Automating Hospital ICU Emergency Signaling

by

Fahmida Zahan	10301012
Nusrat Mahmud	10301015
Amanna Nawshin	10310003
Ferdousi Ahmed	10310010

Departments of Electrical and Electronic Engineering (EEE) &
Computer Science and Engineering (CSE)

Programs - Electronics and Communication Engineering (ECE) &
Computer Science and Engineering (CSE)

BRAC University, Dhaka, Bangladesh

April 2015

DECLARATION

We declare that this thesis and the work presented in it are our own and have been generated by us as the result of our own original research.

____

Fahmida Zahan

____

Nusrat Mahmud

____

Amanna Nawshin

_______________ _______________

Dr. Tarem Ahmed Ferdousi Ahmed

Signature of Advisor **Signature of Students**

DEDICATION

Every challenging work needs self efforts as well as guidance of elders. This dissertation is dedicated to our beloved families for their great support and continuous care, and to all faculties of EEE and CSE departments.

We also appreciate each other's effort for surviving all the stress and sleepless nights together and dedicate this piece of paper to ourselves.

ABSTRACT

This thesis verifies that the proposed Kernel mapping based recursive least square algorithm can detect the slightest deviation of anomaly from the norm, monitor and learn underlying pattern between natural and abnormal multivariate medical parameters of a particular critical ICU patient with high detection accuracy and very low rate of false alarm. This online, automated, sequential, real-time intruder detection algorithm is suitable for any instantaneous detection of accidental emergencies without compromising the patient safety and effectiveness of care. It is an elegant, inexpensive solution, independent of complexity, and also a portable and adaptive approach.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our advisor Dr. Tarem Ahmed, Assistant Professor, BRAC University, who has been a tremendous mentor for us. We are thankful to him for letting us to work on this topic, for his patience, motivation, enthusiasm, and for sharing his immense knowledge. His guidance helped us throughout the research. We are greatly indebted to our co-supervisor Mr. Supriyo Shafkat Ahmed, Senior Lecturer, for his whole-hearted support, cooperation, encouragement and readiness to spare his valuable time for us. We are also grateful to Dr. Md. Zakir Hossain, Cardiac Surgeon, United Hospital, for taking out time from his busy schedule to educate and provide us with medical data from the ICU at United Hospital.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

I. INTRODUCTION

In ICU, the technology available to support the critically ill patient is sophisticated and complex, and the importance of intensive care unit in today's health care system is without question. Simultaneously monitoring multivariate readings becomes exhausting for a human, which may cause a bad consequence and increases the probability of important factors going unnoticed. Research in healthcare has shown that patients frequently experience unnecessary sufferings as a result of poor communications amongst ICU nurses and doctors. There are some simple, automated signaling systems, installed with sounded alarms in developing countries but these mostly use simplistic detection algorithms which cause delay in detection and excessive false alarms. These alarms occur even when the patient moves or probes are loosened and do not necessarily signal a change or any deterioration in the condition of the patient. Considering all these problems, we, hereby propose the Kernel-based Online Anomaly Detection (KOAD) algorithm to detect any sudden signal break in ICU. KOAD quickly learns a normal position in a time series of data and is also lightweight in terms of computational and memory resources it requires. It detects what ranges of values are normal for a particular patient and stores it in a dictionary. It raises an orange alarm if values span beyond the normality space, and either verifying that the value is in the range it resolves the orange alarm to green or if it crosses over the normal range a red alarm is raised. So, it is ideally suited for any critical condition in ICU with limited technological facilities and financial constraints. It also ensures accurate detection with low rate of false alarm rates.

A. THESIS CONTRIBUTION

The application of Kernel-based Online Anomaly Detection algorithm to detect sudden emergencies of a critical patient in an ICU of a hospital is proposed. Vital statistics of a patient from a hospital in developing country is monitored and the proposed algorithm is run over the data set to identify the underlying pattern and find the correlation between medical parameters. It is possible to closely monitor multiple measurements, as KOAD learns the behavior of data, and sequentially constructs and maintains a dictionary of measurement vectors which defines the region of normality. Slightest deviation from the norm is detected and analyzed. This is accomplished through the outcome of KOAD run over multi-dimensional data set, including graphs of sensitivity and detection versus false alarm. To fulfillment of our work, the computational complexity is highlighted and also the important parameters that have been used to achieve the maximum detection performance of KOAD are presented.

B. METHODOLOGY

In this thesis, a set of vital statistics of a cardiac patient are collected at an hourly interval from the ICU of a hospital. The data set is revised by a cardiac surgeon and the correlation between the vital statistics is carefully analyzed. The timesteps at which the patient is in critical condition, when the doctor expects that the alarm should be raised, are manually identified and marked. The data is then integrated to work with the proposed method. KOAD is run over the dataset for thresholds v_1 and $v_2$and the projection error is calculated for each timestep. The best values of the threshold are found out exhaustively using trial and error method. The percentage of the detection and the false alarm rate is then computed to see performance of the algorithm. The proposed algorithm is implemented using the software MATLAB.

C. ORGANIZATION OF PAPER

The rest of the thesis is organized as follows. Section II covers the related works that include the available ICU Emergency Signaling Systems and other applications using algorithm based on Kernel method. Section III presents the theoretical framework giving an overview of the monitoring structure of the ICU Signaling System and the mechanism behind the proposed KOAD algorithm. The pseudo code of this anomaly detection algorithm is also provided. Section IV presents the conducted experiments and their results in details including the complexity analysis and parameter selection. Section V, concludes with the accomplishments and suggests potential future works.

II. RELATED WORKS

There have been intensive research and approaches on automating ICU emergency signaling systems in the recent decade. Different approaches have different algorithms for improved result. In this thesis, we focus on recursive Kernel-based Online Anomaly Detection algorithm. Other than KOAD, there are several algorithms used worldwide for different anomaly detection applications.

