33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

A New Density-Ratio based Approach for Patient-Specific Biomedical Monitoring

Dongmei Chen, *IEEE Student Member,* Max Q.-H. Meng, *IEEE Fellow*

Abstract— **In order to denote the abnormalities of the patients, we propose a novel approach to detect anomaly in biomedical monitoring using density ratio values as the Patient Status Index (PSI). The key idea of the proposed method is to define the ratio of training and testing data densities, where training dataset only consist of normal data and testing dataset consist of both normal and abnormal data, and identify irregular samples for testing patients' dataset. Furthermore, we define four inequalities to denote the interval values of density ratio and give the corresponding status for patients. In addition, the applied Kullback-Leibler based algorithm for calculating density ratio values without involving density estimation is equipped with a cross validation (CV) model selection procedure, allowing us to objectively optimize values of tuning parameters. We select training and testing data from Physionet database to do our pilot experiment. The experimental results for 11901 beats show that the density-ratio based approach work very well in terms of specificity and sensitivity.**

I. INTRODUCTION

MANY patients die in hospital or home every year because deterioration in patients' health status is not identified in advance. One of the most important actions taken to improve patient safety in hospitals or home is to identify patients who are deteriorating and act early [1]. Thus, there is a great need for patient monitoring systems that perform automatic identification of patient status. In order to identify these patients status changes, an approach based on novelty detection (anomaly detection, outlier detection, novelty detection, one-class classification) is proposed, in which a multivariate, multimodal model of the distribution of vital-sign data from normal patients [2] is constructed. Anomaly detection, which aims at detecting uncommon instances in given dataset or finding unusual patterns in time-series [3], [4], gathers a lot of attention these days. However, the anomaly detection problem is vague; it is impossible to universally define what the anomalies are.

Totally, there are three types of methods to deal with anomaly detection problem. The first type is Density Estimator method, such as Kernel Density Estimator (KDE) [5]. In [6] - [8], authors applied KDE-based method into biomedical monitoring to describe the status of the patients. However, $p^{tr}(x)$ is not accessible in practice and density estimation is known to be a difficult problem and Threshold is hard to be determined in the tail of training data distribution. Thus, using densities as anomaly score may not be promising in practice.

To avoid density estimation, One-class Support Vector Machine (OSVM) and Support Vector Data Description (SVDD) are introduced, which belong to the second type of method. Paper [9] shows the results applying OSVM into seizure analysis from intracranial EEG for epilepsy patients. However, the solutions of OSVM (SVDD) depend heavily on the choice of tuning parameters; choosing these parameters seems to be highly subjective. In addition, the training dataset need both normal and abnormal data, such anomalous data for training are not always available in practice and the type of anomalous may be diverse.

In order to overcome the above-mentioned weakness, [10] - [12] employ the ratio of training and testing data density as the scores to denote anomalies. As far as we know, there are no any research groups apply this method into biomedical monitoring system to denote the patients status. Here we regard instances with small density ratio values as anomalies (near to zero). Several methods such as Kernel Mean Matching (KMM) [10], [11], Kullback-Leibler Importance Estimation Procedure (KLIEP) [13] - [16] have been proposed recently to give density ratio without going through density estimation.

The organization of this paper is as follows. In section II, we propose the density-ratio based approach for patient-specific biomedical monitoring and convert this problem into density ratio estimation problem. In section III, we give the mathematical algorithm to show how to estimate the density ratio values. In section IV, we present some experiments results to illustrate the usefulness of our approach which is followed by the conclusions and future work in section V.

II. DENSITY-RATIO BASED APPROACH

In this section, we propose the density-ratio based approach for patient-specific biomedical monitoring and convert this problem into density ratio estimation problem.

A. Approach Description

In practice, normal samples are often available abundantly. Therefore, we can separate the monitoring data into a training set only consisting of normal samples observed in the past and the testing set consisting of recent samples from which we try to detect anomaly. Fig.1 shows a novel approach to detect abnormal event for patient in biomedical monitoring. The key idea of our approach is to use the ratio of training and testing data density as anomaly score to denote the patient status.

