Ad Hoc Networks xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Ad Hoc Networks

journal homepage: www.elsevier.com/locate/adhoc

Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine

Yang Zhang *, Nirvana Meratnia, Paul J.M. Havinga

Pervasive Systems Group, Department of Computer Science, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 29 September 2011
Received in revised form 3 October 2012
Accepted 7 November 2012
Available online xxxx

Keywords:
Outlier detection
Ellipsoidal support vector machine
Spatial correlation
Temporal correlation
Wireless sensor networks

ABSTRACT

Low quality sensor data limits WSN capabilities for providing reliable real-time situation-awareness. Outlier detection is a solution to ensure the quality of sensor data. An effective and efficient outlier detection technique for WSNs not only identifies outliers in a distributed and online manner with high detection accuracy and low false alarm, but also satisfies WSN constraints in terms of communication, computational and memory complexity. In this paper, we take into account the correlation between sensor data attributes and propose two distributed and online outlier detection techniques based on a hyperellipsoidal one-class support vector machine (SVM). We also take advantage of the theory of spatio-temporal correlation to identify outliers and update the ellipsoidal SVM-based model representing the changed normal behavior of sensor data for further outlier identification. Simulation results show that our adaptive ellipsoidal SVM-based outlier detection technique achieves better detection accuracy and lower false alarm as compared to existing SVM-based techniques designed for WSNs.

1. Introduction

With the increasing advances of digital electronics and wireless communications, in the past decade a new breed of tiny embedded systems known as *wireless sensor nodes* has emerged. These wireless sensor nodes are equipped with sensing, processing, wireless communication, and more recently actuation capability. They usually are densely deployed in a wide geographical area and continuously measure various parameters (e.g. ambient temperature, relative humidity, soil moisture, wind speed) of the physical world. A large collection of these sensor nodes forms a *wireless sensor network* (WSN) [1].

Event-driven WSN applications [2,3] require timely data analysis and assessment in order to facilitate (near) real-time, efficient, and accurate critical decision making and situation awareness. Accurate sensor data analysis and decision making process rely heavily on the quality of sensor data as well as additional information and context. However, raw sensor observations collected from sensor nodes often have low data quality and reliability due to the limited capability of sensor nodes in terms of energy, memory, computational power, bandwidth, dynamic nature of network, and harshness of the deployment environment. Use of low quality sensor data in any data analysis and decision making process limits the possibilities for reliable real-time situation-awareness.

A solution to ensure the quality of sensor data is detection of outliers. In the context of WSNs, outliers are defined as *those sensor observations that do not conform to the defined (expected) normal behavior of sensor data* [4]. This definition indicates that a straightforward way for outlier detection in WSNs is to define a normal behavior of sensor data and consider those sensor observations that deviate from the defined normal behavior of sensor data as outliers. An effective and efficient outlier detection technique for WSNs should be able to identify outliers in a distributed and online manner with high detection accuracy and low false alarm, while satisfying WSN constraints in terms of communication, computational and memory complexity [5].

* Corresponding author.
E-mail address: zhangy@cs.utwente.nl (Y. Zhang).

1570-8705/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.adhoc.2012.11.001

A commonly used method to model the normal behavior of given data vectors in data mining and machine learning fields is *similarity measure* [18]. The majority of data vectors with high similarity represents the normal behavior of the given data vectors. Use of such similarity measure avoids making any assumption on statistical distribution of data as well as saves on expensive computational and memory complexity of data distribution estimation. For multivariate sensor data, modeling the normal behavior of data precisely needs taking sensor data correlations into account [26]. In reality sensor data attributes are often correlated, e.g., ambient temperature has certain correlation with relative humidity. Observations indicate that when the air is warmer it can hold more humidity [4]. In this way, the correlated sensor data vectors are not distributed around the center of mass in a spherical manner, instead their distribution has an *ellipsoidal* shape [6]. The direction of the formed ellipsoid reveals the multivariate nature of data distribution trend as well as the strength of the correlation between data attributes.

In this paper, we simplify an ellipsoidal one-class SVM [7] to model the normal behavior of sensor data attributes in resource-constrained WSNs and propose two ellipsoidal one-class SVM-based outlier detection techniques to identify outliers in a distributed and online manner. In the process of detecting outliers, these two techniques also take advantage of the theory of spatio-temporal correlation to precisely detect outliers and changes of the normal behavior of sensor data. Furthermore, the proposed adaptive technique enables to update the ellipsoidal SVM-based model to represent changes of the normal behavior of sensor data for further outlier identification. Simulations with two synthetic datasets and one real environmental dataset from the Grand St. Bernard [8] show that our adaptive outlier detection technique achieves better detection accuracy and lower false alarm as compared to existing SVM-based outlier detection techniques [9,10,23,22] designed for WSNs.