A. ICU Emergency Signaling Systems

The management of the Intensive Care Unit (ICU) in a hospital has its own, very specific requirements that involve, amongst others, issues of risk-adjusted mortality and average length of stay; nurse turnover and communication with physicians; technical quality of care; the ability to meet patient's family needs; and avoid medical errors due to rapidly changing circumstances and work overload. Good ICU management should lead to an improvement in patient outcomes. Decision making at the ICU environment is a real-time challenge that works according to very tight guidelines, which relate to often complex and sensitive research ethical issues. Vicent J. Ribas Ripoll[1] addresses these through the design and development of computer based decision making tools to assist clinicians at the ICU. It focuses on one of the main problems that they must face: the management of the Sepsis pathology. Sepsis is one of the main causes of death for non-coronary ICU patients. The mortality rate can reach almost up to one out of two patients for septic shock, its most acute manifestation. It is a transversal condition affecting people of all ages. The research reported in this document deals with the problem of Sepsis data analysis in general and, more specifically, with the problem of survival prediction for patients affected with Severe Sepsis. The tools at the core of the investigated data analysis procedures stem from the fields of multivariate and algebraic statistics, machine learning and computational intelligence.

In many biomedical studies, a difference in upper quantiles is of specific interest since the upper quantile represents the upper range of biomarkers and is used as the cutoff value for a disease classification. Jihnhee Yu, Albert Vexler, Alan D. Hutson & Heinz Baumann[2] investigated two-group comparisons of an upper quantile based on the empirical likelihood methodology. Two approaches, the classical empirical likelihood and ‘‘plug-in’’ empirical likelihood, are used to construct the test statistics and their properties are theoretically investigated. Although the plug-in method is developed by the framework of the empirical likelihood, the test statistic is not based on maximization of the empirical likelihood and is simplified by using an indicator function in its construction, making it a unique test to investigate. Extensive simulation results demonstrate that the ‘‘plug-in’’ empirical likelihood approach performs better to compare upper quantiles across various underlying distributions and sample sizes. For the actual application, they employed the developed methods to test the differences in upper quantiles in two different studies: oral colonization of pneumonia pathogens for intensive care unit patients treated by two different oral treatments, and biomarker expressions of normal and abnormal bronchial epithelial cells.

Clinical decision support systems are a combination of software techniques to help the clinicians in their medical decision making process via functionalities ranging from basic signal analysis to therapeutic planning and computerized guidelines. The algorithms providing all these functionalities must be very carefully validated on real patient data and must be confronted to everyday clinical practice. One of the main problems when developing these techniques is the difficulty to obtain high-quality complete patient records, comprising data coming both from the biomedical equipment (high-frequency signals) and from numerous other sources (therapeutics, imagery, clinical actions, etc.).L. Allart, C. Vilhelm ,H. Mehdaoui, H. Hubert, B. Sarrazin, D. Zitouni, M. Lemdani, P. Ravaux[3] present an infrastructure for developing and testing such software algorithms. It is based on a bedside workstation where testing different algorithms simultaneously on real-time data is possible in the ward. It is completed by a collaborative portal enabling different teams to test their software algorithms on the same patient records, making comparisons and cross-validations more easily.

Early recognition of abnormalities in the physiological parameters of hospital patients followed by rapid intervention should result in an improvement in functional outcome or mortality rate. L. Trassenko , A. Hann , A. Patterson,E. Braithwaite , K. Davidson , V. Barber , D. Young[4] developed a real-time system, BioSign™, capable of analyzing physiological parameters in order to identify adverse trends in multi-parameter space (departure from "normality") and prompt clinical staff to intervene. The model of normality is based on five vital signs, which can all be recorded non-invasively: the heart rate, blood pressure, arterial oxygen saturation, respiration rate and temperature. We have evaluated the trained model of normality in a two-year randomized controlled trial of 405 patients monitored on general medical and surgical wards. Preliminary results show that BioSign™ can provide early warning of changes in clinical status and deterioration in patient condition.

B. Applications using Kernel-based Algorithm

The Cerebellar Model Articulation Controller (CMAC) neural network[5] is an associative memory that is biologically inspired by the cerebellum, which is found in the brains of animals. The standard CMAC uses the least mean squares algorithm to train the weights. Recently, the recursive least squares algorithm was proposed as a superior algorithm for training the CMAC online as it can converge in one epoch, and does not require tuning of a learning rate. However, the KRLS algorithms computational speed is dependent on the number of weights required by the CMAC which is often large and thus can be very computationally inefficient. Recently also, the use of kernel methods in the CMAC was proposed to reduce memory usage and improve modeling capabilities. Kernel Recursive Least Squares (KRLS) algorithm has been applied to the CMAC. Due to the kernel method, the computational complexity of the CMAC becomes dependent on the number of unique training data, which can be significantly less than the weights required by non-kernel CMACs. Additionally, online scarification techniques are applied to further improve computational speed.

Bernhard Schölkopf, Alex J. Smola, Robert C. Williamson, Peter L. Bartlett[6] proposed a new class of support vector algorithms for regression and classification. In these algorithms, a parameter ν lets one effectively control the number of support vectors. While this can be useful in its own right, the parameterization has the additional benefit of enabling us to eliminate one of the other free parameters of the algorithm: the accuracy parameter ε in the regression case, and the regularization constant C in the classification case. The algorithms, give some theoretical results concerning the meaning and the choice of ν, and report experimental results.

G. Baudat and F. Anouar[7] presented a new method named Generalized Discriminant Analysis (GDA) to deal with nonlinear discriminant analysis using kernel function operator. The underlying theory is close to the support vector machines (SVM) as the GDA method provides a mapping of the input vectors into high-dimensional feature space. Linear properties help to extend and generalize the classical linear discriminant analysis (LDA) to nonlinear discriminant analysis. The formulation is expressed as an eigenvalue problem resolution. It can be covered a wide class of nonlinearities by using different kernel. They gave classification results, as well as the shape of the decision function for both simulated data and alternate kernels.

Networks often face various anomalous behaviors such as attack on large data transfers in IP networks, sudden break at video surveillance systems and ultimately congestion in a road network. Machine Learning Technique is able to develop the anomaly detection algorithms that are non-parametric, adaptive to changes in the relevant network and portable across applications. Tarem Ahmed, Boris Oreshkin and Mark Coates[8] proposed two different database pictures of a highway in Quebec taken by a network of webcams and IP traffic statistics from the Abilene network. They investigate the use of block-based one-class neighbor machine and the recursive Kernel-based Online Anomaly Detection algorithm.

