Research Article

SVM Based Event Detection and Identification: Exploiting Temporal Attribute Correlations Using SensGru

Nauman Shahid,[1] Ijaz Haider Naqvi,[2] and Saad Bin Qaisar[3]

[1] *EPFL Doctoral School of Information and Communication Sciences (EDIC), 1015 Lausanne, Switzerland*
[2] *Department of Electrical Engineering, LUMS SBA School of Science and Engineering, DHA, Lahore 54792, Pakistan*
[3] *School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad 44000, Pakistan*

Correspondence should be addressed to Nauman Shahid; nauman.shahid@epfl.ch

In the context of anomaly detection in cyber physical systems (CPS), spatiotemporal correlations are crucial for high detection rate. This work presents a new quarter sphere support vector machine (QS-SVM) formulation based on the novel concept of *attribute correlations*. Our event detection approach, SensGru, groups multiple sensors on a single node and thus eliminates communication between sensor nodes without compromising the advantages of spatial correlation. It makes use of temporal-attribute (TA) correlations and is thus a TA-QS-SVM formulation. We show analytically that SensGru (or interchangeably TA-QS-SVM) results in a reduced node density and gives the same *event detection* performance as more dense *Spatiotemporal-Attribute Quarter-Sphere SVM* (*STA-QS-SVM*) formulation which exploits both spatiotemporal and attribute correlations. Moreover, this paper develops theoretical bounds on the internode distance, the optimal number of sensors, and the sensing range with SensGru so that the performance difference with SensGru and STA-QS-SVM is negligibly small. Both schemes achieve event detection rates as high as 100% and an extremely low false positive rate.

1. Introduction

In cyber physical systems (CPS), monitoring and control applications using wireless sensor networks (WSN) have been gaining interest [1–3]; however, such systems are vulnerable to anomalies (malicious attacks, noise, and errors) [4–6]. Therefore, it is essential to identify anomalies to provide reliable functioning of network [7–9]. Anomalies may result from sensor faults (errors or outliers) or from an event of interest [10]. Anomaly and event detection in WSNs finds various applications [11–14], especially, disaster prevention, detection, and identification, for instance, in underground mine wells and volcanic sites [4].

A comparison of various types of outlier detection techniques for WSNs has been carried out in [1]. *Classification based techniques* learn a classification model in the training phase and then classify the data instance to one of the training classes [15]. They are grouped into *Support Vector Machine (SVM)* based and *Bayesian Network* based approaches, depending upon the type of classification model

that is used [1, 16]. In wireless sensor networks, a majority of energy is consumed in radio communication rather than in computation [1]. Hence, in order to make the system energy efficient, it is advantageous to decrease the communication overhead at the cost of increasing computational overhead in the network. SVM based approaches naturally fit in for these requirements, as they are unsupervised and nonparametric and perform in-network processing.

The earliest known formulations of Support Vector Machines construct a set of hyperplanes [17, 18], hyperspheres [19], quarter-spheres (QS-SVM) [20], hyperellipsoids (TOCC) [21], centered hyperellipsoids [22] (CESVM), or conic segments (CS-SVM) [23] in a high dimensional space, which can be used for separation of normal and anomalous data (outliers and events). QS-SVM and CE-SVM have an advantage of linear optimization formulation as compared to the quadratic optimization for other. CS-SVM is more suited to multiclass problems; therefore, it is not used for outlier detection (one-class problem). A detailed analysis of one-class SVM based outlier detection techniques has

(a) Region R for STA-QS-SVM
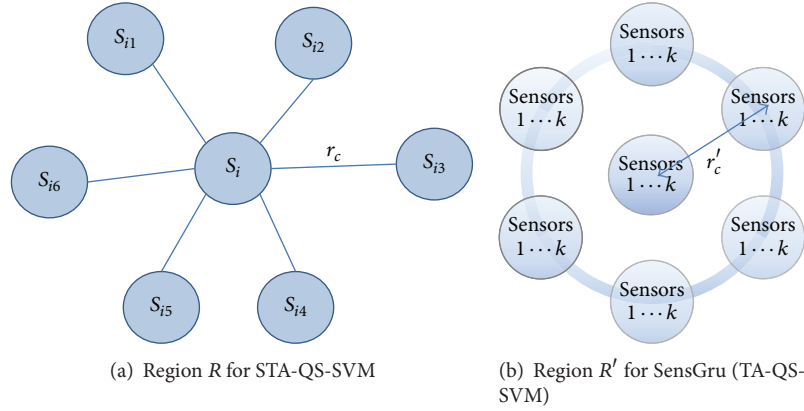
(b) Region R' for SensGru (TA-QS-SVM)

FIGURE 1: Regions R and R' for STA-QS-SVM and TA-QS-SVM approaches.

been carried out in [24]. Amongst online schemes, QS-SVM scheme is more communication efficient as compared to online CE-SVM [25] and is therefore preferred over CE-SVM for outlier detection in WSNs [26]. QS-SVM techniques for outlier detection *use only spatiotemporal correlations of nodes' data,* (hence, we call them Spatiotemporal Quarter-Sphere SVM (ST-QS-SVM)), ignoring the attribute correlations which must be incorporated for better outlier and event detection rates [1] (event detection rates have never been presented for these techniques).

In our previous work, STA-QS-SVM formulation has been proposed [27] which incorporates attribute correlations into ST-QS-SVM and achieves better detection and false positive rates as compared to ST-QS-SVM. This paper introduces SensGru, a novel energy efficient, temporal-attribute based (TA-QS-SVM) formulation, which groups multiple sensors at a single node, and presents a thorough analysis of event detection performance of STA-QS-SVM and SensGru. We show analytically that SensGru (or interchangeably TA-QS-SVM) results in a reduced (as much as 75%) node density and gives the same *event detection* performance as more dense STA-QS-SVM formulation. Moreover, this paper develops theoretical bounds on the internode distance, the optimal number of sensors, and the sensing range with SensGru so that the performance difference with SensGru and STA-QS-SVM is negligibly small. Both schemes achieve event detection rates as high as 100% and an extremely low false positive rate.

The main contributions of this paper are as follows.

(1) It establishes that SensGru can give a performance which is within a negligible difference to STA-QS-SVM. With this benchmark performance with Sens-Gru, this paper further derives.

(2) The ratio of internode distance with SensGru to internode distance with STA-QS-SVM.

(3) Maximum sensing range with both schemes.

(4) The optimal value of the number of sensors with SensGru.

(5) The worst case performance analysis and distance ratio.

(6) The complexity analysis of both schemes.

2. Problem Statement

Figure 1 presents the network model for both STA-QS-SVM and SensGru. The model for STA-QS-SVM considers a densely deployed homogeneous wireless sensor network in a stationary environment (whose spatial and temporal characteristics do not change), where the sensor nodes are localized and connected to each other, such that the sensor data are correlated in time and space. The nodes in the network are represented as S_i. In a neighborhood, represented as $N(S_i)$, k boundary nodes are connected to a central node with an internode distance of r_c units (as shown in Figure 1(a)). For TA-QS-SVM based SensGru, the internode distance is r_c'; however, k sensors per attribute are grouped together on each node (as shown in Figure 1(b)). Hence, no communication is required for event detection based on temporal and attribute correlations. In case an anomaly gets detected, the nodes transmit with a higher power in order to increase their transmission radius and communicate with each other (details follow).

2.1. Spatiotemporal and Attribute Correlations Based Outlier Detection in WSNs (STA-QS-SVM). For STA-QS-SVM, the central node S_i is considered to be within the radio transmission range r_c of all other nodes in $N(S_i)$. The k spatially neighboring nodes of S_i are represented by S_{ij}, such that $\{j = 1, 2, \ldots, k\}$. At each time interval m, each sensor node S_i in the set $N(S_i)$ measures a d-dimensional data vector $\overrightarrow{x_m}$. Thus, a sensor node equipped with d different types of sensors will sense a d-point data vector. Let $\overrightarrow{x_m^0}, \overrightarrow{x_m^1}, \overrightarrow{x_m^2}, \ldots, \overrightarrow{x_m^k}$ denote the d-point data vectors at $S_i, S_{i1}, S_{i2}, \ldots, S_{ik}$, in the set $N(S_i)$, at the mth time instant, respectively. The goal is to identify every new measurement arriving at S_i as normal or anomalous in real-time. STA-QS-SVM determines the radius of quarter-sphere based on *attribute* correlations in addition to the *spatiotemporal* correlations.