Manuscript received April 11, 2011. Dongmei Chen is currently a PHD student at Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong. (Phone: 00852-9584-7162; e-mail:dmchen@ee.cuhk.edu.hk).

Max Q.-H. Meng is a professor at Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T. Hong Kong. (Phone: 00852- 2609-8282; fax: 00852-2603-5558 ; e-mail: max@ee.cuhk.edu.hk).

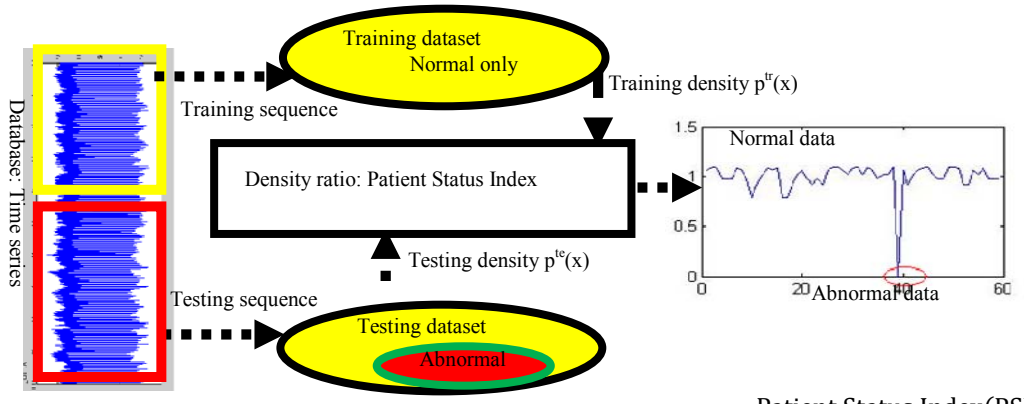

Fig.1 Approach to detect abnormal event for patient in biomedical monitoring

B. Formulation and Notation

Illumined by the density definition from [10], [13], here we can define our two sets of samples: training dataset $\{x_j^{tr}, x \in \mathbb{R}^d\}_{j=1}^{n_{tr}}$ and testing dataset $\{x_i^{te}, x \in \mathbb{R}^d\}_{i=1}^{n_{te}}$ in the d-dimensional domain ($D \in \mathbb{R}^d$). All samples in the training dataset $\{x_j^{tr}, x \in \mathbb{R}^d\}_{j=1}^{n_{tr}}$ are normal, while some anomalies are in the testing dataset $\{x_i^{te}, x \in \mathbb{R}^d\}_{i=1}^{n_{te}}$. We suppose training samples $\{x_j^{tr}, x \in \mathbb{R}^d\}_{j=1}^{n_{tr}}$ are independent and identically distributed (i.i.d.) following a training data distribution with nonnegative density- $p^{tr}(x)$, and testing samples $\{x_i^{te}, x \in \mathbb{R}^d\}_{i=1}^{n_{te}}$ are i.i.d. following a test data distribution with strictly positive density-$p^{te}(x)$. Via above two densities, the density ratio can be defined by [10], [13]:

$$w(x) = \frac{p^{tr}(x)}{p^{te}(x)} \quad (1)$$

In real situation, training dataset cannot consist of all types of normal data; the training probability density function can only represent part of normal distribution. So we propose the following (2):

$$\begin{cases} w(x) = \frac{p^{tr}(x)}{p^{te}(x)} > 1, p^{tr}(x) > p^{te}(x) & (a) \\ w(x) = \frac{p^{tr}(x)}{p^{te}(x)} \cong 1, p^{tr}(x) \cong p^{te}(x) & (b) \\ T < w(x) = \frac{p^{tr}(x)}{p^{te}(x)} < 1, p^{tr}(x) < p^{te}(x), T = \text{threshold} & (c) \\ w(x) = \frac{p^{tr}(x)}{p^{te}(x)} < T, ==> \frac{p^{tr}(x)}{p^{te}(x)} \ll 1, p^{tr}(x) \ll p^{te}(x) & (d) \end{cases} \quad (2)$$

From above inequality we can know that the ratio values tend to be small (near to be zero) in the regions where the test data density is high and the training data density is low.