During the last years, the task of automatic event analysis in video sequences has gained an increasing attention among the research community. The application domains are disparate, ranging from video surveillance to automatic video annotation for sport videos or TV shots. Whatever the application field, most of the works in event analysis are based on two main approaches: the former based on explicit event recognition, focused on finding high-level, semantic interpretations of video sequences, and the latter based on anomaly detection. This work deals with the second approach, where the final goal is not the explicit labeling of recognized events, but the detection of anomalous events differing from typical patterns. Piciarelli, C; Micheloni, C ; Foresti, G.L.[9] proposed this work addressing anomaly detection by means of trajectory analysis, an approached with several application fields, most notably video surveillance and traffic monitoring. The proposed approach is based on single-class support vector machine (SVM) clustering, where the novelty detection SVM capabilities are used for the identification of anomalous trajectories. Particular attention is given to trajectory classification in absence of a priori information on the distribution of outliers.

III. THEORITICAL FRAMEWORK

A. MONIRTORING ARCHITECTURE

Patients in an ICU are monitored using several devices depending on their condition. Patients are connected to equipment to monitor heart rate, blood pressure, respiratory rate, etc. These parameters are often monitored by the bedside unit which is connected to a local monitoring unit next to every critical patient. As there is already a local monitoring unit for each patient, we suggest a distributed network for the system. In this distributed approach, there is local monitoring unit, central monitoring unit and central data repository. The local monitoring unit collects measured data from medical devices connected to the bedside unit. KOAD is locally run in each of these units, for each patient, and the result is send to the central monitoring unit (CMU). If there is an anomaly detected an alarm is raised, and the trained staff on duty in the monitor room is responsible to take an action. Nurses or doctors are informed depending on the situation and immediate actions are taken. The monitor room is at a distance from the ICU rooms. This is because continuous alarms could be disturbance to the patients. It is desirable to have remote access to patient data and this distributed system allows monitoring measurements and alarms from any given location within the network through CMU. This central monitoring unit has limited access, can be accessed only by trained staffs and nurses, doctors and specialists. The central data repository stores the data monitored by the local monitoring unit for each patient. This is the central server where all the ICU data are stored for a fixed period of time. Figure 1 presents the monitoring architecture of the system using a distributed approach and Figure 2 presents the overview of the overall system.

