitoring time series (16 studies), whereas telemonitoring alarms were investigated in only 5 studies primarily using statistical approaches of trend detection and curve fitting. Usual alarm generation upon exceeding fixed measurement thresholds results in numerous false alarms [5]. No standard exists for setting the default alarm thresholds for a particular monitored parameter [6]. Furthermore, there is no gold standard for alarm classification [7, 8]. The existing statistical techniques are limited by interpretability of the high-dimensional data [9], while the artificial intelligence approaches lack predictability needed for regulatory approvals [5].

In our previous publication, we proposed an optimization of the alarm-management system for telemonitoring of heart failure patients to help clinicians focus on the clinically significant data [6]. Observing the number of exceeded thresholds over consecutive days, the developed system provided a prioritization of alarms (emitted from different devices) based on the significance level (a hierarchy of the alarm importance).

Our present work focuses on examining alarm threshold features towards reduction of the false alarms and enhancement of the alarm-management system.

2. Methods

Data of the MOBITEL study, which was conducted in Austria from 2003 to 2008, were used for optimizing and validating our algorithm. A total of 9128 measurement records achieved by 65 patients were available, containing physiological parameters as well as lower and upper threshold values as set by the physicians. Interventions by the physicians were documented as well. As a response to the MOBITEL system alarms, the physicians indicated performing one of the following activities: (1) contact the patient, (2) adjust medications, (3) other action, (4) adjust alarm threshold, and (5) no action. The first three responses were considered to indicate true alarms (intervention necessary), whereas the last two indicated false alarms (no intervention necessary) together with the cases when no physician responses were recorded.

In the current study, we focused on the statistical data analysis methods and considered reducing the number of false alarms analyzing alternative procedures for setting up thresholds in comparison to the alarm generation based upon manually adjustable fixed thresholds. The dynamic thresholds we introduced were automatically adjusted for each new measurement symmetrically with respect to the current estimated reference state of patient conditions.

A “reference state” of each monitored parameter was estimated for each measurement day, based on smoothing of the measured data within an appropriate window size of consecutive measurement days preceding the current measurement day. Smoothing of the data was done using statistical measures of location, i.e. mean value (moving average), and Kalman filtering. We applied the proposed dynamic threshold adjustment methodology (reference state ± deviation limit) on the existing data from the MOBITEL study including four telemonitoring variables of chronic heart failure patients: Weight, Heart rate, Systolic and Diastolic blood pressure [10]. The difference in occurrence of true and false alarms was observed when dynamically adjusting the alarm thresholds. Various window sizes and deviation limits (absolute and relative values) were used to test the performance of the algorithm. Investigation of the possible dynamic threshold bounds was based on the calculated means and medians of the existing MOBITEL fixed thresholds. A comparison of the original method used in the MOBITEL study and our new approach is presented in *Figure 1*.