The remainder of this paper is organized as follows. Related work on developing one-class SVM-based outlier detection in WSNs as well as in data mining and machine learning fields is described in Section 2. Principles of modeling the ellipsoidal one-class SVM classifier are addressed in Section 3. How to lower down the complexity of traditional hyperellipsoidal one-class SVM classifier to fit requirements of WSNs is addressed in Section 4. Our proposed two ellipsoidal SVM-based outlier detection techniques are presented in Section 5. Simulation results and performance evaluation of our techniques with other existing SVM-based outlier detection techniques are reported in Sections 6 and 7. Finally this paper is concluded in Section 8 with plans for future research.

2. Related work

SVM-based techniques are original from the family of classification-based techniques in data mining and machine learning fields. The main idea of classification-based techniques is to learn a classifier using data vectors in the training phase and classify an unseen instance into one of the learned classes in the testing phase. SVM-based techniques specifically separate the data vectors belonging to different classes by fitting a hyperplane, which produces a maximal margin in a high-dimensional data space. SVM-based techniques are commonly used for the purpose of outlier detection due to the fact that they have three main attracting advantages, i.e., (i) do not require an explicit statistical model and complex parameter estimation, (ii) use an optimum solution to produce a more reliable normal boundary to precisely distinguish between normal data and outliers, and (iii) avoid the curse of data dimensionality problem for computing the similarity measure among data vectors.

However, traditional SVM-based outlier detection techniques suffer from two disadvantages: (i) they require error-free or labeled data for training, and (ii) they require a computationally expensive quadratic optimization. One-class (unsupervised) SVM-based techniques can solve the first disadvantage as they can model the normal behavior of the unlabeled data while ignoring the anomalies existing in the training set. Their main idea is to use a non-linear function to map the data vectors collected from the original *input space* to a higher dimensional space called *feature space*. Then a decision boundary of normal data is found, which encompasses the majority of data vectors in the feature space. Those data vectors falling outside the normal boundary are classified as outliers. To this end, Scholkopf et al. [11] have proposed a hyperplane-based one-class SVM, which identifies outliers by fitting a hyperplane from the origin. Those data vectors near the origin are declared as outliers. Tax and Duin [12] have proposed a hypersphere-based one-class SVM, which identifies outliers by fitting a hypersphere with a minimum radius. Those data vectors falling outside the hypersphere are declared as outliers. Wang et al. [7] have proposed a hyperellipsoid-based one-class SVM, which identifies outliers by fitting multiple hyperellipsoids with minimum effective radii. Those data vectors falling outside the hyperellipsoids are declared as outliers. However, these one-class SVM-based techniques still require a computationally expensive quadratic optimization.

In order to reduce high computational cost of the quadratic optimization, Campbell and Bennett [13] have formulated a linear programming approach for the hyperplane-based SVM proposed in [11], which is based on attracting the hyperplane towards the average of the distribution of mapped data vectors. Laskov et al. [14] have extended work in Tax and Duin [12] by proposing a quarter-sphere one-class SVM, which converts the quadratic optimization problem to a linear optimization problem by fitting a hypersphere centered at the origin, and consequently reduces computational complexity of learning the normal boundary of data vectors.

Rajasegarar et al. [9] use the quarter-sphere one-class SVM proposed in [14] to present a distributed outlier detection technique for WSNs. In their technique, each node analyzes sensor data in an offline manner only after all observations are collected within a day, which obviously causes a considerable outlier detection delay. This is not suitable for detecting outliers in critical real-time applications of WSNs. Rajasegarar et al. [10] have further

Please cite this article in press as: Y. Zhang et al., Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.11.001

extended work in [7,14] by proposing a hyperellipsoidal one-class SVM using a linear optimization. However, this technique is neither distributed nor online. It only operates in a single location in an offline manner.

In this paper, we propose two distributed and online outlier detection techniques based on a simplified hyperellipsoidal one-class SVM using a linear optimization. Our techniques enable to identify outliers online and adapt to changes of the normal behavior of sensor data in real-time.

3. Principles of modeling hyper-ellipsoid one-class SVM

In this section, we describe the principles of modeling the hyperellipsoidal one-class SVM proposed in [7] for multivariate data vectors and further compare the differences between hyper-ellipsoid SVM and hyper-sphere SVM.