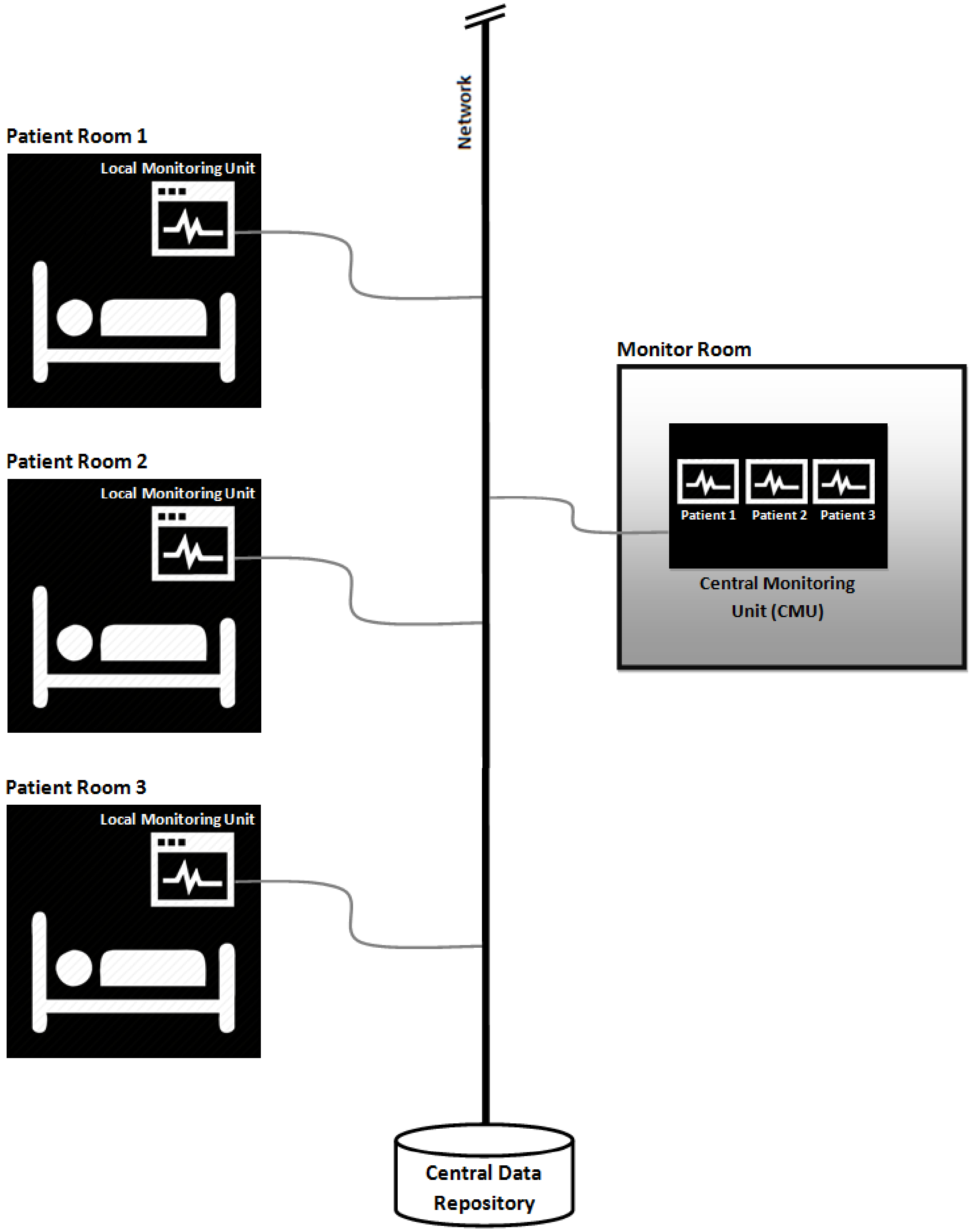

Figure 1 – Monitoring Architecture

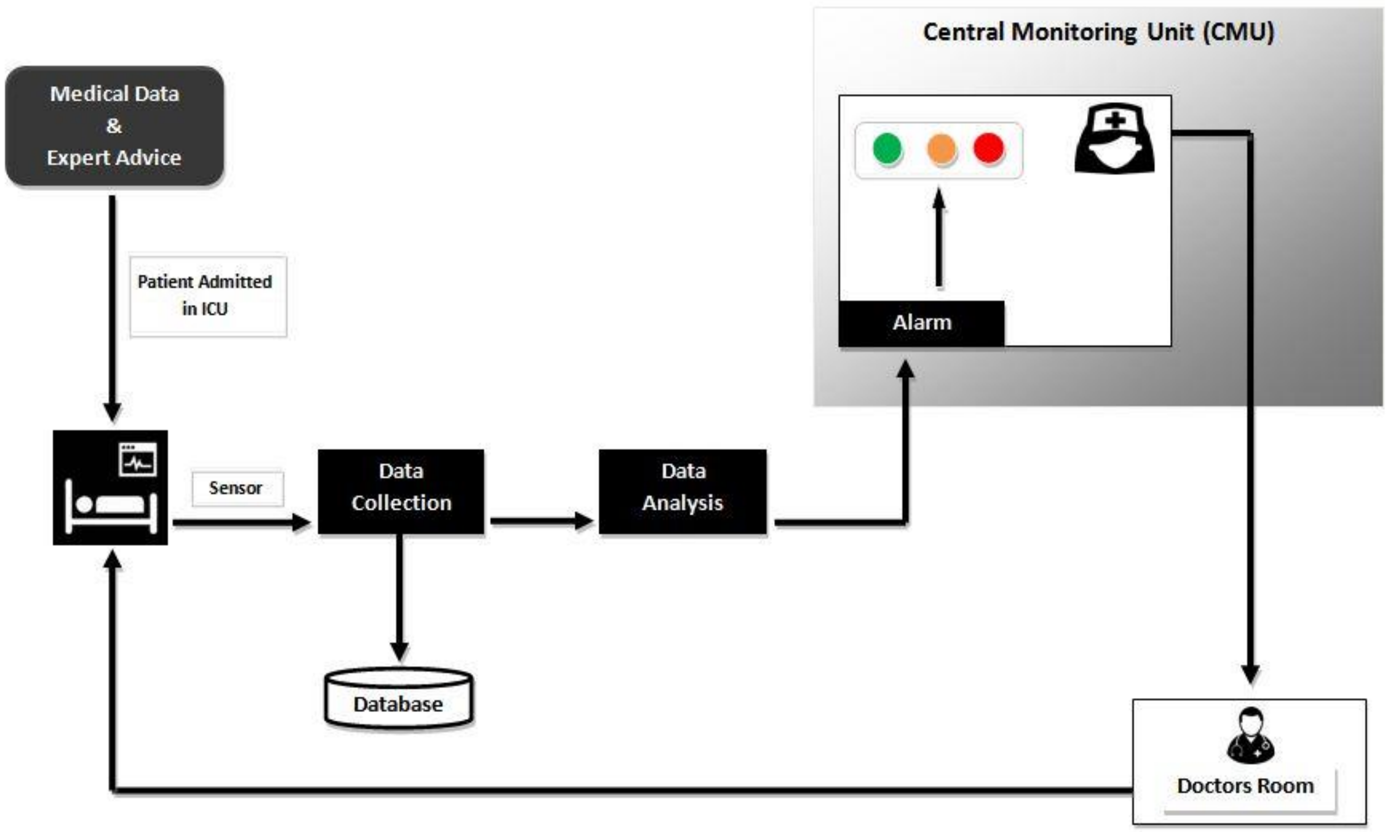

Figure 2 – System Overview

B. KERNEL-BASED ONLINE ANOMALY DETECTION ALGORITHM

Algorithms based on the so-called "kernel trick" involve using a *kernel* function that maps the input data onto a *feature space* of much higher dimension[10], with the expectation that points depicting similar behavior would cluster in the higher dimensional feature space. The idea is that a suitable kernel function, when applied to a pair of input vectors, may be interpreted as an inner product in the feature space. This subsequently allows inner products in the feature space (inner products of the *feature vectors*) to be computed without explicit knowledge of the feature vectors themselves, by simply evaluating the kernel function[10]:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle \qquad (1)$$

Where $\mathbf{x}_i$, $\mathbf{x}_j$ denote the input vectors and Φ represents the mapping onto the feature space.

Consider a set of multivariate measurements $\{\mathbf{x}_t\}_{t=1}^{T}$. In an appropriately chosen feature space F with an associated mapping Φ, the feature vectors corresponding to the points in $\{\mathbf{x}_t\}_{t=1}^{T}$ that depict normal behavior, $\{\Phi(\mathbf{x}_t)\}_{t=1}^{T}$, are expected to cluster. Then, it should be possible to explain the region of normality (in the feature space) using a relatively small *dictionary* of *approximately* linearly independent elements $\{\Phi(\tilde{\mathrm{x}}_j)\}_{j=1}^{m}$ [11]. Here $\{\tilde{\mathrm{x}}_j\}_{j=1}^{m}$ represent those $\{\mathbf{x}_t\}_{t=1}^{T}$ that are entered into the dictionary. The size of the dictionary, m, is expected to be much less than T, thereby leading to computational and storage savings. Feature vector $\Phi(\mathbf{x}_t)$ is said to be *approximately* linearly dependent on $\{\Phi(\tilde{\mathrm{x}}_j)\}_{j=1}^{m}$ with approximation threshold ν, if the projection error $\delta_t$satisfies[11]:

$$\delta_t = \min_{\mathbf{a}} \left\| \sum_{j=1}^{m} a_j \phi(\tilde{\mathbf{x}}_j) - \phi(\mathbf{x}_t) \right\|^2 < \nu. \qquad (2)$$

where $\mathbf{a} = \{a_j\}_{j=1}^{m}$ is the optimal coefficient vector.

The Kernel-based Online Anomaly Detection (KOAD) algorithm operates at each timestep t on a measurement vector $\mathbf{x}_t$. It begins by evaluating the error δ_t in projecting the arriving $\mathbf{x}_t$ onto the current dictionary (in the feature domain). Observe that equation (2) involves an L2 norm, which may be simplified exclusively in terms of the inner products of $\Phi(\tilde{\mathrm{x}}_j)$ and $\Phi(\mathbf{x}_t)$[12], and thus evaluated using the kernel function without explicit knowledge of the feature vectors themselves.