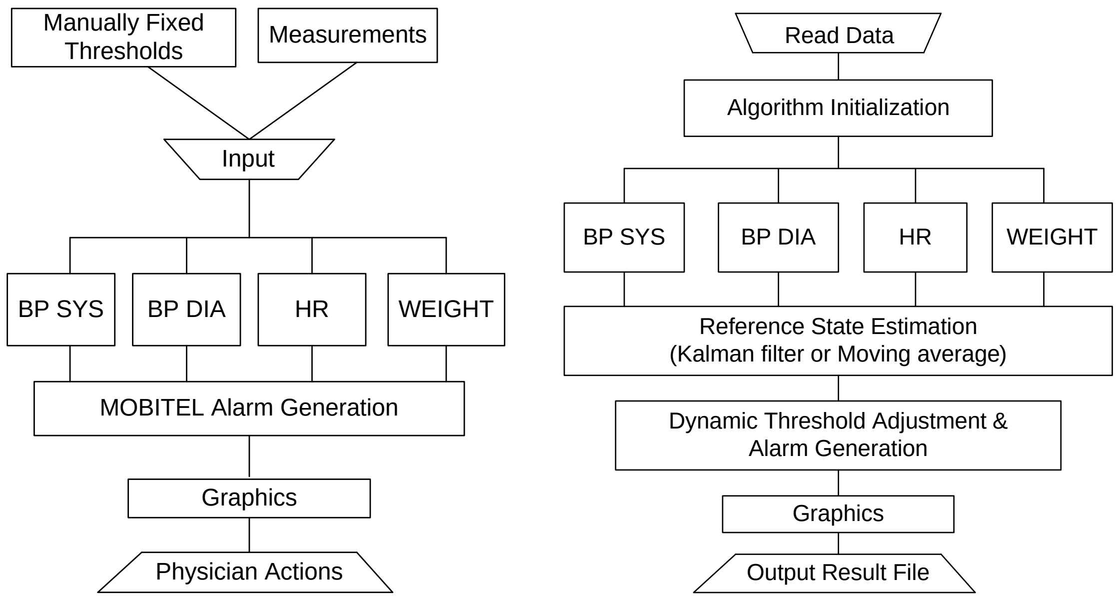

Figure 1: MOBITEL (left) and dynamic threshold adjustment (right) procedures

3. Results

Optimal values for the deviation limit (absolute or relative) and window sizes for reference state estimation are presented in *Table 1*. The presented bounds and window sizes were selected to achieve the largest agreement between the newly generated and MOBITEL true positive alarms. As a reference, fixed individual thresholds from the MOBITEL study are shown as well. Threshold ranges (e.g. lower systolic blood pressure 50-120 mmHg) correspond to minimum (50 mmHg) and maximum (120 mmHg) values of thresholds (lower systolic blood pressure), as specified by all physicians for all patients throughout the MOBITEL study.

Table 2 presents the comparison of the results from the original MOBITEL algorithm and from the analyses using the dynamic threshold adjustments on the same dataset. The presented results use moving average and Kalman filtering for patient reference state estimation. Deviation limits and window sizes were selected according to *Table 1*. The number of true and false alarms is shown for each corresponding physician action as well as the overall sensitivity and specificity.

Figure 2 presents the comparison between the results of the original MOBITEL and dynamically adjusted threshold algorithms for the selected patient weight data.

Table 1: Settings for the original MOBITEL and optimal values of the new statistical data processing algorithms using dynamic thresholds

	Fixed individual thresholds – MOBITEL	Dynamic thresholds	
		Deviation limit	Measurement window size [preceding days]
Systolic blood pressure [mmHg]	Lower 50-120 Upper 110-200	28	15
Diastolic blood pressure [mmHg]	Lower 35-200 Upper 80-210	20	14
Heart rate [bpm]	Lower 20-300 Upper 60-310	33 %	13
Weight [kg]	Lower 40-117 Upper 41-150	4	13