Let R^i and R^{*i} be temporal and attribute radii of quarterspheres that have to be determined during the training phase from a set of n measurements $\mathbf{X}_i = \{\overrightarrow{x_1^i}, \overrightarrow{x_2^i}, \ldots, \overrightarrow{x_n^i}\}$ at the nodes S_i in the set $N(S_i)$, corresponding to n time instants, where d is the number of attributes/data points corresponding to each vector. Let the vectors $\overrightarrow{x_m^i}$ at each time instant m be mapped onto feature space via some function $\phi(x_m^i)$.

The constrained optimization problem of the one-class centered quarter-sphere SVM based on spatiotemporal correlations (ST-QS-SVM) is formulated as in [26] to obtain the radius R^i and identify support, nonsupport, and border support vectors based on the Lagrangian multiplier α_m values. The other part of this problem consists of applying the quarter-sphere SVM formulation along the d attributes of received vectors $\overrightarrow{x_m^i}$. This is done at each node S_i of the set $N(S_i)$ and at each time instant m, to determine the radius R^{*i} and the set of support vectors based on attribute correlations [27].

Consider the training data set $\mathbf{X}_i$ consisting of n, d point data vectors $\overrightarrow{x_1^i}, \overrightarrow{x_2^i}, \overrightarrow{x_3^i}, \ldots, \overrightarrow{x_n^i}$, at node S_i and let the vectors $\overrightarrow{x_m^i}$ at each time instant m be mapped onto feature space via some function $\phi^*(x_m^i)$. We divide the received data set $\mathbf{X}_i$ of dimension $n \times d$ into $\lfloor n/d \rfloor$ portions of $d \times d$ dimensions each named as $\mathbf{X}_i^p$. Each row of $\mathbf{X}_i^p$ corresponds to a data measurement at a specific time instant, whereas each column corresponds to a specific attribute over different time instants. Our approach, based on attribute correlations, applies the constrained optimization problem of one-class quarter-sphere to each $\mathbf{X}_i^p$, where $p = 1, 2, \ldots, \lfloor n/d \rfloor$, using each column of $\mathbf{X}_i^p$ as a d-point data vector. Thus, we treat a single attribute of the d consecutive time measurements as a vector for optimization purpose, in contrast to the previous spatiotemporal approach (ST-QS-SVM), which takes into account each row of $\mathbf{X}_i$ as a vector for optimization. For details on formulation of $\lfloor n/d \rfloor$ dimensional vector optimization problem and its solution and the algorithm for outlier and event detection, the reader is referred to [27].

By using some effective linear optimization techniques, we can compute the Lagrangian multipliers, $\alpha_{p,q}^{*i}$, $\forall q = 1, 2, \ldots, d$, $p = 1, 2, \ldots, \lfloor n/d \rfloor$. We obtain $\lfloor n/d \rfloor$ sets of $\alpha_{p,q}^{*i}$, corresponding to $\lfloor n/d \rfloor$ portions of $\mathbf{X}_i$, each set containing d number of $\alpha_{p,q}^{*i}$. The vectors present in each of the pth portion of $\mathbf{X}_i$ can be further classified depending on the results of $\alpha_{p,q}^{*i}$. The data vectors with $\alpha_{p,q}^{*i} = 0$, which fall inside the quarter-sphere and whose distances from the origin are smaller than the radius of the quarter-sphere, are called nonsupport vectors. The data vectors with $\alpha_{p,q}^{*i} > 0$ are called margin support vectors. Their distances to the origin indicate the radius of the quarter-sphere. Support vectors with $\alpha_{p,q}^{*i} = 1/vd$, which fall outside the quarter-sphere and whose distances from the origin are larger than the radius of the quarter-sphere, are called nonmargin support vectors. These data vectors are called outliers. The norm of support

vectors with $0 < \alpha_{p,q}^{*i} < 1/vd$ determines the radius R^{*i}. Since we have $\lfloor n/d \rfloor$ sets of support vectors, we determine $\lfloor n/d \rfloor$ values of radius R_p^{*i} from each set using these support vectors

$$R_p^{*i2} = k_c^* \left(\overrightarrow{x_{p,q}^i}, \overrightarrow{x_{p,q}^i} \right), \tag{1}$$

where $k_c^* (\overrightarrow{x_{p,q}^i}, \overrightarrow{x_{p,q}^i})$ is a centralized kernel function. The final attribute radius is then determined by taking the median of all $\lfloor n/d \rfloor$ radii:

$$R^{*i} = \text{median} \left\{ R_1^{*i}, R_2^{*i}, \ldots, R_{\lfloor n/d \rfloor}^{*i} \right\}. \tag{2}$$

2.2. SensGru: The Communication Independent Approach to Anomaly Detection. Now, consider the problem of a wireless sensor network, consisting of localized nodes S_i. Assume that the nodes do not lie in a closed neighborhood to any other nodes and are separated by a distance r_c'. Also, let $r_c' = C \cdot r_c$, where $C > 1$. Also, assume that each region R' contains only one node S_i (as shown in Figure 1(b)) with k sensors for each attribute. Hence, no communication is required for majority voting based event detection. The SVM based approaches identify outliers based on temporal deviations and then exploit spatial information of neighboring nodes to identify events from the detected outliers. This approach, however, considers exploiting only the temporal and attribute information at each node S_i, to identify both outliers and events. Assuming that an event at node S_i is equally likely to be detected in the neighborhood $N(S_i)$ due to spatial correlations [12], events are identified by introducing a consensus of multiple sensors at S_i. The probability of a faulty node reporting a low measurement in an event region cannot be ignored [12]. Thus, the presence of multiple sensors for each attribute at each node incorporates the flexibility of some faulty sensors at each S_i. Assuming that the probability of all sensor nodes being faulty together is very low, this system detects events at a high rate based on majority voting.

The node S_i in R', thus, has k sensors for measuring each of the d attributes, where $k = \{2, 3, \ldots, t\}$, t being the maximum number of sensors which can be deployed at S_i. We define a sensor array as a group of sensors which measure d different attributes. Let s_j^i represent the jth sensor array for measuring d attributes at ith node:

$$\mathbf{S}_i = \left\{ \mathbf{s}_1^i, \mathbf{s}_2^i, \ldots, \mathbf{s}_k^i \right\}, \tag{3}$$

where each $\mathbf{s}_j^i$ measures d different attributes

$$\mathbf{s}_j^i = \begin{bmatrix} s_{1j}^i & s_{2j}^i & \cdots & s_{dj}^i \end{bmatrix}^T. \tag{4}$$

Let $\mathbf{x}_{dm}^j$ be the d attribute data vector corresponding to the jth sensor array measured at mth time instant at ith node, such that $j = \{1, 2, \ldots, k\}$ and $m = \{1, 2, \ldots, n\}$. Thus, a set of k data vectors $\mathbf{X}_m$, at the jth node, corresponding to k sensor arrays will be obtained at the mth time instant:

$$\mathbf{X}_m = \left\{ \mathbf{x}_{dm}^1, \mathbf{x}_{dm}^2, \ldots, \mathbf{x}_{dm}^k \right\},$$

$$\text{where } \mathbf{x}_{dm}^j = \begin{bmatrix} x_{1m}^j & x_{2m}^j & \cdots & x_{dm}^j \end{bmatrix}^T. \tag{5}$$

The problem is to identify each newly arrived data vector $\mathbf{x}_{dm}^{j}$ at each of the k sensor arrays as normal or anomalous in real time using temporal and attribute correlations only, based on quarter-sphere one-class SVM formulation.