This error measure δ_t is then compared with two thresholds ν_1 and ν_2, where $\nu_1 < \nu_2$. If $\delta_t < \nu_1$, KOAD infers that $\mathbf{x}_t$ is sufficiently linearly dependent on the dictionary, and represents normal behavior. If $\delta_t > \nu_2$, it concludes that $\mathbf{x}_t$ is far away from the realm of normality and immediately raise a "Red1" alarm to immediately signal an anomaly.

If $\nu_1 < \delta_t < \nu_2$, KOAD infers that $\mathbf{x}_t$ is sufficiently linearly independent from the dictionary to be considered an unusual event. It may indeed be an anomaly, or it may represent an expansion or migration of the space of normality itself. In this case, KOAD does the following: it raises an "Orange" alarm, keeps track of the contribution of the relevant input vector $\mathbf{x}_t$ in explaining subsequent arrivals for ℓ timesteps, and then takes a firm decision on it. At timestep $t + \ell$, KOAD re-evaluates the error δ in projecting $\mathbf{x}_t$ onto dictionary $D_{t+\ell}$ corresponding to timestep $t + \ell$. Note that the dictionary may have changed between timesteps t and $t + \ell$, and the value of δ at this re-evaluation may consequently be different from the δ_t at timestep t. If the value of δ after the re-evaluation is found to be less than ν_1, KOAD lowers the orange alarm and keeps the dictionary unchanged.

If the value of δ is found instead to be greater than ν_1 after the re-evaluation at timestep $t + \ell$, KOAD performs a secondary "usefulness" test to resolve the orange alarm. The usefulness of $\mathbf{x}_t$ is assessed by observing the kernel values of $\mathbf{x}_t$ with $\{\mathbf{x}_i\}_{i=t+1}^{t+\ell}$. If a kernel value is high (greater than a threshold d), then $\Phi(\mathbf{x}_t)$ is deemed close enough to $\Phi(\mathbf{x}_i)$. If a significant *number* of the kernel values are high, then $\mathbf{x}_t$ cannot be considered anomalous; normal traffic has just migrated into a new portion of the feature space, and $\mathbf{x}_t$ should be entered into the dictionary. Contrarily if almost all kernel values are low, then $\mathbf{x}_t$ may be concluded to be a reasonably isolated event, and should be heralded as an anomaly.

Equation (3) is evaluated[12]:

$$\left[\sum_{i=t+1}^{t+\ell} \mathbb{I}(k(\mathbf{x}_t, \mathbf{x}_i) > d)\right] > \epsilon\ell, \qquad (3)$$

Where $\mathbb{I}$ is the indicator function and $\varepsilon \in (0, 1)$ is a selected constant.

In this manner, by employing this secondary "usefulness test", KOAD is able to distinguish between an arrival that is an anomaly, from one that is a result of a change in the region of normality. If equation (3) evaluates true, then KOAD lowers the relevant orange alarm to green (no anomaly) and adds $\mathbf{x}_t$ to the dictionary. If equation (3) evaluates false, it elevates the relevant orange alarm to a “Red2” alarm. KOAD also deletes obsolete elements from the dictionary as the region of normality expands or migrates, thereby maintaining a sparse and *current* dictionary. In addition, it incorporates exponential forgetting so that the impact of past observations is gradually reduced.

A pseudo code of the algorithm is presented as Algorithm 1.

Algorithm 1: OUTLINE OF KOAD[12]

Set thresholds v_1, v_2;
Enter $\boldsymbol{x_1}$ to $\mathcal{D}$;
For t=2,3,….**do**
    Evaluate δ_t;
    If $\delta_t > v_2$ **then**
        Raise Red Alarm;
    else if $v_1 < \delta_t \leq v_2$
        Raise Orange Alarm;
        Add $\boldsymbol{x_t}$ to $\mathcal{D}$;
    else /* $\delta_t \leq v_1$ */
    /* Do nothing */;
    endif
    If *Orange*($\boldsymbol{x_{t-\ell}}$) **then**
        Evaluate Usefulness of $\boldsymbol{x}_{t-\ell}$ over past ℓ timesteps;
        If *NOT useful* **then**
            Elevate Orange ($\boldsymbol{x_{t-\ell}}$) to Red;
            Remove $\boldsymbol{x_{t-\ell}}$ from $\mathcal{D}$;
        else
            Delete Orange ($\boldsymbol{x_{t-\ell}}$);
        endif
    endif
    Remove any useless element from $\mathcal{D}$;
endfor

IV. EXPERIMENTS

Intensive Care Unit patients are severely ill and have to be monitored continuously until they reach a stable and improved health condition. Several monitoring devices are used to monitor their vital statistics. Commonly, ICU's practice paper charts to write down their vital statistics for every patient as it easily gives an overview of their condition when the doctor revisits patients. These paper charts with side notes are updated by the appointed nurses time to time. These paper charts helped us to obtain real data for this thesis. The patient that had been monitored for our experiment was admitted in ICU of United Hospital, Dhaka for about four days. During the period, the data had to be collected and tabulated manually from the paper charts. This data, revised by the cardiac surgeon, is manipulated and integrated to work with our application. After final revision, a total of 50 measurement vectors for each of 25 vital statistics were acquired. This data set is tested for anomaly detection using the proposed KOAD algorithm. The objective is to prove that it is possible to obtain high detection accuracy with low false alarm rates. The results include the graphs of sensitivity and detection versus false alarm rates. The complexity analysis is also given as it an important factor in terms space and time. The performance measure of KOAD is stated and the parameter selection throughout the experiment is also mentioned in this section.

A. DATA

Data for testing was collected from the vital statistics paper chart of a cardiac patient in the ICU. As the patient was a cardiac, only statistics related to heart were available. These vital statistics were monitored and tabulated on hourly basis as long as long as the patient was in the ICU. The collected data consisted of 25 parameters recorded in a total of 50 hours of the patient's critical condition. The parameters include the pulse rate, the systolic, diastolic and mean blood pressures, the venous and arterial pressures, the cardiac values, the peripheral temperature, the systemic vascular resistance, the partial pressures and concentrations of oxygen and carbon dioxide, the pH value and some important ion concentrations such as sodium, potassium, chloride and bicarbonate. Among these instances, the most sensitive parameters that depend on one another were carefully analyzed by the cardiac surgeon to identify when it is appropriate to raise an alarm marking those manually to be anomalous. The relationship between different parameters that are dependent, such as the blood pressure directly depends on the cardiac output and the systemic vascular resistance, these had to be carefully monitored to identify critical condition. The other independent parameters, such as the pH and the ion concentrations, were said to be unusual only when they were out of their normal range. The normal range for each of the parameter was also noted down for that particular patient as some variables vary from the average medical range according to their age, height, weight, etc. 16 anomalies had been identified and marked from the data set. After a revision, spreadsheet was used to tabulate the data accordingly to be implemented using KOAD.