3.1. Modeling hyper-ellipsoid one-class SVM

The quadric optimization problem of modeling the hyperellipsoidal SVM classifier has been converted to the linear optimization problem in [10] by fixing the center of mapped data vectors in the feature space at the origin. The geometry of hyperellipsoidal one-class SVM-based approach is shown in Fig. 1. The general process of modeling the hyperellipsoidal SVM classifier for multivariate data vectors is addressed below.

Assume that m data vectors $\{x_i \epsilon \Re^d, i = 1, \ldots, m\}$ of d variables in the input space are mapped into the feature space using some non-linear mapping function ϕ. The hyperellipsoidal SVM aims at enclosing the majority of mapped data vectors $\phi(x_i)$ in the feature space by fitting a hyperellipsoid centered at the origin with a minimum effective radius R. Thus, the optimization problem in this hyperellipsoidal SVM classifier is represented as:

$$\min_{R\epsilon\Re,\xi\epsilon\Re^m} R^2 + \frac{1}{\upsilon m}\sum_{i=1}^{m}\xi_i \tag{1}$$

$$\textit{subject to}: \phi(x_i)\Sigma^{-1}\phi(x_i)^T \leqslant R^2 + \xi_i, \xi_i \geqslant 0, \ i = 1, 2, \ldots m$$

where $\upsilon \epsilon (0,1)$ is a parameter that controls the fraction of mapped data vectors that can be outliers. The slack variables $\{\xi_i : i = 1, 2, \ldots, m\}$ allow some of mapped data vectors

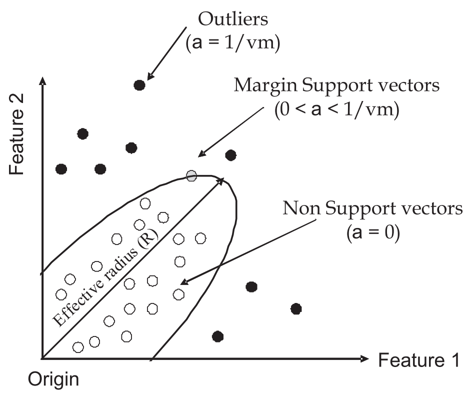


Fig. 1. Geometry of the hyper-ellipsoidal formulation of one-class SVM [10].

to lay outside the hyperellipsoid. Σ^{-1} is the inverse of the *covariance matrix* Σ of mapped data vectors, which is computed as follows:

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m}(\phi(x_i) - \mu)(\phi(x_i) - \mu)^T, \quad \mu = \frac{1}{m}\sum_{i=1}^{m}\phi(x_i) \tag{2}$$

Using Mercer Kernels [15], the inner products of mapped data vectors in the feature space can be computed in the input data space without needing any knowledge about the non-linear function ϕ. Let $K \epsilon \Re^{m\times m}$ be the kernel matrix of the original data vectors. Mapped data vectors can be centered in the feature space by subtracting the *mean*. Then the centered kernel matrix K_c can be obtained in terms of the kernel matrix K using $K_c = K - 1_m K - K 1_m + 1_m K 1_m$, where 1_m is the $m \times m$ matrix with all its values equal to $\frac{1}{m}$.

The eigen structures of K_c is denoted by $K_c = A\Omega A^T$, where Ω is a diagonal matrix with positive eigenvalues as the diagonal elements. A is the eigenvector matrix corresponding to the positive eigenvalues [16]. Hence the covariance matrix Σ can be denoted as $\Sigma = \left(\Omega^{-\frac{1}{2}}A\phi(x)^T\right) \times \left(\frac{\Omega}{m}\right)\left(\Omega^{-\frac{1}{2}}A\phi(x)^T\right)^T$, where X is the data vectors in feature space. By calculating the pseudo inverse Σ^{+}, we can approximate Σ^{-1} as $\Sigma^{-1} = \Sigma^{+} = mX^TA\Omega^{-2}A^TX$ [7]. Consequently, Eq. (1) will become:

$$\min_{R\epsilon\Re,\xi\epsilon\Re^m} R^2 + \frac{1}{\upsilon m}\sum_{i=1}^{m}\xi_i \tag{3}$$

$$\textit{subject to}: \left\|\sqrt{m}\Omega^{-1}A^TK_c^i\right\|^2 \leqslant R^2 + \xi_i, \xi_i \geqslant 0, \ i = 1, 2, \ldots m$$

where K_c^i is the ith column of the kernel matrix K_c. Using similar Lagrange function and deviations, finally the dual formulation of hyper-ellipsoidal SVM will become a linear optimization problem represented as:

$$\min_{\alpha\epsilon\Re^m} -\sum_{i=1}^{m}\alpha_i\|\sqrt{m}\Omega^{-1}A^TK_c^i\|^2 \tag{4}$$

$$\textit{subject to}: \sum_{i=1}^{m}\alpha_i = 1, 0 \leqslant \alpha_i \leqslant \frac{1}{\upsilon m}, \ i = 1, 2, \ldots m$$