When a new data measurement x_{new}^{j} arrives at node S_i, it computes the distance $d(x)_c^{j}$, which is the mean of the distances between x_{new}^{j} and origin of its centered quarter-spheres in the feature space, using the kernels. Thus, a set of k distances $d(x)_c^{j}$, corresponding to k sensor arrays, is obtained at each node S_i. Now, we make use of the reading from k sensors to find the spatial radius R^i. The temporal radius R^{*i} is found in a similar manner described above. Once we have both these radii, we use the algorithm presented in [27] to identify an outlier or event at each sensor node. The two approaches are linked by the analysis that follows. This paper addresses several unanswered questions. The most important contribution of this paper is that it develops theoretical bounds on important parameters with SensGru required to achieve the same performance as the spatiotemporal-attribute approach. It serves as a mathematical justification of the fact that SensGru is feasible (much more cost efficient), at a cost of deploying a few additional sensors at each node.

3. Analysis of Event Detection in STA-QS-SVM and TA-QS-SVM

Let R be a uniform circular region consisting of exactly k sensors for the STA-QS-SVM approach. $k - 1$ of the k sensors are uniformly distributed around the boundary at a distance of r_c from the sensor present at the center of the region. Also, consider another similar region R', consisting of a single sensor node with exactly k sensors, for the TA-QS-SVM approach. Let $r_c' = C \cdot r_c$, where $C > 1$, (we call C the distance ratio) be the distance between the node of the region R' and nodes in similar neighboring regions. Each sensor in regions R and R' has a radio communication range r_c and a sensing range r_s. The problem is to evaluate C, such that the performance of the event detection techniques in both the deployments is approximately the same:

$$\left| \text{DR}_{\text{STA}} - \text{DR}_{\text{TA}} \right| < \epsilon, \tag{6}$$

where $\epsilon \ll 1$ and DR_{STA} and DR_{TA} are the detection rates of STA-QS-SVM and TA-QS-SVM techniques, respectively, and C is defined as the distance ratio:

$$C = \frac{\text{Distance between the nodes in TA-QS-SVM } \left(r_c' \right)}{\text{Distance between the nodes in STA-QS-SVM } \left(r_c \right)}. \tag{7}$$

Let n be the total number of measurements in time interval T and let $n(E_a)$ be the number of events in a region R or R'. Then, $n(E_a) = n\mu$, where μ is the fraction of events occurring in T. Also, let $n(E_{\text{STA}})$ and $n(E_{\text{TA}})$ be the number of

events correctly detected by STA-QS-SVM and TA-QS-SVM approaches, respectively. Then,

$$\text{DR}_{\text{STA}} = \frac{n(E_{\text{STA}})}{n(E_a)},$$
$$\text{DR}_{\text{TA}} = \frac{n(E_{\text{TA}})}{n(E_a)}. \tag{8}$$

Substituting (8) in (6), we get

$$\left| n(E_{\text{STA}}) - n(E_{\text{TA}}) \right| < n\epsilon\mu. \tag{9}$$

Proposition 1. *Let DR_{STA} and DR_{TA} be the event detection rates for STA-QS-SVM and TA-QS-SVM approaches; then, $DR_{STA} = p(E_{STA} \mid E_a)$ and $DR_{TA} = p(E_{TA} \mid E_a)$, where $p(E_{STA} \mid E_a)$ and $p(E_{TA} \mid E_a)$ are the probabilities of event detection by STA-QS-SVM and TA-QS-SVM approaches, respectively, given that an event has occurred.*

Proof. Let p_d be the probability of the detection of a sensor given as

$$p_d = \begin{cases} e^{-\beta x} & \text{if } x \leq r_s, \\ 0 & \text{else,} \end{cases} \tag{10}$$

where x is the distance of a sensor from event/outlier location and β is a constant that depends on the physical characteristics of an environment. We make the following assumptions in order to prove this proposition.

(1) Let p_e be the probability of the error of a sensor; then, $p_e \approx 0$.

(2) An outlier (extreme condition) can occur uniformly in a region.

(3) An outlier occurring in a given region is independent of other regions.

(4) An event is reported if there is an outlier on $\lceil k/2 \rceil + 1$ sensors in the region.

(5) An event is within the sensing range of all the sensors in the region.

(6) The probability of detection of a sensor at $x = r_s$, that is, $p_d(x = r_s) \geq 1/2$.

(7) The probability of detection of an event, given that assumption (6) is satisfied, is

$$p(E_{\text{STA}} \mid E_a) \geq 1 - \delta_{\text{STA}},$$
$$p(E_{\text{TA}} \mid E_a) \geq 1 - \delta_{\text{TA}}, \tag{11}$$

where δ_{STA} and δ_{TA} are the SVM generalization errors for STA-QS-SVM and TA-QS-SVM approaches, respectively.

We analyze a set of n measurements in a time interval T. Thus, each instant corresponds to the small interval $\tau = T/n$. Let λ be the rate of arrival of an outlier in a region per τ. The rate of arrival of outliers can be modeled as a Poisson process

with parameter λ. Assuming that $\lambda \ll 1$ and only one outlier arrives at an instant τ, the probability of the arrival of one outlier in an instant τ is

$$p(1, \tau) = \lambda \tau e^{-\lambda \tau}. \tag{12}$$

Further, we assume that an arriving outlier is equally likely to touch all the sensors. Thus, the probability that an outlier occurs and affects the sensor is

$$p_o = \frac{\lambda \tau e^{-\lambda \tau}}{k}. \tag{13}$$

Thus, the probability of outlier detection at an instant τ, using (10) and (13), is given as

$$p_{O_d} = \frac{\lambda \tau e^{-\lambda \tau} e^{-\beta x}}{k}; \tag{14}$$

also,

$$p_{O_d} = p(O_d \mid O) p_o + p(O_d \mid O^c) p'_o, \tag{15}$$

where $p(O_d \mid O)$ is the probability of detection of an outlier given that an outlier occurs and $p(O_d \mid O^c)$ is the probability of detection of an outlier given that an outlier does not occur. From the first assumption $p_e \approx 0$, thus, $p(O_d \mid O^c) = p_e = 0$. Comparing (13), (14), and (15) we get

$$p(O_d \mid O^c) = e^{-\beta x}. \tag{16}$$

Again from assumption (1), as $p(O_d \mid O^c) = p_e = 0$, we need not consider the interarrival times of outlier in the outlier detection or event detection performance. Hence, the problem of comparing the event detection rates of both the techniques becomes independent of the interarrival times. Hence,

$$p(E_{\text{STA}} \mid E_a) = \sum_{j=\lceil k/2 \rceil + 1}^{k} \sum_{l=1}^{\binom{k}{j}} \left(e^{-\beta x_l^j} \right)^j \left(1 - e^{-\beta x_l^j} \right)^{k-j},$$

$$p(E_{\text{TA}} \mid E_a) = \sum_{j=\lceil k/2 \rceil + 1}^{k} \binom{k}{j} \left(e^{-\beta x_j} \right)^j \left(1 - e^{-\beta x_j} \right)^{k-j}. \tag{17}$$

Thus, using the assumptions (5) and (6) and the abovementioned facts, the detection rates DR_{STA} and DR_{TA} are equal to the probabilities of the detection of an event given in which an event occurs, that is, $p(E_{\text{STA}} \mid E_a)$ and $p(E_{\text{TA}} \mid E_a)$. □

Thus, (6) can be written as

$$\left| p(E_{\text{STA}} \mid E_a) - p(E_{\text{TA}} \mid E_a) \right| < \epsilon. \tag{18}$$

Using (17) in (18) we obtain

$$\left\| \sum_{j=\lceil k/2 \rceil + 1}^{k} \sum_{l=1}^{\binom{k}{j}} \left(e^{-\beta x_l^j} \right)^j \left(1 - e^{-\beta x_l^j} \right)^{k-j} \right. \tag{19}$$

$$\left. - \sum_{j=\lceil k/2 \rceil + 1}^{k} \binom{k}{j} \left(e^{-\beta x_j} \right)^j \left(1 - e^{-\beta x_j} \right)^{k-j} \right\| < \epsilon. \tag{20}$$