B. RESULTS

KOAD is run over the data set of 25 parameters across 50 timesteps to obtain the graphs of sensitivity and detection versus false alarm. The objective is to detect the 16 manually identified anomalies out of the 50 timesteps. The two graphs obtained are of the detection versus false alarm rate and the sensitivity graph.

The detection versus false alarm rate graph shows the trade-off between the probability of detection (P_D) and the probability of false alarms (P_{FA}) as a Receiver Operating Characteristics (ROC) curve. This curve gives an idea of where to set the thresholds to obtain an acceptable balance between detection and false alarm rates. These threshold values can be determined over a training period using a set of normal data and known anomalies. The values differ from patient to patient as every patient does not follow the average normal range for medical parameters due to various issues such as age, height, weight, etc. This can be manually set for each patient according to their health conditions. Here, the lower threshold v_1 is set to fixed value and the upper threshold v_2 is varied to obtain the graph.

The second graph is the sensitivity graph, which shows the projection error δ_t across t timesteps. This shows the value of the error δ_t which is compared to the thresholds to identify the alarm to be raised. When the error is below v_1 it is considered as normal case, raising green alarm. Error above v_2, red alarm is raised as the value is high. In between v_1 and v_2 there is a possibility of either anomaly or normal case, hence an orange alarm is raised and this is further tested to resolve the orange alarm. These thresholds are set manually to obtain best outcome. The magnitude of δ_t for inconsistent values is supposed to be higher than the rest in order to be detected. The thresholds can be varied accordingly to detect higher or lower values as required.

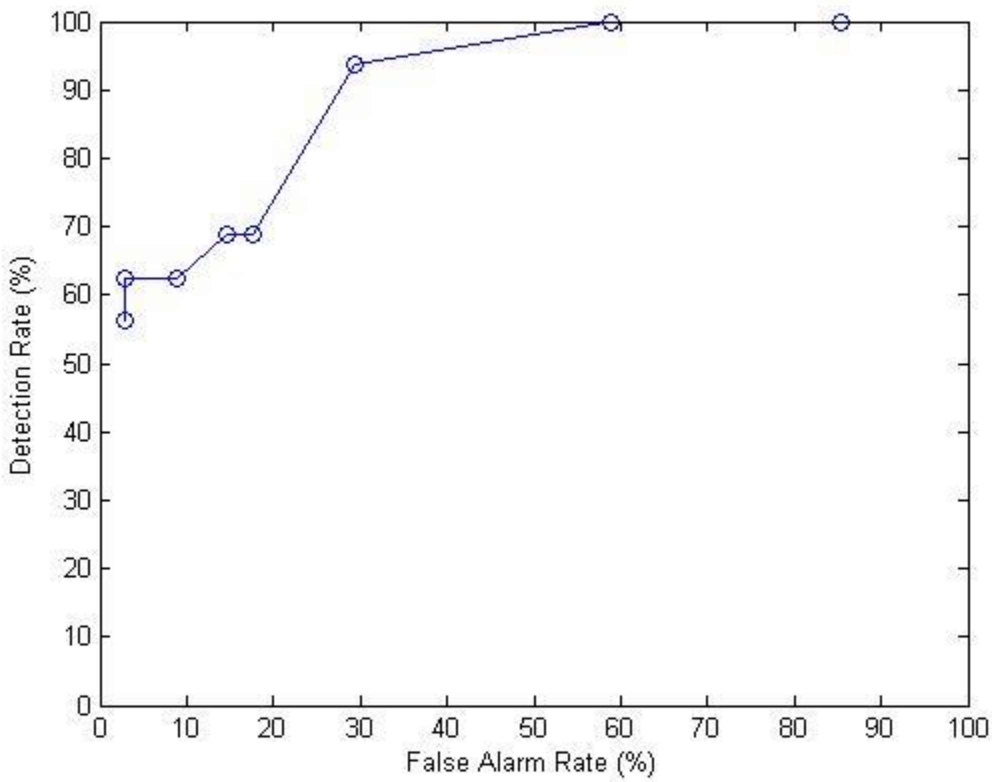

Figure 3 – ROC curve showing KOAD performance

Figure 3 presents the detection versus false alarm rate trade-off as a function of v_2 for a fixed v_1. Here v_1 is set to 0.03 and v_2 is iterated for a range of values higher than v_1. It is observed from Fig.3 that high detection rate is possible at the cost of false alarm rate. 100 percent detection is possible with 58.8 percent being false alarm. Table 1 provides the breakdown KOAD performance for various representative values of v_1 and v_2.

Thresholds		Detected		False Alarm		Missed	
v1	**v2**	**Count**	**%**	**Count**	**%**	**Count**	**%**
0.03	0.04	16	100.0	20	58.8	0	0.0
	0.06	15	93.7	10	29.4	1	6.2
	0.08	11	68.7	6	17.6	5	31.2
	0.10	9	56.2	3	8.8	7	43.7

TABLE 1: Detection of the 16 anomalies for various settings of v_1 and v_2

Detection of all 16 anomaly is possible at thresholds v_1 =0.03 and v_2= 0.04. Though it is important to have a 100 percent detection rate as the patient is in ICU, any critical condition unnoticed could be of high risk, it is not always likely to achieve. There has to be a compromise between detection and false alarm rates. For about 93.7 percent detection rate, the false alarm is 29.4 percent when v_1 =0.03 and v_2= 0.06, which is satisfactory with lower false alarm rate.