Those data vectors with $\alpha_i = 0$ falling inside the hyperellipsoid will be considered as normal. Those data vectors with $0 < \alpha_i < \frac{1}{\upsilon m}$ will reside on the surface of the hyperellipsoid. Their distances to the hyperellipsoidal center indicate the minimum effective radius R, which can be obtained by calculating $R^2 = \left\|\sqrt{m}\Omega^{-1}A^TK_c^i\right\|^2$ for any *margin support vectors*. Those data vectors with $\alpha = \frac{1}{\upsilon m}$ whose distances to the origin are larger than R of the hyperellipsoid are considered as outliers.

3.2. Hyper-ellipsoid SVM vs. hyper-sphere SVM

Both hyper-ellipsoid SVM and hyper-sphere SVM are used to model the normal behavior of given data vectors. A significant difference between these two SVMs is that they use different *distance measures* to determine the similarity of data vectors before modeling the normal behavior of the data vectors. More specifically, hyper-sphere SVM uses *Euclidean distance* (ED) while hyper-ellipsoid SVM uses *Mahalanobis distance* (MD). These two distance measures are both commonly used to measure the similarity between any two data vectors [17]. Euclidean distance does not consider the correlation between attributes but calculates the distance in terms of individual attribute. On the contrary, Mahalanobis distance considers the correlation between attributes and calculates the distance by combining all attributes together. Consequently, the correlation between attributes learned by Mahalanobis distance can be represented by covariance matrix, where variance of a variable itself and covariance between any two variables are included. Formally, Mahalanobis distance of given multivariate data vectors $x = (x_1, x_2, x_3, \ldots, x_N)$ with mean $\mu = (\mu_1, \mu_2, \mu_3, \ldots, \mu_N)$ and covariance matrix Σ is defined as:

$$MD(x) = \sqrt{(x-\mu)\Sigma^{-1}(x-\mu)^T} \tag{5}$$

If the covariance matrix Σ is the identity matrix I, where all diagonal elements are set to 1, Mahalanobis distance reduces to the Euclidean distance for two data vectors x and y and is represented as:

$$EM(x,y) = \sqrt{\sum_{i=1}^{N}(x_i - y_i)^2} \tag{6}$$

Compared to Euclidean distance, Mahalanobis distance, which takes into account both distances from the center of mass and the direction, has a better understanding of multivariate data structure. This is due to the fact that Euclidean distance is blind to attribute correlation and assumes all data vectors have equal distance from the center of mass. Moreover, Mahalanobis distance is scale-invariant meaning that it is insensitive to the scale of data attributes, while Euclidean distance is extremely sensitive to the scale of data attributes. However, the computational and memory complexity of Mahalanobis distance is much higher than Euclidean distance due to computation of covariance matrix. Table 1 presents the major differences between these two distance measures.

Using Euclidean distance as similarity measure would generate a sphere at 2-D data space, where data vectors are equally distributed around the center of mass. Using Mahalanobis distance as similarity measure, however, would generate an ellipse at the 2-D data space, where data vectors are distributed in directional linear trend [18] indicating the correlation between variables. These two different shapes actually define the normal behavior of data vectors. We here introduce an example to represent the results of outliers using these two shapes. Fig. 2 illustrates the normal behavior of data vectors modeled by hyper-ellipsoid SVM and hyper-sphere SVM using corresponding distance measures. It can be clearly seen that those outliers detected by the sphere may not be considered as outliers by the ellipse (e.g., *point B*); whereas, those data vectors that are not declared as outliers may be considered as outliers by the ellipse (e.g., *point A*). Therefore, using an appropriate shape and its corresponding distance measure to model the normal behavior of data vectors is significantly important for accurate outlier detection. The choice of using Euclidean distance or Mahalanobis distance depends on data characteristics and application requirements.