The exact value of $p(E_{\text{STA}} \mid E_a)$ in (19) is complex to evaluate, since the value of x_l^j is different for different combinations of sensors. Hence, we perform an average case analysis for $p(E_{\text{STA}} \mid E_a)$. We find the expected value of distance, that is, x_{STA}, between the event and sensors, for the STA-QS-SVM approach. Let (ρ_1, θ_1) be the coordinates of the event in the region R, centered at the origin. Then, PDF of ρ_1 and θ_1 are

$$f_{\rho_1}(\rho_1) = \begin{cases} \dfrac{1}{r_c} & \text{for } 0 \leq \rho_1 \leq r_c, \\ 0 & \text{else}, \end{cases}$$

$$f_{\theta_1}(\theta_1) = \begin{cases} \dfrac{1}{2\pi} & \text{for } 0 \leq \theta_1 \leq 2\pi, \\ 0 & \text{else}. \end{cases} \tag{21}$$

Let (ρ_2, θ_2) be the coordinates of the sensors present at the boundary of region R:

$$f_{\rho_2}(\rho_2) = \begin{cases} \dfrac{k-1}{k} & \text{for } \rho_2 = r_c, \\ \dfrac{1}{k} & \text{for } \rho_2 = 0, \\ 0 & \text{else}. \end{cases} \tag{22}$$

Thus, for STA-QS-SVM, the expected value of the distance of a sensor from the event location x_{STA} is given using the cosine law and Jensen's inequality as

$$E\left[\sqrt{\rho_1^2 + \rho_2^2 - 2\rho_1\rho_2 \cos(\theta_1 - \theta_2)} \right]$$

$$\geq \sqrt{E[\rho_1^2] + E[\rho_2^2] - 2E[\rho_1\rho_2 \cos(\theta_1 - \theta_2)]}. \tag{23}$$

Assuming that $\cos(\theta_1 - \theta_2)$ is independent of ρ_1, we get

$$x_{\text{STA}} = r_c \sqrt{\frac{1}{3} + \left(\frac{k-1}{k} \right)^2}. \tag{24}$$

For the TA-QS-SVM approach, we note that r_c' is the distance between the sensor nodes. Since the sensor nodes need not communicate with each other, in the worst case, a sensor node can detect an event at a distance of $r_c'/2$. Assume that the event is equally likely to occur in the range of 0 to $r_c'/2$. Thus, the expected distance of a sensor from event location

$(x_{TA}) = r'_c/4 = (C \cdot r_c)/4$. Since we are evaluating $p(E_{STA} \mid E_a)$ and $p(E_{TA} \mid E_a)$ for the expected value of distances, that is, x_{STA} and x_{TA}, (19) transforms to the following form:

$$\left\| \sum_{j=\lceil k/2 \rceil + 1}^{k} \binom{k}{j} \left(e^{-\beta x_{STA}} \right)^j \left(1 - e^{-\beta x_{STA}} \right)^{k-j} \right.$$

$$\left. \times \sum_{j=\lceil k/2 \rceil + 1}^{k} \binom{k}{j} \left(e^{-\beta x_{TA}} \right)^j \left(1 - e^{-\beta x_{TA}} \right)^{k-j} \right\| < \epsilon. \tag{25}$$

Further, we assume that the central sensor always detects the event

$$\sum_{j=\lceil k/2 \rceil}^{k-1} \binom{k-1}{j}$$

$$\times \left\| e^{-\beta j r_c \sqrt{1/3 + ((k-1)/k)^2}} \left(1 - e^{-\beta r_c \sqrt{1/3 + ((k-1)/k)^2}} \right)^{k-j} \right. \tag{26}$$

$$\left. - e^{-j\beta((C \cdot r_c)/4)} \left(1 - e^{-\beta((C \cdot r_c)/4)} \right)^{k-j} \right\| < \epsilon.$$

Solving the above equation using Taylor expansion,

$$\sum_{m=1}^{\infty} \underbrace{\left\| (-1)^m \right\| \sum_{j=\lceil k/2 \rceil}^{k-1} \binom{k-1}{j} \sum_{i=1}^{m} \frac{j^{m-i} (k-j-1)^i}{i! (m-i)!} \beta^m}_{\lambda_m}$$

$$\times \left\| r_c^m \left(\frac{1}{3} + \left(\frac{k-1}{k} \right)^2 \right)^{m/2} - \left(\frac{C \cdot r_c}{4} \right)^m \right\| < \epsilon. \tag{27}$$

Bounding the above summation by a finite number of terms, that is, l, we obtain the following expression for C:

$$C = 4 \sqrt[l]{\left(\frac{1}{3} + \left(\frac{k-1}{k} \right)^2 \right)^{l/2} - \epsilon}. \tag{28}$$

Case II (all the sensors detect the event). In this case, a sensor present at the maximum distance, that is, $2r_c$, from the event must detect the event.

Hence, the range of r_s for event detection is

$$\Omega \leq r_s \leq 2r_c. \tag{31}$$

$\square$

Proposition 2. *Let r_s and r_c be the sensing and communication ranges of each of the k sensors deployed in a region R' at*

The above equation gives an expression of distance ratio (C). Thus, the events would be detected with approximately the same detection rates for both STA-QS-SVM and TA-QS-SVM approaches if $r'_c = C \cdot r_c$. A plot of C for different values of k is shown in Figure 2. The figure shows that C increases with an increase in the corresponding value of k.

3.1. Maximum Sensing Ranges r_s for STA-QS-SVM and TA-QS-SVM (SensGru). Let r_s and r_c be the sensing and communication ranges of each of the k sensors deployed in a region R consisting of a central sensor at a distance of r_c from uniformly deployed sensors at the border of the region for the STA-QS-SVM approach. Then, the maximum value of r_s required for event detection is $2r_c$.

Proof. Given that region R consists of exactly k sensors, $k-1$ of which are deployed uniformly along the circumference of the region and are at a distance r_c apart from the central sensor; event detection occurs when at least $\lceil k/2 \rceil + 1$ sensors detect it. In the worst case, an event occurring at the border of region R should be detected. Two cases arise.

Case I ($\lceil k/2 \rceil + 1$ sensors detect the event). In this case, the $\lceil k/2 \rceil + 1$ nearest sensors detect the event. Thus, a sensor present at the distance $\sqrt{2} r_c \sqrt{1 - \cos(\lceil (k-1)/2 \rceil \pi/(k-1))}$ should be able to detect it. Moreover, from assumption (7) and (27) and using finite number of l terms to bound the summation, we have

$$\lambda_1 \beta x_{STA} + \lambda_2 \beta^2 x_{STA}^2 + \cdots + \lambda_l \beta^l x_{STA}^l \geq 1 - \delta_{STA}. \tag{29}$$

Let κ be the real root of this equation; then, using the above relation between the minimum possible r_s, k, and r_c, we get $r_s \geq \kappa \sqrt{2} \sqrt{1 - \cos(\lceil (k-1)/2 \rceil \pi/(k-1))} / 2\beta \sqrt[l]{(1/3 + ((k-1)/k)^2)^{l/2} - \epsilon}$. Hence,

$$r_s \geq \underbrace{\max \left(r_c, \frac{\kappa}{2\beta \sqrt[l]{\left(1/3 + ((k-1)/k)^2 \right)^{l/2} - \epsilon}} \right) \sqrt{2} \sqrt{1 - \cos \left(\frac{\lceil (k-1)/2 \rceil \pi}{(k-1)} \right)}}_{\Omega}. \tag{30}$$

a central node for SensGru. Each node is at a distance of r'_c from nodes in the neighboring regions. Then, the maximum value of r_s required for event detection is $(C \cdot r_c)/2$.