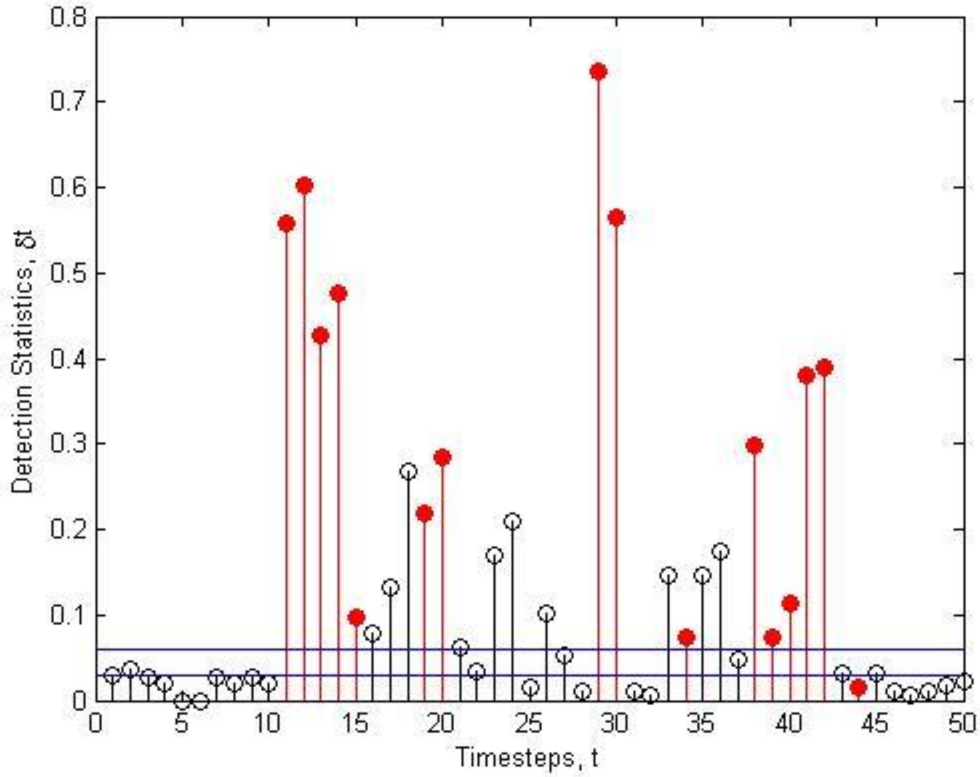

Figure 4 – Progression of the detection statistics *δt*

Figure 4 present the plot of detection statistics δ_t for the 25 vital statistics across 50 timesteps. The timesteps corresponding to the manually identified critical conditions are marked as red, filled stems. The projection error for these marked anomalies show higher value than the rest as expected. Initially the dictionary is being built and so the initial few stems are slightly raised. The most appropriate outcome is obtained using Gaussian kernel function[1] with standard deviation $\sigma = 0.1$. The value of lower threshold set to v_1=0.03 using trial and error method. It is observed from Fig.2 that at thresholds v_1=0.03 and v_2=0.06 maximum detection is possible. 15 out of 16 anomalies are detected successfully and one is missed at timestep t=40. At this point, the patient is in serious health condition and requires an alarm to be raised, but goes unnoticed. With these thresholds, there are a good number of false positives occurring at different timesteps. False alarms are generated because of various reasons, one of which is, at times, a patient might have slight fluctuations in their measurements of specific vital statistics due to effect of medicine or other apparent reason which is normal in medical terms, but mathematically considered as an abnormality. Compromising detection increases the false alarm rate. It is more important to have maximum accurate detection rate with least possible false alarms, hence 93.7 percent detection with 29.4 percent false alarm is an acceptable result.

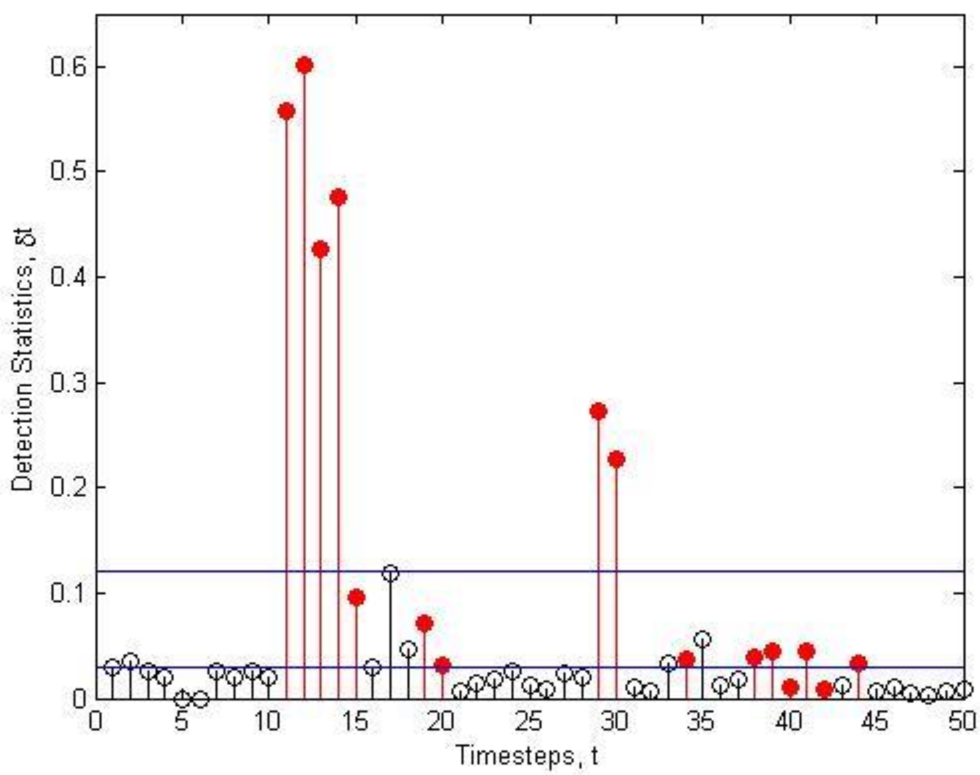

Figure 5 – Progression of the δt at thresholds v_1= 0.03 and v_2 = 0.12

Figure 5 presents the detection statistics δ_t with thresholds set at v_1=0.03 and v_2 = 0.12. This shows that is possible to achieve about 2 percent false alarm rate, but with much lower detection rate with most of the labeled anomaly overlooked.

C. COMPLEXITY ANALYSIS

Intensive care unit is data intensive and hence complexity analysis is an important issue when dealing with related applications. A large number of data arriving at regular intervals, continuously being monitored and analyzed by an application could require huge resource in terms of space and time. However, KOAD algorithm is lightweight concerning computational and memory storage requirements. In terms of storage, the maximum dimensions of the variables that KOAD stores are $m \times m$, where m is the dictionary size. KOAD also retains the input vectors that raise orange alarms for l timesteps, and an additional $L \times m$ binary matrix. The computational complexity is $O(m^2)$ for every standard timestep, and $O(m^3)$ on the rare occasions when an element removal occurs. KOAD complexity is thus independent of time, making it naturally suited for online use.