C. Patient Status Index

The motivation of designing an automated biomedical monitoring system is to give the indication of patient status. In order to quantify the status of patients, a Patient Status Index is defined to generate alerts during periods of abnormal physiology:

$$\text{Patient Status Index(PSI)} = w(x) = \frac{p^{tr}(x)}{p^{te}(x)} \quad (3)$$

Alerts are generated when the PSI is below the threshold of density ratio values. In real situation, the density ratio w(x) is usually unknown, so the key issue of our approach is how to accurately estimate w(x). In the following part, we will show how to estimate density ratio values.

III. DENSITY RATIO ESTIMATION ALGORITHM

In this section, we give the mathematical algorithm to show how to estimate the density ratio values.

Illuminated by Kullback-Leibler Importance Estimation Procedure (KLIEP) [10] - [13] and [16], we can model the density ratio w(x) by the following linear model [10], [13]:

$$\tilde{w}(x) = \sum_{l=1}^{b} \alpha_l \varphi_l(x) \quad (4)$$

Where $\{\alpha_l\}_{l=1}^{b}$ are parameters to be learned from data samples and $\{\varphi_l(x)\}_{l=1}^{b}$ are basis functions such that $\varphi_l(x) \geq 0$ for all $x \in \mathbb{R}^d$ and for $l = 1,2,\cdots,b$. Using the model-$\tilde{w}(x)$, we can estimate the training data density $p^{tr}(x)$ by

$$\widetilde{p^{tr}}(x) = \tilde{w}(x)\, p^{te}(x) \quad (5)$$

The parameters $\{\alpha_l\}_{l=1}^{b}$ are determined so that the Kullback-Leibler divergence from $p^{tr}(x)$ to $\widetilde{p^{tr}}(x)$ can be minimized [10], [13]:

$$\min_{\{\alpha_l\}_{l=1}^{b}} KL[p^{tr}(x)||\widetilde{p^{tr}}(x)] = \min_{\{\alpha_l\}_{l=1}^{b}} \int p^{tr}(x) log \frac{p^{tr}(x)}{\widetilde{p^{tr}}(x)} dx$$

$$= \min_{\{\alpha_l\}_{l=1}^{b}} \int p^{tr}(x) log \frac{p^{tr}(x)}{\tilde{w}(x) p^{te}(x)} dx$$

$$= \min_{\{\alpha_l\}_{l=1}^{b}} \left\{ \int p^{tr}(x) log \frac{p^{tr}(x)}{p^{te}(x)} dx - \int p^{tr}(x) log \tilde{w}(x)\, dx \right\} \quad (6)$$

The first term in (6) is constant, so we only focus on the second term-inconstant term. Since $\widetilde{p^{tr}}(x) = \tilde{w}(x)\, p^{te}(x)$ is probability densities function, it should satisfy [10] - [13]:

$$\int \widetilde{p^{tr}}(x) dx = \int \tilde{w}(x) p^{te}(x) dx$$

$$= \frac{1}{n_{te}} \sum_{l=1}^{b} \alpha_l \left(\sum_{i=1}^{n_{te}} \varphi_l(x_i^{te}) \right) = 1 \quad (7)$$

Considering the second inconstant term and the restriction (7), the original problem can be converted into the following optimization problem:

$$\max_{\{\alpha_l\}_{l=1}^{b}} \left[\sum_{j=1}^{n_{tr}} \log\left(\sum_{l=1}^{b} \alpha_l \varphi_l\left(x_j^{tr}\right)\right)\right] \quad (8)$$

$$\text{s.t. } \frac{1}{n_{te}} \sum_{l=1}^{b} \alpha_l \left(\sum_{i=1}^{n_{te}} \varphi_l\left(x_i^{te}\right)\right) = 1, \alpha_1, \cdots, \alpha_b \geq 0$$

This is a convex optimization problem and the global solution can be obtained by performing gradient ascent and feasibility satisfaction iteratively.