Proof. Given that the region R' consists of exactly k sensors at a central node and each node is at a distance of $C \cdot r_c$ from other neighboring nodes, event detection occurs when at least $\lceil k/2 \rceil + 1$ sensors detect it. In the worst case, an event occurring at the border of the region R' should be
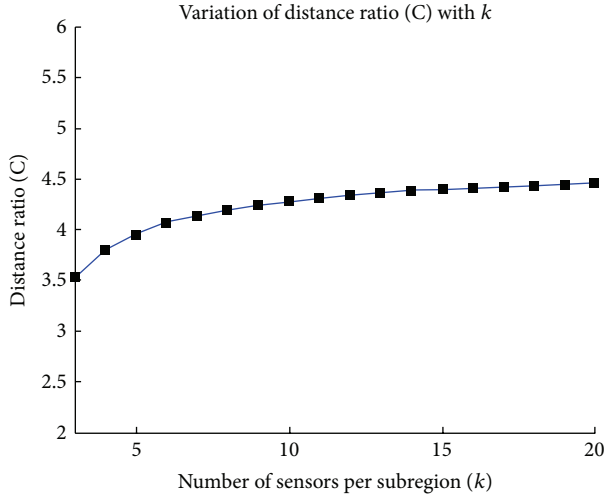
Figure 2: Variation of distance ratio (C) with number of sensors (k).

Figure 3: Variation of maximum sensing range with number of sensors (k).

detected. Since the border of the region R' is at the distance of $(C \cdot r_c)/2 = 2r_c\sqrt[l]{(1/3 + ((k-1)/k)^2)^{l/2}} - \epsilon$ from the nodes, the event at the border will be detected if all the sensors have

$$r_s = 2r_c\sqrt[l]{\left(\frac{1}{3} + \left(\frac{k-1}{k}\right)^2\right)^{l/2}} - \epsilon. \tag{32}$$

$\square$

A plot of maximum r_s for various values of k and $\epsilon = 0.001$ for STA-QS-SVM and SensGru (TA-QS-SVM) approaches is shown in Figure 3. The figure indicates that the maximum value of r_s required for SensGru increases with an increase in the number of sensors per subregion k. However, as the value of k increases, the increase in the maximum r_s reduces.

Proposition 3. *The optimal value of k, that is, k_{opt}, for event detection in STA-QS-SVM and SensGru is 7.*

Proof. The optimal value of k is defined as the value for which maximum reduction in computation complexity and sensor node density is attained in SensGru, given that the event detection performance with both schemes is approximately the same. The maximum reduction in computation complexity and cost is attained when the sensors deployed in both regions R and R' have the same maximum sensing ranges r_s (as shown in Figure 3). Thus, equating the maximum r_s values for STA-QS-SVM and SensGru, we get

$$2r_c\sqrt[l]{\left(\frac{1}{3} + \left(\frac{k_{opt} - 1}{k_{opt}}\right)^2\right)^{l/2}} - \epsilon = 2r_c,$$

$$k_{opt} = \frac{1}{1 - \sqrt{(1+\epsilon)^{2/l} - 1/3}}. \tag{33}$$

$\square$

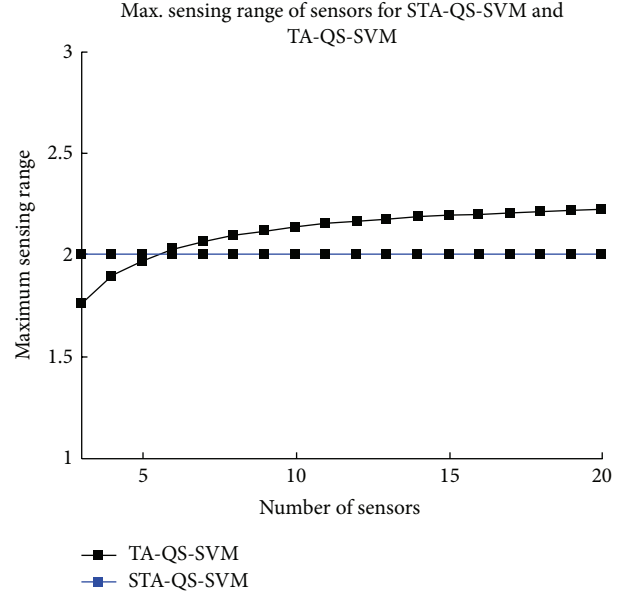The above equation indicates that the optimal value of k, that is, k_{opt}, depends only on ϵ and l. Hence, as evident from the figures, the optimal value $k_{opt} = 5.5$. Thus, for experimental purposes, $k_{opt} = 6$ can be chosen. However, since we prove in the following that the detection probabilities for even number of sensors are less as compared to odd number of sensors, we choose $k_{opt} = 7$.

3.2. Worst Case Analysis

Proposition 4. *For the worst case analysis,*

$$\{p(E_{STA} \mid E_a)\}|_{k=odd} > \{p(E_{STA} \mid E_a)\}|_{k=even}. \tag{34}$$

Proof. The worst case analysis is made assuming that the event is located at the boundary of region R' for the SensGru and in between the two sensors present along the boundary of R for the STA-QS-SVM approach. The worst case probabilities of event detection in regions R are greater for odd number of sensors as compared to the even number of sensors. When the number of sensors are even, one of the k sensors will always be located at maximum distance, that is, $2r_c$ from the event location, whereas when the number of sensors is odd, no sensor is located at maximum distance from the event location; hence, the probability of detection increases. This is shown in Figures 4 and 5 for the STA-QS-SVM approach. $\square$

Proposition 5. *The worst case distance ratio (C_{worst}) is equal to 2.36.*

Proof. The worst case distance ratio C_{worst} was evaluated assuming that the event is located at the boundary of region R' for SensGru and in between the two sensors present along the boundary of R for the STA-QS-SVM approach. The optimal value k_{opt} was assumed for R and R'. The worst
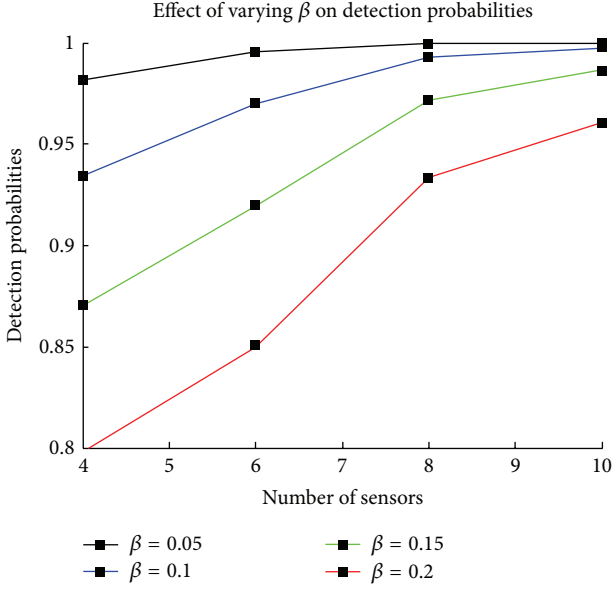
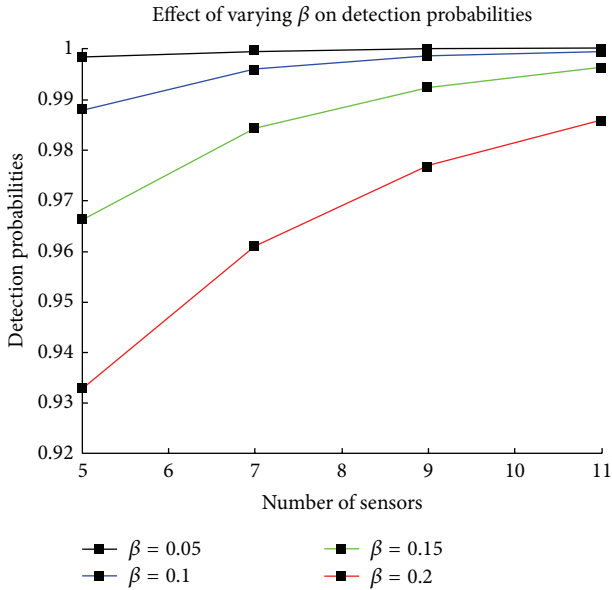FIGURE 4: Worst case probability of detection for STA-QS-SVM for even number of sensors (k).