4. Fitting hyper-ellipsoid one-class SVM modeling to resource-constraint WSNs

In the previous section, we compared hyper-ellipsoid SVM and hyper-sphere SVM regarding modeling the normal behavior of data vectors and identifying outliers. We are aware that modeling hyper-ellipsoid SVM has high computational and memory complexity due to the fact that it considers attribute correlation, generates kernel matrix, and requires the transformation of central kernel matrix. To reduce the cost of modeling hyper-ellipsoid SVM in the feature space, we instead model hyper-ellipsoid SVM in the input space and fix the center of hyperellipsoid at the origin. For doing so, raw sensor data has to be first transformed to a better symmetric data distribution using Box–Cox method [19]. Then due to the fact that Mahalanobis distance is scale-invariant, the data vectors can be centered at the origin just by subtracting the mean. For a transformed data vector x_i', its mean-centered value is formulated as $x_i'' = (x_i' - \mu)$. Considering that mean-centered values may be sensitive to outliers, we replace the arithmetic mean by the *median*.

After the above data preprocessing, the data vectors are centered at the origin in the input space, which lowers down the computational and memory complexity of modeling hyper-ellipsoid SVM in the feature space. Consequently, the dual formulation of Eq. (4) in the input space will be simplified to:

$$\min_{\alpha \in \Re^m} -\sum_{i=1}^{m}\alpha_i\left(x_i''\Sigma^{-1}x_i''^T\right) \tag{7}$$

$$\textit{subject to}: \sum_{i=1}^{m}\alpha_i = 1, 0 \leqslant \alpha_i \leqslant \frac{1}{\upsilon m}, \; i = 1, 2, \ldots m$$

where $x_i''\Sigma^{-1}x_i''^T$ represents the Mahalanobis distances of mean-centered data vectors in the input space from the origin. We further present a basic decision function to

Table 1
Comparison between hyper-ellipsoid SVM and hyper-sphere SVM.

Classifiers	Distance measure	Characteristics	Shape
Hyper-ellipsoid SVM	Mahalanobis distance	Considers attribute correlation Scale-insensitive High complexity	Ellipse
Hyper-sphere SVM	Euclidean distance	Ignores attribute correlation Scale-sensitive Low complexity	Sphere

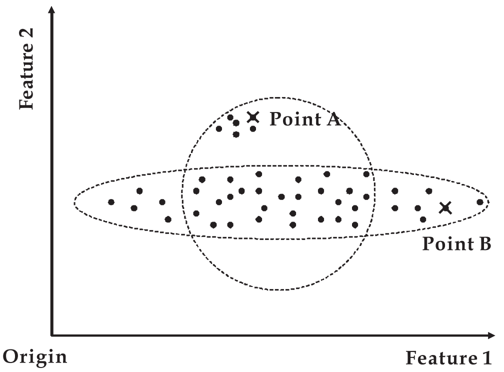


Fig. 2. Use of Mahalanobis distance and Euclidean distance as similarity measure. Mahalanobis distance generates an ellipse while Euclidean distance generates a sphere at 2-D space.

determine whether a new arriving sensor observation x is an outlier using the modeled hyper-ellipsoid SVM in the input space. According to Eq. (7), the decision function can be formulated as:

$$f(x) = sgn(R^2 - d(x'')^2) = sgn(R^2 - x''\Sigma^{-1}x''^T) = sgn\left(R^2 - \left\|\Sigma^{-\frac{1}{2}}x''^T\right\|\right) \quad (8)$$

where R^2 is the square of the effective radius of the hyper-ellipsoid and can be computed by the inner product of any margin support vectors in the input space. Those observations with a negative value are classified as outliers since their square distances from the origin depending on the direction in the input space are larger than R^2. It can be obviously seen from Eq. (4) and (7) that the computation of the decision function in the input space is cheaper than in the feature space. Also, modeling hyper-ellipsoid SVM in the input space solves the problem of impossible calculation of inverse matrix of Σ when Σ is singular [19].

5. Proposed ellipsoidal SVM-based outlier detection techniques

In this section, we propose our distributed and online outlier detection techniques using the addressed simplified hyper-ellipsoid SVM model to identify outliers online and detect changes of the normal behavior of sensor data in real-time.

Let us first consider a small sensor sub-network, which can be easily extend to a cluster-based or a hierarchal network topology. This sub-network consists of densely deployed n sensor nodes $\{s_1, \ldots, s_n\}$, in which observations are made at (nearly) equal time intervals and all nodes directly communicate with each other. Moreover, spatial and temporal correlations are assumed to exist in sensor observations collected in this sub-network. Fig. 3 illustrates an example of such a sub-network.

At a time instant t, $x(s_1,t), \ldots, x(s_n,t)$ denote data vectors measured at nodes $s_1, \ldots, s_n$, respectively. Each data vector at the corresponding node is composed of multiple attributes $x^l(s_i,t)$, where $x^l(s_i,t) = \{x^l(s_i,t){:}i = 1 \ldots, n, l = 1, \ldots, d\}$ and $x(s_i,t) \epsilon \Re^d$.