In terms of actual run time, processing 50 multi-dimensional timesteps took about 1.67 seconds when run on our personal laptop with i5 processor and standard configurations. Hence, it can be concluded that data arriving at one second intervals can be processed smoothly in real time at an ICU.

D. PARAMETER SELECTION

KOAD algorithm depends on several parameters. Kernel parameters are the ones that appear in mapping into feature space. Different kernel functions have different parameters. In our experiment, we worked with Linear, Gaussian and Polynomial kernel functions amongst which Gaussian kernel proved to be the most suitable with $\sigma = 0.1$. Other parameters are the thresholds ν_1 and ν_2, which directly controls the detection performance of KOAD. These thresholds are set accordingly using trial and error for the dataset with known anomalies so that the detection is maximized and the error minimized. Currently this is set manually for testing, which can be also be automatically set in future using training set of data with known anomalies to achieve an acceptable compromise between detection and false alarm rates. The parameters such as l and ε for the orange alarm resolution or the dropping parameters L and d, they do not have much effect on the performance of the algorithm. Their choice is not particularly sensitive and thus is set to the default values that suit. We worked with values

l=10, ε =0.2, L=50, d=0.9 to accomplish our results.

V. CONCLUSION AND FUTURE WORK

Technology has led to obvious improvement in Intensive Care Unit worldwide within last few years. Despite the pace of technological promise, there is a very slow progress in developing countries in terms of ICU Automation. Bedside technologies are used in an ICU of a private hospital but due to lack of resources, networking facilities, training, etc. these cannot be integrated to provide much effective services. ICU emergency signaling could be at step taken towards automation we say. For developing countries this application using KOAD is a sophisticated, inexpensive solution for the detection of sudden break in signal pattern in an ICU.

As stated earlier, we believe that this recursive algorithm based on Kernel mapping can achieve faster time-to-detection with high accuracy. Alarms in medical devices are a matter of concern in critical and preoperative care. The high rate of false alarms can compromise patient safety and effectiveness of care. So, we propose a great detection performance which works almost perfectly with very low computational complexity. We restricted our research only to detect anomalous events immediately, which alerts when measurements cross pre-set limits with negotiable alarm rates. We emphasize on maximum detection and make it fully suitable for online use in hospital ICUs.

Moving ahead for future works, investigation could be carried out implementing KOAD using more unstable, complex real data from neonatal ICU (NICU).Supplementary algorithm could be adopted to automatically set the KOAD thresholds[13]. The time interval between arriving data could be reduced to minutes or seconds and dataset of 200 or more timesteps could be used for improved statement. A comparative study of KOAD with other common anomaly detection algorithms such as Kernel Principal Component Analysis (KPCA) could be performed.

REFERENCES

[1] Vicente J. Ribas Ripoll, (2012).“*Study of the Prognosis in the Intensive Care Unit*”.Centre de RecercaMatemàtica.

[2] Jihnhee Yu, Albert Vexler, Alan D. Hutson & Heinz Baumann, (published online 01 Feb. 2014). “*Empirical Likelihood Approaches to Two-Group Comparisons of Upper Quantiles Applied to Biomedical Data”.* pp. 30-40.

[3] L. Allart, C. Vilhelm,H. Mehdaoui, H. Hubert, B. Sarrazin, D. Zitouni, M. Lemdani, P. Ravaux,(July 2008). “*Anarchitecture for online comparison and validation of processing methods and computerized guidelines in intensive care units”.*

[4] L. Trassenko, A. Hann, A. Patterson,E. Braithwaite, K. Davidson, V. Barber, D. Young, (2005). “*multi-parameter monitoring for early warning of patient deterioration*”.3rd IEE International Seminar on Medical Applications of Signal Processing, pp.71 – 76.

[5] C. W. Laufer and G. Coghill, (June 2013).“*Kernel Recursive Least Squares for the CMAC Neural Network”.*International Journal of Computer Theory and Engineering, Vol. 5, No. 3.

[6] Bernhard Schölkopf, Alex J. Smola, Robert C. Williamson, Peter L. Bartlett,(May 2000,Posted Online March 13, 2006).“*New Support Vector Algorithms”.*Vol. 12, No.5, pp.1207-1245.

[7] G. Baudat**and**F.Anouar, (Oct. 2000, Posted Online March 13, 2006). “*Generalized*

Discriminant Analysis Using a Kernel Approach".Vol.12, No.10, pp. 2385-2404.

[8] **T. Ahmed, B. Oreshkin and M. Coates, (April 2007)."***Machine Learning Approaches to Network Anomaly Detection*".

[9] C.Piciarelli, C.Micheloni, G.L.Foresti, (Sept. 2008). "*Trajectory-Based Anomalous Event Detection*", Circuits and Systems for Video Technology, IEEE Transactions on (Volume.18, Issue: 11),pp.1544 – 1554.

[10] B. Scholkopfand, A. Smola, (Dec. 2001). "*Learning with Kernels*", Cambridge, MA: MIT Press.

[11] Y. Engel, S. Mannor, and R. Meir, (Aug. 2004). "*The kernel recursive least squares algorithm*". *IEEE Trans. Signal Proc.*, Vol. 52, No. 8, pp. 2275–2285.

[12] T. Ahmed, M. Coates, and A. Lakhina, (May 2007)."*Multivariate online anomaly detection using kernel recursive least squares*", in Proc. IEEE INFOCOM, Anchorage, AK.

[13] T. Ahmed, (Nov. 2009). "*Online anomaly detection using KDE*", in Proc. IEEE, Global Communications Conf. (GLOBECOM), Honolulu, HI, USA.

[14] T. Ahmed, Sabrina Ahmed, SupriyoAhmed,M. Motiwala, (May 2010). "*Real-time Intruder Detection in Surveillance Networks using Adaptive Kernel Methods*", in Communications (ICC), IEEE International Conf., Cape Town.

[15] R. Bengali, N. Yeazdani, N. Zaman, T. Ahmed, S. Ahmed, (2013) "*Taking Meredith out of Grey's Anatomy: Automating Hospital ICU Emergency Signaling*"