FIGURE 6: Variation of C_{worst} with β.

in R). Thus, if this condition is satisfied, $p(E_{STA} \mid E_a)$ and $p(E_{TA} \mid E_a)$ will differ by ϵ. $\square$

4. Complexity Analysis

Let n be the total number of data measurements at a node and let d be the number of data attributes. Let k be the number of sensors in the region R for STA-QS-SVM approach and region R' for SensGru approach.

4.1. Computation Complexity

4.1.1. ST-QS-SVM. ST-QS-SVM requires the computation of $n \times n$ kernel matrix K which poses a complexity of $O(n^2)$. The radius is computed at each node via a linear optimization problem. Hence, the radius computation poses a complexity of $O(n)$. Each node then broadcasts its radius information to the central node. On the receipt of radius information, the central node computes the global radius in $O(k)$ complexity (k being the number of nodes). Hence, the total computation complexity is $O(n^2)$.

4.1.2. STA-QS-SVM. Online STA-QS-SVM requires the computation of $n \times n$ kernel matrix K and $d \times d$ matrix K^* which poses a complexity of $O(n^2 + nd^2)$. The radii (temporal and attribute) are computed at each node via a linear optimization problem. Hence, the radius computation poses a complexity of $O(n+nd)$. Each node then broadcasts its radius information to the central node. On the receipt of radius information, the central node computes the global radius in $O(k)$ complexity (k being the number of nodes). Hence, the total computation complexity is $O(n^2 + nd^2)$. In offline STA-QS-SVM, the matrix K^* is computed at the beginning only; hence, the computation complexity is $O(n^2)$.

FIGURE 5: Worst case probability of detection for STA-QS-SVM for odd number of sensors (k).

case distance ratio C_{worst} was found such that probabilities of detection $p(E_{STA} \mid E_a)$ and $p(E_{TA} \mid E_a)$ differ by ϵ. $p(E_{STA} \mid E_a)$ and $p(E_{TA} \mid E_a)$ were evaluated for various values of β. An example plot of worst case distance ratio for various values of β is shown in Figure 6. The figure shows that the worst case distance ratio C_{worst} does not vary significantly with β. The results shown in Figure 6 also prove that the worst case distance ratio is equal to 2.36. The value of C_{worst} = 2.36 implies that, for the worst case scenario, where the event is located at the boundary of regions R and R', distance between sensor nodes in R' = 2.36 × (distance between sensor nodes
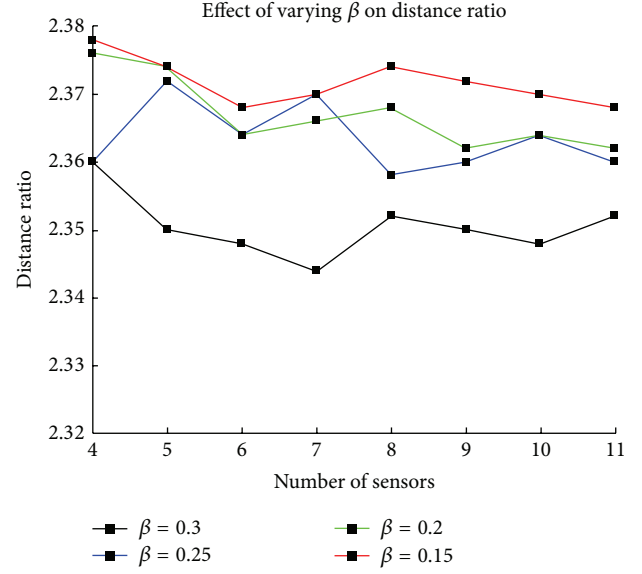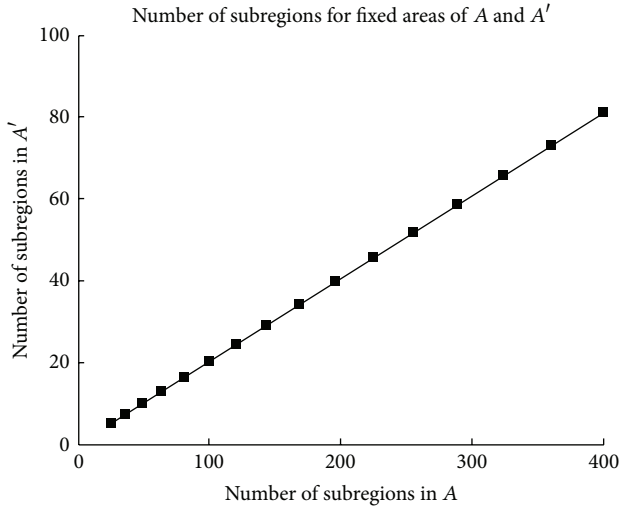
Number of subregions for fixed areas of A and A'

FIGURE 7: Number of subregions in A and A' for a fixed area.

Total number of sensors

- ■ STA-QS-SVM
- ■ TA-QS-SVM

FIGURE 8: Variation of the total number of sensors in A and A' with the parameter b.

4.1.3. SensGru (TA-QS-SVM). SensGru requires the computation of $n \times n$ kernel matrix K and $d \times d$ matrix K^* which poses a complexity of $O(n^2 + nd^2)$. The radii (temporal and attribute) are computed at each sensor via a linear optimization problem. Hence, the radius computation poses a complexity of $O(n + nd)$. Each sensor then sends its radius information to the central sensor. On the receipt of radius information, the central sensor computes the global radius in $O(k)$ complexity (k being the number of nodes). Hence, the total computation complexity is $O(n^2 + nd^2)$. Note that SensGru can have online and offline versions.
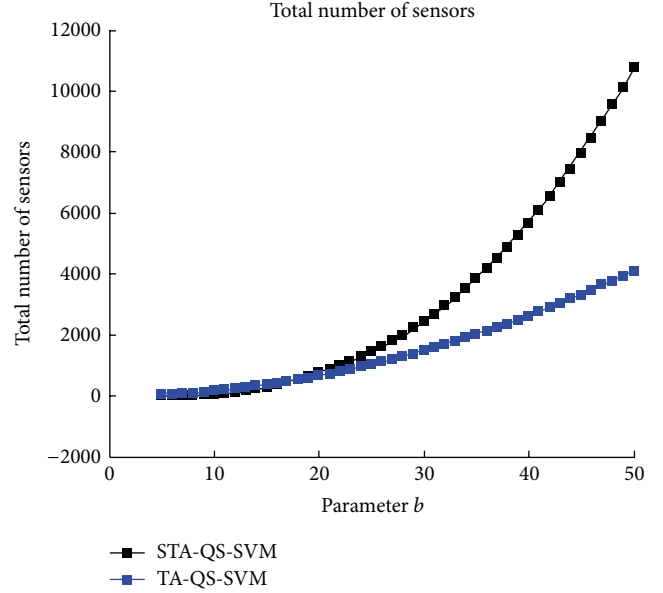
4.2. Communication Complexity

4.2.1. ST-QS-SVM. ST-QS-SVM requires the broadcast of temporal radius to the central node which poses a communication complexity of $O(n)$.

4.2.2. STA-QS-SVM. STA-QS-SVM requires the broadcast of the temporal and the attribute radii to the central node which poses a communication complexity of $O(n + d)$.

4.2.3. TA-QS-SVM. TA-QS-SVM poses no communication complexity as it does not require the broadcast of radii to the central sensor.