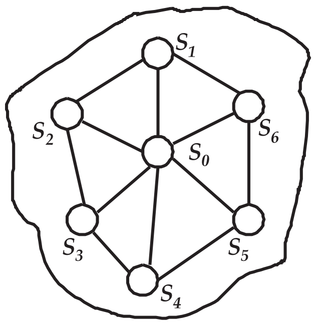


Fig. 3. Example of a sensor sub-network.

According to the requirements of applications, outliers can be identified as *local outliers* or *global outliers*. Local outliers represent those outliers that are detected at individual sensor node only using its local data. Global outliers represent those outliers that are detected in a more global perspective [24] by considering a cluster of sensor nodes. Specifically, global outliers can be identified at a parent node, cluster-head node, or even a central station, by collecting many data from its assigned sensor nodes. Alternatively, global outliers can be identified at individual sensor node using a well-defined normal behavior of sensor data, which is modeled in a global view. In this paper, we use this strategy in our outlier detection techniques to identify global outliers at individual sensor node. One should note that a local outlier may not be identified as a global outlier and vice versa [5]. For instance, a local outlier is an observation collected by a sensor node that is significantly different with respect to other observations of this sensor node, but may not be a global outlier in a global view of a cluster of neighboring nodes.

5.1. Ellipsoidal SVM-based online outlier detection technique (EOOD)

Our ellipsoidal SVM-based online outlier detection technique (EOOD) enables each node to determine its every new observation as normal or outlier in real-time. Specifically, each node models its own hyper-ellipsoid SVM after collecting sensor data during a time interval. As a result, each node obtains the effective radius of the modeled hyperellipsoid together with the median and the covariance matrix of centered data. Based on temporal correlation of sensor data, each node uses the modeled hyper-ellipsoid SVM to determine whether its new arrived observations are normal or outliers in the time domain. Moreover, each node communicates the effective radius of the modeled hyperellipsoid, the median and the covariance matrix parameters with its neighboring nodes to cooperatively identify outliers based on spatial correlation of sensor data. The main steps of EOOD are:

- *Step 1.* Each node s_i models its own hyper-ellipsoid SVM for m sequential observations and then calculates the effective radius R_i as well as the corresponding parameters of the median and the covariance matrix of centered data. Then the local outliers at each sensor node

s_i can be determined in real-time by comparing the distances between new arrived observations and the origin with R_i using Eq. (8).

- *Step 2*. Each node s_i communicates its parameters, i.e., R_i as well as the median and the covariance matrix to its neighbors. R_i and the median are transmitted by 1 element, respectively. For the covariance matrix, each node s_i only needs to transmit $d(d+1)/2$ elements due to the fact that the covariance matrix is symmetric, where d is the dimension of a sensor observation.
- *Step 3*. Each node s_i collects these R_i, median and covariance matrix parameters from its neighbors and combines these parameters together with its own R_i, median and covariance matrix. Specifically, these effective radii and medians are merged by arithmetic average and these covariance matrixes need to be merged by the formula used in [6]. These merged parameters are denoted as the global effective radius R_g, the global median, and the global covariance matrix Σ_g.
- *Step 4*. Each node s_i uses these global parameters to determine the global outlierness of its new arrived observations in real-time. For a new observation x, the decision function indicating whether it is a global outlier in the input space can be defined as:

$$f(x) = sgn\left(R_g^2 - \left\|\Sigma_g^{-\frac{1}{2}} x''^T\right\|\right) \tag{9}$$

EOOD scales well with the increased number of nodes due to its distributed processing nature. It enables to update the global parameters by communicating among neighboring nodes at the end of each time interval. It lowers down communication overhead and computational complexity, especially no need to transmit any actual observations between sensor nodes except the required parameters, i.e., R_i, median and covariance matrix.

EOOD can detect the change of the normal behavior between two consecutive time windows when most of observations measured in the second time window are detected as outliers. However, EOOD cannot detect changes of the normal behavior of data within one time window or adapt to a new normal behavior of sensor data due to the fact that it does not update the existing SVM model until the end of the entire time window. In this way, EOOD may suffer from a possibly high rate of false alarm when new observations arrived in the same time window representing a new normal behavior of sensor data are detected as outliers. In order to alleviate this problem, our adaptive outlier detection technique incorporates new arrived observations and updates the modeled hyper-ellipsoid SVM for more reliable outlier detection.