The performance of above algorithm depends on the choice of basic functions $\{\varphi_l(x)\}_{l=1}^{b}$. In KLIEP, a good model may be chosen by Likelihood Cross Validation (LCV) [10], [13], and which is the procedure corresponding to choosing the model with the minimal $KL[p^{tr}(x)||\widetilde{p^{tr}}(x)]$.

IV. EXPERIMENTS EVALUATION

In this section, we do pilot experiment and report the results of anomalies for selected patients from MIT-BIH Arrhythmia Database.

The training and testing data employed to verify the availability and efficacy of the proposed diagnosis system are selected from MIT-BIH Arrhythmia Database [17], which is available at Physionet [18]. Here, we have one criterion to select patient records in our experiment. The record, where the duration of continuous normal data is longer than 6.30 minutes, can be selected as our database. So there are totally 10 records from 10 patients are satisfied.

The features are extracted from the sequence of records then converted into data samples in the training dataset and testing dataset. Here, we extracted 10 features from ECG signal. The features lists are shown in Table I. The numbers of involved heart beats are listed in table II. Clearly, it is a large scale experiment, containing totally 11,901 beats.

All the experiments processes are performed. Due to page limitation, here, we randomly choose 4 patients to show the results for anomalies detection. Fig.2 – Fig.5 are abnormal points' detection results for patient 101, 113, 115, 121. Among them, there are no any anomalies occurring during monitoring process for records 115. While there are some abnormal beats in records 101, 113, and 121. In the following figures, we use blue circle symbols to denote normal samples and red star symbols to denote abnormal samples.

TABLE I
DESCRIPTION FOR EXTRATED FEATURES

Feature symbol	Feature description
RR	The time duration between the adjacent beat R peak
QRS-dur	The time duration between *Q* and *S* in a QRS complex
TeSend	The time duration between *S* end and *T* end
H-QR	The amplitude between *Q* and *R* in a QRS complex
H-RS	The amplitude between *R* and *S* in a QRS complex
RP	The time duration between P and R
TR	The time duration between T and R
QsPstart	The time duration between Q start and P start
H-PR	The amplitude between *R*and *P* in the same beat
H-RT	The amplitude between *R* and *T* in the same beat