4.3. Sensor Node Density. Consider regions A (for STA-QS-SVM) and A' (for SensGru) having the same physical characteristics (the same values of β). Assume that region $A = b^2 \cdot R$ and region $A' = c^2 \cdot R'$. Thus, A and A' (of the same area) contain b^2 and c^2 number of closely packed subregions R and R'. Further, we assume that the subregions R and R' are hexagon shaped, since they can then be closely packed to form A and A'. Let each of the subregions R and R' have radii r_c and r'_c; then, the radii of A and A' are $b \cdot r_c$ and $(c \cdot r'_c)/2$ approximately. Further, assume that each of the subregions R in A has k number of sensors, of which one is the central and

$k - 1$ are located uniformly along the boundary of R. Each of the subregions R will then have some sensors in common with the neighboring subregions. Also let the subregions R' in A' have a single node located at the center and have k sensors on it. Thus, the number of sensors in region A will be given as

$$k + \sum_{i=1}^{b/2} \left(\frac{k}{2} \left(\frac{i^2 + 7i}{2} - 3 \right) \right) = \frac{kb^3}{96} + \frac{kb^2}{32} - \frac{71kb}{48}. \quad (35)$$

Since we want to evaluate the number of sensors for equal areas of A and A', equating the two areas and using (28), we get the number of sensors in A':

$$N\left(A'\right) = \left(\frac{b}{2\sqrt[l]{\left(1/3 + ((k-1)/k)^2\right)^{l/2}} - \epsilon} \right)^2 k. \quad (36)$$

Figure 7 shows the number of subregions required for the same area of A and A'. The figure clearly shows that the number of subregions R' in A' is less as compared to the number of regions R in A. Figure 8 shows the variation of total number of sensors in regions A and A' with the parameter b of the regions. Clearly the number of sensors for region A is much greater as compared to the number of sensors for region A' for larger values of b. The results also indicate that with an increase in the parameter b, the increase in $N(A)$ will be much greater as compared to $N(A')$. Similar results are shown in Figure 9, which gives the total number of sensors in regions A and A' for different values of k. The number of sensors $N(A)$ and $N(A')$ increases linearly with an increase in k. Thus, the sensor density required in TA-QS-SVM approach is significantly less as compared to that required for the STA-QS-SVM approach.
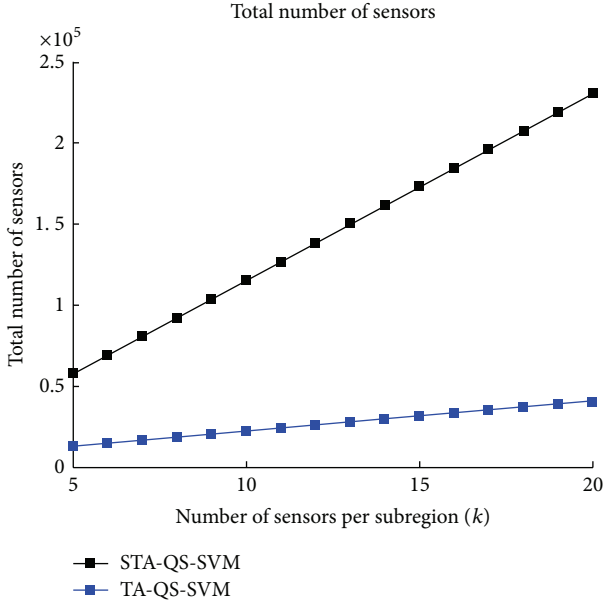
FIGURE 9: Variation of the total number of sensors in A and A' with the number of sensors per subregion k.

5. Simulations and Results

Experimental evaluation was performed on matlab with a synthetic data set that consists of a mixture of Type-1, Type-2, Type-3, and Type-4 outliers. Type-1 and Type-2 outliers are isolated errors which may occur due to noise at random time instants in the data set, whereas Type-3 and Type-4 outliers consist of sequences of Type-1 and Type-2 outliers [10, 28–30]. The data set consists of a mixture of various Gaussian distributions, with mean randomly selected from (0.03–0.6) and standard deviation 0.03, along each of the attributes. The number of Gaussian distributions depends on the number of attributes. The testing data set consisted of 2500 measurements. The training window n was kept equal to 250. Performance evaluation was done for data vectors with 2 to 15 attributes (A), with 5% of uniformly distributed outliers in the range [0.65, 1] along each of the attributes. RBF Kernel function is used to generate the results with sigma 0.2, because it has infinite VC dimension. The simulation environment assumed a central node surrounded by 6–9 (N or k) spatially correlated neighboring nodes. The distance between the nodes in STA-QS-SVM is kept equal to the maximum r_s as given in (32). Further, all the assumptions have been taken into account as described in the above sections.

The comparison of event detection rates for STA-QS-SVM and SensGru is presented in Figures 10 and 11. Assuming that the probability of all the sensors being faulty at the same time instant at a node S_i is negligible and that the spatial correlations of erroneous measurements indicate an event, our technique identifies a measurement as an event if it has temporal-attribute deviation with itself and a strong correlation with a majority of sensors deployed at the node.
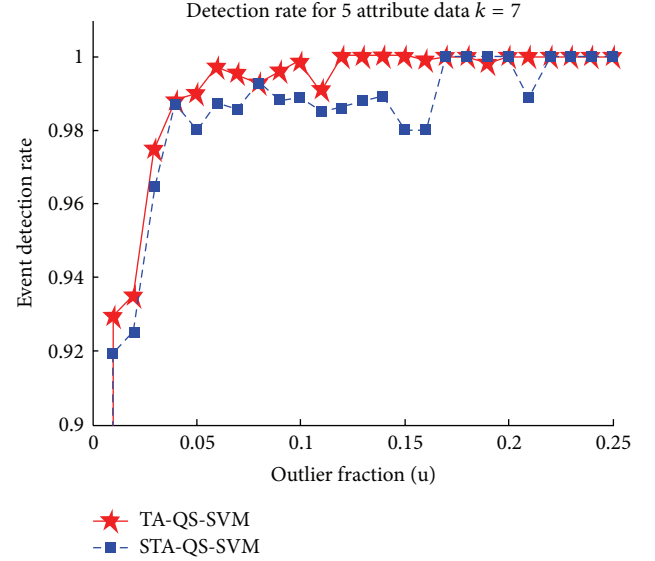

FIGURE 10: Event detection rates for STA-QS-SVM and TA-QS-SVM (nodes = 7).
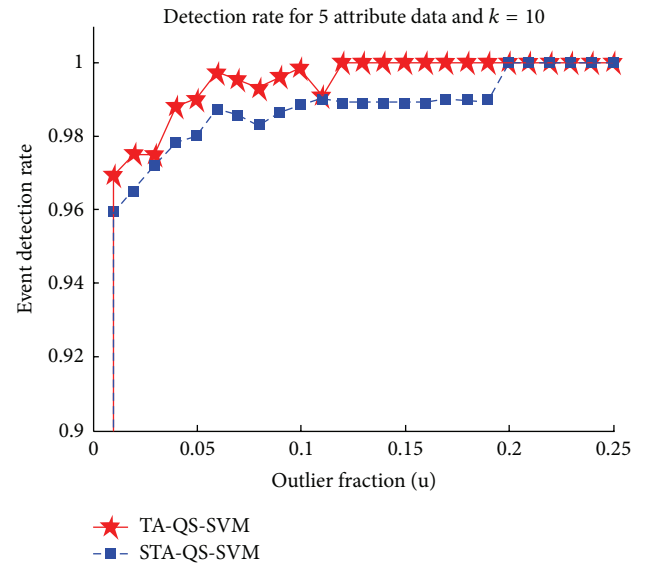

FIGURE 11: Event detection rates for STA-QS-SVM and TA-QS-SVM (nodes = 10).