5.2. Ellipsoidal SVM-based adaptive outlier detection technique (EAOD)

Our ellipsoidal SVM-based adaptive outlier detection technique (EAOD) enables each node to detect changes of the normal behavior of sensor data within a time window based on decision results achieved in a sliding window and then update the ellipsoidal SVM-based model to represent the changed normal behavior of sensor data for further outlier detection. The use of the sliding window in EAOD is to incorporate new arrived observations and meanwhile remove the oldest observations. Consequently, the hyper-ellipsoid SVM can be updated using the observations representing a change of the normal behavior in the sliding window. Initially the sliding window includes all m sequential observations for modeling a hyper-ellipsoid SVM.

Before describing when and how to detect a new normal behavior using the sliding window, we recall the parameter v, which is very important to model a hyper-ellipsoid SVM in Eq. (1). It controls the fraction of data vectors that can be outliers and it significantly impacts the performance of one-class SVM-based outlier detection techniques. The v actually denotes the upper bound on the fraction of detected outliers in a dataset [11], which indicates the maximum number of outliers allowed by a SVM model. In this ways, if the fraction of outliers detected by the hyper-ellipsoid SVM model exceeds the given upper bound within a time period, it indicates that a new normal behavior is emerged and the previously modeled SVM is not suitable to represent the current normal behavior of sensor data any more so as to be updated. On the other hand, we consider reducing the computational and communication cost of updating the SVM model and being able to compare the performance of our outlier detection techniques with the existing techniques described in [9,10,23]. Thus our proposed EAOD decides to update the modeled hyper-ellipsoid SVM when the same amount of m new sequential observations inserted in the sliding window is instantly detected as normal or outlier and also the fraction of detected outliers in this sliding window exceeds the given upper bound (v).

After detecting a new arriving observation as normal or outlier, each node does not immediately update the effective radius (R) of the model but instead only updates the median and covariance matrix [21] of the changed sliding window, in which the oldest observation is removed and replaced by the new observation, for further outlier detection. There is no need for communication among nodes until the effective radius (R) of the model changes. One should note that all new arrived observations, regardless of being detected as normal or outlier, can be incorporated into the sliding window due to the fact that the parameter v of the one-class SVM model allows anomalous observations in the training set. Moreover, removing anomalous observations from the set would bias the normal boundary of the one-class SVM model [20]. When all new observations are instantly detected as normal or outlier and inserted in the entire sliding window, each node checks if the fraction of detected outliers exceeds the given upper bounder (v). If so, a new normal behavior is detected in this sliding window and the SVM model needs to be updated, i.e., the effective radius R. After the hyper-ellipsoid SVM is updated, each observation in the sliding window can be labeled as normal or outlier using the updated SVM model and Eq. (9).

EAOD enables robust detection of outliers using sequential observations and also detects changes of the normal behavior of sensor data for reliable outlier detection. It recognizes the previously detected outliers as the

Please cite this article in press as: Y. Zhang et al., Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.11.001

indication of a new normal behavior and reduces the false alarm rate in EOOD. Moreover, EAOD is generated efficiently in terms of communication while requires less computational time. Fig. 4 illustrates the update policy of EAOD. The corresponding pseudocode for EAOD is shown in Table 2.

6. Simulation results

This section describes simulation results of our EOOD and EAOD, compared to the spherical SVM-based adaptive outlier detection technique (SAOD) we proposed in [22], and the ellipsoidal SVM-based batch outlier detection technique (EBOD) and the spherical SVM-based batch outlier detection technique (SBOD) presented in Rajasegarar et al. [9,10,23]. Our proposed SAOD in [22] uses the quarter-sphere one-class SVM to model the normal behavior of sensor data. It takes the similar strategy as EAOD to detect changes of the normal behavior of sensor data and update the modeling quarter-sphere SVM for reliable outlier detection.

For EBOD and SBOD proposed in [9,10,23], as we described before, they cause a considerable outlier detection delay, in which each node analyzes sensor data in an offline manner only after all observations are collected within a day. They also do not provide online outlier detection for new arriving observations. More specifically, EBOD is neither online nor distributed. It only identifies outliers in a single node without any parameter or raw data exchange so that the achieved outlier detection results are not sufficiently reliable in case of node failure. SBOD uses the quarter-sphere one-class SVM for distributed outlier detection but exchanges just only radius information among neighboring nodes for outlier detection while ignoring the other important modeling parameters, i.e., the mean and the standard deviation. These modeling parameters of the quartersphere SVM contribute to provide more reliable outlier detection results. Furthermore, EBOD and SBOD that model the SVM in the feature space need the generation of kernel matrix and the transformation of centered kernel matrix so that they bring very high computational and memory complexity for WSNs.