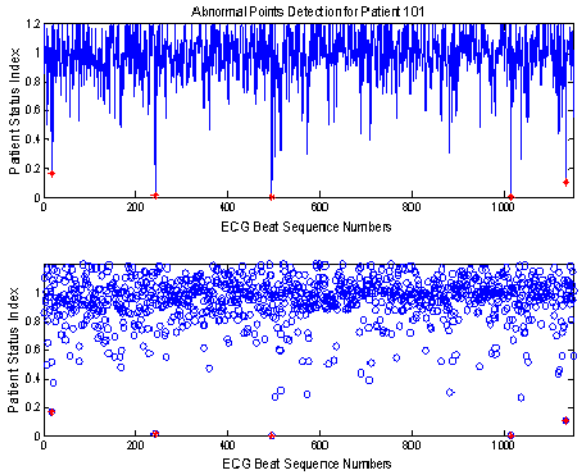

Fig.2 Patient Status Index for patient 101

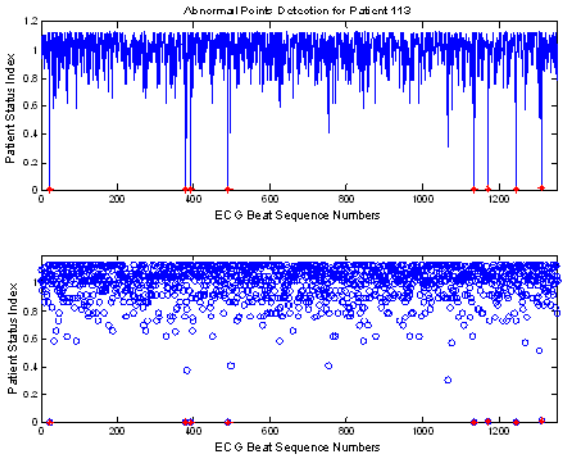

Fig.3 Patient Status Index for patient 113

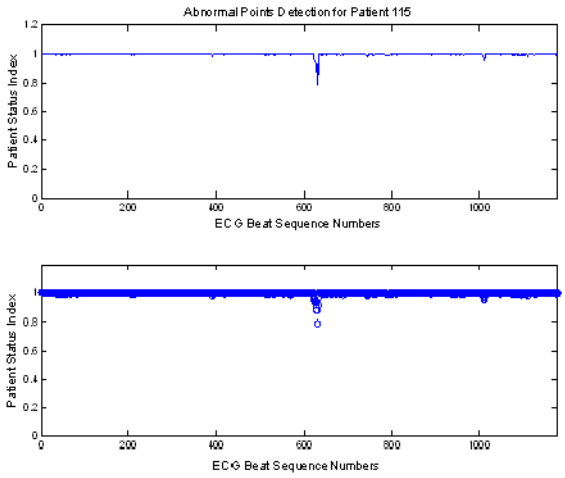

Fig.4 Patient Status Index for patient 115

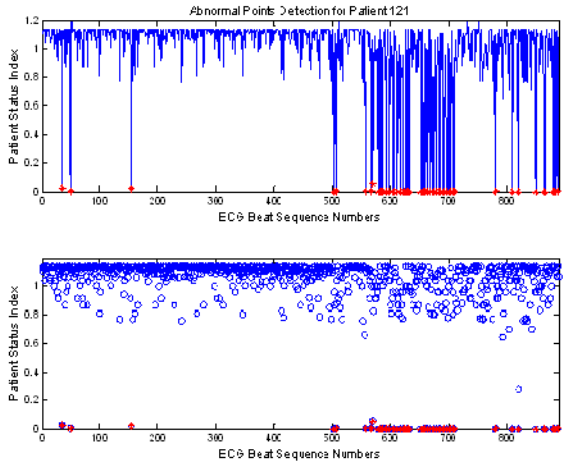

Fig.5 Patient Status Index for patient 121

TABLE II
SENSITIVITY AND SPECIFICITY FOR TEN PATIENTS
(Threshold=0.2)

Patient Record	Test Beat Numbers	Sensitivity (Se)	Specificity (Sp)
101	1153	**100%**	**100%**
103	1302	**100%**	99.77%
112	1597	**100%**	98.50%
113	1354	**100%**	99.85%
115	1179	**100%**	**100%**
117	912	**100%**	**100%**
121	890	**100%**	95.90%
122	1065	**100%**	**100%**
123	1165	**100%**	**100%**
234	1284	90.57%	99.62%
Total	**11901**	**99.06%**	**99.37%**

The sensitivity (Se) and specificity (Sp) for the ten patients are shown in table II when selecting 0.2 as the Threshold.

The experimental results for the ten patients datasets showed that our approach work very well in terms of specificity and sensitivity. From the table II we can see Patient record 234, exhibited poor performance, this is because there are many noises in this record, and those noises can be regarded as anomaly when performing our approach.

V. CONCLUSIONS AND FUTURE WORK

In this paper we propose a novel approach to the problem of abnormal event detection in biomedical monitoring using density ratio as the Patient Status Index (PSI) to denote the anomaly of the patients. KLIEP algorithm performing with a model selection procedure is more optimal and promising. So we use the method KLIEP to estimate density ratio parameter without involving directly density estimation. Via extensive experiment with some patients, the proposed approach can be demonstrated to be efficient.

In the future we will further apply this approach in longer time records and more patients. In addition, we try to find more effective algorithm to estimate density ratio values and apply it in disease detection for biomedical monitoring. In practice, undetected abnormal event may exist in the training set. In order to simulate realistic patient situation, we will add a small fraction of 'before-disease' samples to the training set. In the future, we will consider this realistic situation and detect anomaly using our proposed approach.