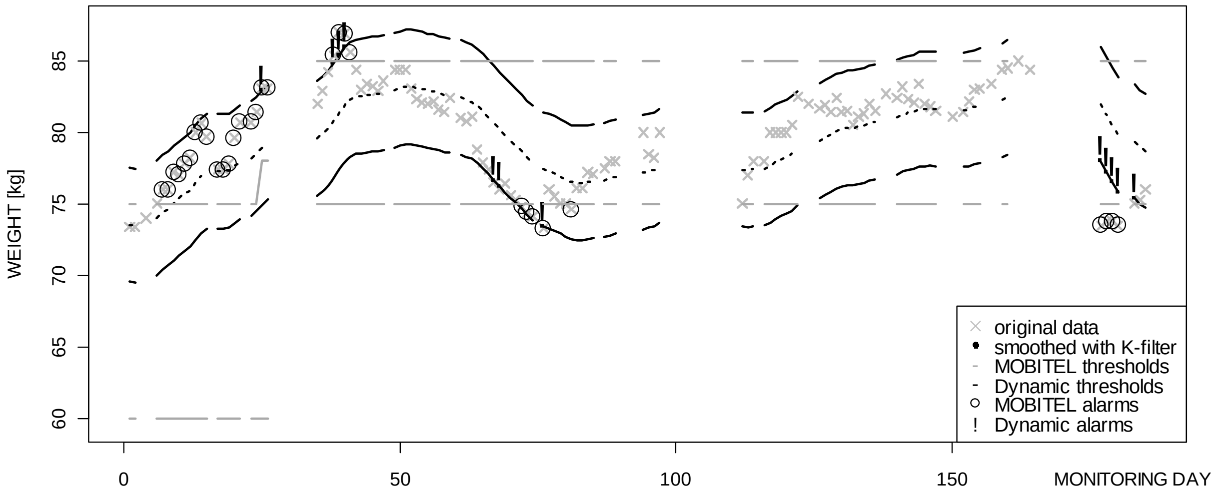

Figure 2: Sample comparison between MOBITEL and dynamic threshold adjustment results

The results presented in *Table 2* showed 91.5% and 84.7% reduction of false positive alarms for the two applied automated threshold adjustment algorithms based on moving average and Kalman filtering, respectively. However, such reductions of false positive alarms were followed by a reduction of 86% and 73.6% in true positive alarms, respectively. The portion of true positive alarms in the total number of alarms increased from 13% in MOBITEL to 20% using the dynamic threshold adjustment. False negative alarms also increased, decreasing the sensitivity to 0.131 and 0.248 in the cases of moving average and Kalman filtering, respectively.

Table 2: Results of the original data analysis and statistical data processing algorithms on MOBITEL data

	Physician responses / Type of alarms / Statistics	Fixed individual thresholds – MOBITEL	Dynamic thresholds	
			Moving average	Kalman filter
TRUE ALRMS	Patient contact	133	26	42
	Medication adjust	89	9	22
	Other action	36	1	4
	Total true alarms (TP)	258	36	68
FALSE ALARMS	No action	1233	82	162
	Treshold adjust	108	6	11
	No recorded response	387	59	92
	Total false alarms (FP)	1728	147	265
NO ALARMS	False negative (FN)	16	238	206
	True negative (TN)	6016	7597	7479
	Total no alarms	6032	7835	7685
	Specificity = TN/(TN+FP)	0.777	0.981	0.966
	Sensitivity = TP/(TP+FN)	0.942	0.131	0.248

4. Discussion

The settings of the original MOBITEL patient monitoring system presented in *Table 1* illustrate how unrealistic alarm thresholds might occur during the course of monitoring if medical practitioners are expected to perform manual adjustment. In some cases the setup was done in such a way that certain alarms were completely disabled, e.g. diastolic blood pressure lower limit of 35, upper limit of 210, or heart rate upper limit above 200. As exceeding such threshold limits would hardly occur, the number of alarms would be reduced, but also some possible adverse patient conditions could not be detected. With the proposed automated dynamic alarm threshold adjustment, all the measurements were compared against reasonable upper and lower threshold limits.

The results presented in *Table 2* illustrate the reduction of false positive alarms that can be achieved via statistical data processing approaches including dynamic threshold adjustment. However, the false alarm reduction comes with the price reflected in the reduction of true positive alarms. It is important to notice the number of unrecorded physician responses to the alarm situations. As the physicians had the possibility to review each alarm situation in MOBITEL, all the cases with no recorded physician responses were classified as false alarms. However, in the dynamic threshold adjustment approaches the physicians did not review any alarm cases, as the algorithms were not tested in practice but only on the recorded MOBITEL data. Therefore, it is questionable how many such alarms classified as false because of no recorded physician responses would have actually been false if the physicians had a chance to examine them. Most of such alarms were uniquely created by the dynamic threshold adjustment algorithms and differed from alarm occurrences in MOBITEL. Furthermore, out of the three physician actions classified as relevant: Patient contact, Medication adjust and Other action, the largest relative influence of the proposed dynamic threshold adjustment is on Other action. The reduction of "Other action" interventions was around 90% in both dynamic threshold adjustment approaches. It is not documented which "Other actions" the physicians took and how relevant such actions were for the patient condition. Due to such considerations the presented results are highly conservative in the estimates of true positive alarms and sensitivity. The net benefit of a particular statistical approach can be viewed via the false positive minus true positive alarm reduction difference, although true positive alarm reduction may not be acceptable for all the monitored cases. No approach was found to eliminate just the false positive alarms without affecting also the true positive alarm occurrences. Consequently, as the sum of true positive and false negative alarms remained constant the reduction in true positive alarms was followed by an increase in false negative occurrences and decrease in sensitivity. The effects of the reduced sensitivity on the patients might influence the stability of their health conditions if the medical staff would rely solely on the telemonitoring alarm generation. Therefore, the physicians would be expected to inspect and follow the patient conditions frequently with care when using the current dynamic threshold adjustment approaches. The use of the currently developed automated threshold adjustment algorithms is recommended in cases of reduced medical staff effectiveness due to the large number of false alarms in common telemonitoring approaches.