5.1. Discussion on Results. Figures 10 and 11 show that SensGru always gives equal or better performance than STA-QS-SVM at a small cost of slightly increasing the sensing range r_s and increasing the number of sensors per node. Communication overhead gets totally eliminated and node density gets greatly reduced. Since the spatial correlations are only used to determine an event and temporal-attribute correlations determine outliers, TA-QS-SVM based SensGru exploits the temporal and attribute correlations to identify outliers and invoke a consensus of all the k sensor arrays to identify an event, without using spatial correlations of neighboring nodes. The consensus of k sensor arrays at a node S_i is thus equivalent to the spatial information of k

neighboring nodes of S_i in a region and thus makes the outlier and event detection process free of communication cost and highly energy efficient. Since increasing the sensing range is cost effective as compared to the communication cost, SensGru is much more communicationally and computationally effective as compared to STA-QS-SVM approach. However, there is a trade-off associated with the increase in the sensing range and detection probabilities of event in STA-QS-SVM and TA-QS-SVM based SensGru. The trade-off can be explained in terms of Figures 4, 5, and 3. The figures show that on choosing the optimal value of $k = k_{\mathrm{opt}} = 7$ the sensing range for STA-QS-SVM and SensGru will be r_s. However, the detection probabilities for both the approaches (Figures 4 and 5) are less than ideal 100%. If we increase the number of sensors k, the sensing range of sensors for STA-QS-SVM can be maintained equal to r_s, and an increase in $p(E_{\mathrm{STA}} \mid E_a)$ is obtained. However, in order to obtain the corresponding increase in $p(E_{\mathrm{TA}} \mid E_a)$, an increase in the sensing range of sensor is required. However, reduction in sensor node density outweighs this increase in the sensing range and the overall computation cost is reduced. A more involved reader might come across the thought that using the optimal value of $k = k_{\mathrm{opt}}$ for both STA-QS-SVM and SensGru, the sensor node density will become the same. However, as evident from (28), when the sensing ranges are the same, the distance between the nodes in SensGru will still be twice that of the distance between the nodes in STA-QS-SVM. Hence, the sensor density will still increase quadratically with b for SensGru and cubically with b for STA-QS-SVM.

6. Conclusion

This paper proposes an energy efficient event detection and identification approach for cyber physical systems. A thorough analysis on the impact of grouping multiple sensors on a single node has been conducted where each node runs an event detection and identification algorithm based on quarter-sphere SVM formulation. Huge power savings are guaranteed in the proposed scheme as it gets rid of the communication overhead before declaring an event with a node density reduced by 75%. Furthermore, all these gains are achieved while maintaining the same performance as power inefficient STA-QS-SVM formulation.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] Y. Zhang, N. Meratnia, and P. Havinga, "Outlier detection techniques for wireless sensor networks: a survey," *IEEE Communications Surveys and Tutorials*, vol. 12, no. 2, pp. 159–170, 2010.

[2] C. F. Garca-Hernndez, P. H. Ibargengoytia-Gonzlez, J. Garca-Hernndez, and J. A. Prez-Daz, "Wireless sensor networks and applications: a survey," *IJCSNS International Journal of Computer Science and Network Security*, vol. 7, no. 3, pp. 264–273, 2007.

[3] R. Cardell-Olivera, M. Kranza, K. Smettemb, and K. Mayerc, "A reactive soil moisture sensor network: design and field evaluation," *International Journal of Distributed Sensor Networks*, vol. 1, no. 2, pp. 149–162, 2005.

[4] P. Misra, S. Kanhere, D. Ostry, and S. Jha, "Safety assurance and rescue communication systems in high-stress environments: a mining case study," *IEEE Communications Magazine*, vol. 48, no. 4, pp. 66–73, 2010.

[5] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," *Ad Hoc Networks*, vol. 3, no. 3, pp. 257–279, 2005.

[6] I. F. Akyildiz, Ö. B. Akan, C. Chen, J. Fang, and W. Su, "Inter-planetary internet: state-of-the-art and research challenges," *Computer Networks*, vol. 43, no. 2, pp. 75–112, 2003.

[7] P. Tan, M. Steinback, and V. Kumar, *Introduction to Data Mining*, Addison-Wesley, New York, NY, USA, 2006.

[8] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: a survey," *ACM Computing Surveys*, vol. 41, no. 3, article 15, 2009.

[9] G. H. John, "Robust decision trees: removing outliers from databases," in *In Knowledge Discovery and Data Mining*, pp. 174–179, AAAI Press, 1995.

[10] S. Rajasegarar, J. C. Bezdek, C. Leckie, and M. Palaniswami, "Elliptical anomalies in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 6, no. 1, article 7, 2009.

[11] M. Zoumboulakis and G. Roussos, "Escalation: Complex event detection in wireless sensor networks," in *Smart Sensing and Context*, pp. 270–285, 2007.

[12] B. Krishnamachari and S. Iyengar, "Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks," *IEEE Transactions on Computers*, vol. 53, no. 3, pp. 241–250, 2004.

[13] M. Ding, D. Chen, K. Xing, and X. Cheng, "Localized fault-tolerant event boundary detection in sensor networks," in *Proceedings of the IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '05)*, vol. 2, pp. 902–913, March 2005.

[14] M. Ding and X. Cheng, "Robust event boundary detection in sensor. Networks—a mixture model based approach," in *Proceedings of the 28th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '09)*, pp. 2991–2995, IEEE, April 2009.

[15] N. Shahid, I. Naqvi, and S. Qaisar, "Characteristics and classification of outlier detection techniques for wireless sensor networks in harsh environments: a survey," *Artificial Intelligence Review*, 2012.

[16] V. J. Hodge and J. Austin, "A survey of outlier detection methodologies," *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, 2004.

[17] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[18] V. Gomez-Verdejo, J. Arenas-Garcia, M. Lazaro-Gredilla, and A. Navia-Vazquez, "Adaptive one-class support vector machine," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2975–2981, 2011.

[19] D. M. J. Tax and R. P. W. Duin, "Data domain description using support vectors," in *Proceedings of the European Symposium on Artificial Neural Networks (ESANN '99)*, pp. 251–256, Bruges, Belgium, April 1999.

[20] P. Laskov, C. Schfer, and I. Kotenko, "Intrusion detection in unlabeled data with quarter-sphere support vector machines," in *Proceedings of the Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA '04)*, pp. 71–82, 2004.

[21] D. Wang, D. S. Yeung, and E. C. C. Tsang, "Structured one-class classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 36, no. 6, pp. 1283–1295, 2006.

[22] S. Rajasegarar, C. Leckie, and M. Palaniswami, "CESVM: centered hyperellipsoidal support vector machine based anomaly detection," in *Proceedings of the IEEE International Conference on Communications (ICC '08)*, pp. 1610–1614, May 2008.

[23] S. Rajasegarar, A. Shilton, C. Leckie, R. Kotagiri, and M. Palaniswami, "Distributed training of multiclass conic-segmentation support vector machines on communication constrained networks," in *Proceedings of the 6th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP '10)*, pp. 211–216, December 2010.

[24] N. Shahid, I. H. Naqvi, and S. B. Qaisar, "One-class support vector machines: analysis of outlier detection for wireless sensor networks in harsh environments," *Artificial Intelligence Review*, pp. 1–45, 2013.

[25] Y. Zhang, N. Meratnia, and P. J. M. Havinga, "Hyperellipsoidal SVM-based outlier detection technique for geosensor networks," in *GeoSensor Networks*, vol. 5659 of *Lecture Notes in Computer Science*, pp. 31–41, Springer, Berlin, Germany, 2009.

[26] S. Rajasegarar, C. Leckie, J. C. Bezdek, and M. Palaniswami, "Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 518–533, 2010.

[27] N. Shahid, I. H. Naqvi, and S. B. Qaisar, "Quarter-Sphere SVM: attribute and Spatio-Temporal correlations based outlier & event detection in wireless sensor networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference: Mobile and Wireless Networks (WCNC '12)*, pp. 2048–2053, Shanghai, China, April 2012.

[28] Y. Zhang, *Observing the unobservable—distributed online outlier detection in wireless sensor networks [Ph.D. thesis]*, University of Twente, June 2010.

[29] J. C. Bezdek, T. C. Havens, J. M. Keller et al., "Clustering elliptical anomalies in sensor networks," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ '10)*, pp. 1–8, Barcelona, Spain, July 2010.

[30] S. Suthaharan, M. Alzahrani, S. Rajasegarar, C. Leckie, and M. Palaniswami, "Labelled data collection for anomaly detection in wireless sensor networks," in *Proceedings of the 6th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP '10)*, pp. 269–274, Brisbane, Australia, December 2010.