REFERENCES

[1] R. Thompson, et al., "Safer Care for the Acutely Ill patient: learning from serious incidents, fifth report from the Patient Safety Observatory", National Patient Safety Agency (NPSA), August 2007.
[2] Clifton, D.A., Hugueny, S., and Tarassenko, L.Novelty Detection with Multivariate Extreme Value Theory,Part I:A Numerical Approach to Multimodal Estimation IEEE Machine Learning in Signal Processing, Grenoble, France, 2009, pp. 1-6.
[3] Markos Markou, Sameer Singh. Novelty detection: a review—part 1: statistical approaches. Signal Processing Volume 83, Issue 12, December 2003, pp. 2481-2497.
[4] Markos Markou, Sameer Singh.Novelty detection: a review—part 2: neural network based approaches Signal Processing, Volume 83, Issue 12, December 2003, Pages 2499-2521
[5] R.O.Duda,P.E.Hart,D.G.Stork,Pattern Classication,Wiley, NY, USA, 2001.
[6] Hugueny, S., Clifton, D.A., and Tarassenko, L.Probabilistic Patient Monitoring Using Extreme Value Theory IEEE Biomedical Engineering Systems and Technologies, Valencia, Spain, 2010, pp. 5-12.
[7] Clifton, D.A., Clifton, L., and Tarassenko, L. Patient-Specific Biomedical Condition Monitoring for Post-operative Cancer Patients IET Condition Monitoring, Dublin, Ireland, 2009, pp. 424-433
[8] Tarassenko L, Hann A, Young D. Integrated monitoring and analysis for early warning of patient deterioration. British Journal of Anaesthesia. 97(2): pp.64-68, 2006.
[9] Gardner AB, Krieger A, Vachtsevanos G, Litt B. One-Class Novelty Detection for Seizure Analysis from Intracranial EEG. Journal of Machine Learning Research 7, pp:1025-1024, 2006.
[10] Hido, S., Tsuboi, Y., Kashima, H., Sugiyama, M., & Kanamori, T. Statistical outlier detection using direct density ratio estimation. To appear in Knowledge and Information Systems, 2010. http://sugiyama-www.cs.titech.ac.jp/~sugi/publications.html
[11] Sugiyama, M., Kanamori, T., Suzuki, T., Hido, S., Sese, J., Takeuchi, I., & Wang, L. A density-ratio framework for statistical data processing. IPSJ Transactions on Computer Vision and Applications, vol.1, pp.183-208, 2009.
[12] Hido, S., Tsuboi, Y., Kashima, H., Sugiyama, M., & Kanamori, T. Inlier-based outlier detection via direct density ratio estimation. Proceedings of IEEE International Conference on Data Mining (ICDM2008), pp.223-232, Pisa, Italy, Dec. 15-19, 2008.
[13] Sugiyama, M., Suzuki, T., Nakajima, S., Kashima, H., von Bünau, P. & Kawanabe, M. Direct importance estimation for covariate shift adaptation. Annals of the Institute of Statistical Mathematics, vol.60, no.4, pp.699-746, 2008.
[14] Sugiyama, M., Nakajima, S., Kashima, H., von Bünau, P. & Kawanabe, M. Direct importance estimation with model selection and its application to covariate shift adaptation Advances in Neural Information Processing Systems 20, pp.1433-1440, Cambridge, MA, MIT Press, 2008.
[15] http://sugiyama.www.cs.titech.ac.jp/~sugi/software/KLIEP/index.html
[16] Kanamori, T., Hido, S., & Sugiyama, M. Efficient direct density ratio estimation for non-stationarity adaptation and outlier detection, Advances in Neural Information Processing Systems 21, pp.809-816, Cambridge, MA, MIT Press, 2009.
[17] http://www.physionet.org/physiobank/database/mitdb/
[18] http://www.physionet.org/