Limitations of the presented results are related to the unequal number of measurements between the participating patients effectively increasing the influence of the patients with more measurement days on the results. Furthermore, the measurements were often irregular, sometimes leaving large gaps within the data time series. In all such cases the measurement gaps were removed, as if the measurements were part of the continuous time series. Such limitations restricted the capabilities for developing greater reliability of the proposed automated data processing algorithms. Finally, the data analyses used all the collected measurements and did not differentiate between the individual patients in the sense of calculating personalized effectiveness of the developed algorithms for each patient separately. This is particularly significant as 30% of the patients in the referent study had no true positive alarms.

5. Conclusion

Large numbers of false alarms are a characteristic of current patient health status monitoring systems. Despite statistical and artificial intelligence approaches to reduce the number of false alarms, no gold standard for alarm classification exists. The explored possibilities to use dynamic alarm thresholds effectively reduce the false alarm occurrences at the expense of true alarm reductions and reducing sensitivity. Although the sensitivity estimates are conservative, as the developed algorithms are not tested in medical practice, usage of the proposed automated threshold adjustment may be considered as a complement to the common telemonitoring approaches in cases of reduced medical staff effectiveness due to the large numbers of false alarms. Future research will focus on increasing reliability and precision of the automatic threshold adjustment algorithms, as well as on early detecting adverse patient conditions by improving the alarm system intelligence.

6. References

[1] ECRI institute. Top 10 health technology hazards for 2012, Health devices, Nov 2011; 41(11)

[2] Emergency Care Research Institute ECRI. Hazard: Gaymar Hypothermia Machine, Health Devices, July 1974; 3(9): 229-230.

[3] ACCE healthcare technology foundation. Impact of clinical alarms on patient safety. A report from the American College of Clinical Engineering Healthcare Technology Foundation. Plymouth meeting PA Journal of Clinical Engineering Jan/Mar 2007.

[4] Atzema C, Schull MJ, Borgundvaag B, Slaughter GR, Lee CK. ALARMED: adverse events in low-risk chest pain receiving continuous electrocardiographic monitoring in the emergency department. A pilot study. Am J Emerg Med 2006 Jan; 24(1):62-7

[5] Imhoff M, Kuhls S, Alarm algorithms in critical care monitoring. Anest Analog 2006; 102:1525-37

[6] Vukovic M, Drobics M, Hayn D, Kastner P, Schreier G. Optimization of the alarm-management of a heart failure home-monitoring system. Computing in Cardiology 2010; 37:53-56

[7] Zhang Y. Real-time analysis of physiological data and development of alarm algorithms for patient monitoring in the intensive care unit. MEng thesis. Massachusetts Institute of Technology. Avg 2003; 29-31.

[8] Chambrin MC. Alarms in the intensive care unit: how can the number of false alarms be reduced? Critical Care 2001; 5:184-188.

[9] Gather U, Imhoff M, Fried R. Graphical models for multivariate time series from intensive care monitoring. Stat Med 2000; 21(18): 2685-2701.

[10] Scherr D, Kastner P, Kollmann A, Hallas A, Auer J, Krappinger H, Schuchlenz H, Stark G, Grander W, Jakl G, Schreier G, Fruhwald F. Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. J Med Internet Res 2009; 11(3):e34.

Corresponding Author
Marija Vukovic
Safety and Security Department, Austrian Institute of Technology GmbH
Donau City Strasse 1
A-1220 Vienna
Email: marija.vukovic @ ait.ac.at