The goals of our simulation in this section are threefold, i.e., (i) to test the accuracy of our distributed and online outlier detection techniques compared to the existing SVM-based outlier detection techniques and their robustness in terms of parameter selection, (ii) to compare the accuracy between ellipsoidal and spherical SVM-based outlier detection techniques, and (iii) to investigate impact of three commonly used labeling techniques, i.e., Mahalanobis distance, density and running average, on performance of outlier detection techniques.

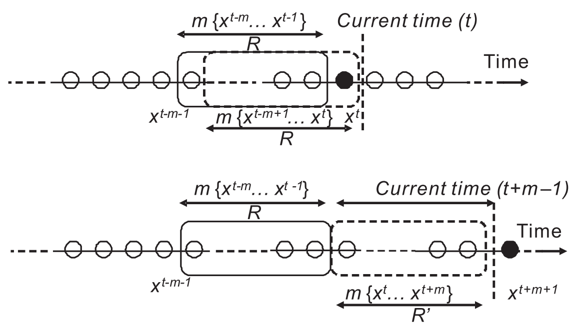


Fig. 4. The update policy of EAOD. The circles represent sensor observations. The sliding window is composed of the last m observations. The black dot represents the observation identified at current time t.

6.1. Simulation datasets

In our simulation, we use two c datasets as well as a real dataset gathered at the Grand St. Bernard [8]. The synthetic datasets we used are similar to the one used in [24]. We consider a sensor sub-network consisting of seven sensor nodes, as illustrated in Fig. 3, which can be within radio transmission range of each other or communicated with a cluster-head in this sub-network.

The first 2-D synthetic dataset is composed of 200 data vectors for each node having a mixture of three Gaussian distributions with uniform distribution of outliers. The mean value of this dataset is randomly selected from (0.32, 0.35, 0.38) while the standard deviation is set to be 0.03. Subsequently, 10 uniform outliers (i.e., 5% of the normal data) are introduced and uniformly distributed in the $[0.5, 1]$ interval. The total number of the data vectors in this dataset is 2940 including the 5% outliers. Fig. 5 illustrates the data distribution of dataset of a single node.

The second 2-D synthetic dataset changes the mean of a mixture of three Gaussian distributions into (0.25, 0.35, 0.45). The standard deviation is still 0.03 and 5% (of the normal data) anomalous data is introduced and uniformly distributed in the $[0.5, 1]$ interval. Fig. 5 illustrates the data distribution of dataset of a single node. These two synthetic datasets aim to evaluate the accuracy of ellipsoidal and spherical SVM-based outlier detection techniques for different data distributions.

The real dataset is collected from the small cluster of neighboring sensor nodes, i.e., nodes 25, 28, 29, 31, 32 at the Grand St. Bernard, as illustrated in Fig. 6. In our simulations, we test the real data collected during the period of 6am–14am on 1st October 2007 with two attributes: ambient temperature and relative humidity for each sensor observation. We label this dataset using three different labeling techniques, i.e., Mahalanobis distance, density and running average. Labeling results of applying these labeling techniques are illustrated in Fig. 7. More details about labeling techniques for sensor data are referred to [25].

6.2. Simulation results

We evaluate two important accuracy metrics, (i) the detection rate (DR), which represents the percentage of outliers that are correctly detected, and (ii) the false alarm rate, also known as false positive rate (FPR), which represents the percentage of normal data that are incorrectly considered as outliers. DR represents the ratio between the number of correctly detected outliers and the total number of outliers, while FPR represents the ratio between the number of normal data detected as outliers and the total number of normal data.

Table 2
Pseudocode of EAOD.

```
procedure ModellingSVMProcess()
each node models the hyper-ellipsoid SVM;
each node locally broadcasts the modeled hyper-ellipsoid R_i as well as the median and the covariance matrix to its spatially neighboring nodes;
each node then computes the global R_g as well as the global median and covariance matrix;
initiate OutlierDetectionProcess(R_g, the global median and covariance matrix);
return;
procedure OutlierDetectionProcess(R_g, the global median and covariance matrix)
  when x(t) arrives
    compute d(x);
    if (d(x) > R_g)
      x(t) indicates an outlier;
    else
      x(t) indicates a normal observation;
    endif;
    initiate UpdatingSVMProcess(x(t));
    set t ← t + 1;
    if (m data observations are collected)
      if (the fraction of detected outliers  > the given upper bound (v))
        update the SVM model for outlier detection of (x(t − m + 1) ... x(t));
      endif;
    endif;
  return;
procedure UpdatingSVMProcess(x(t))
update the sliding window: the oldest observations x(t − m) is removed and replaced by x(t);
update the median and the covariance matrix of the sliding window;
return;
```

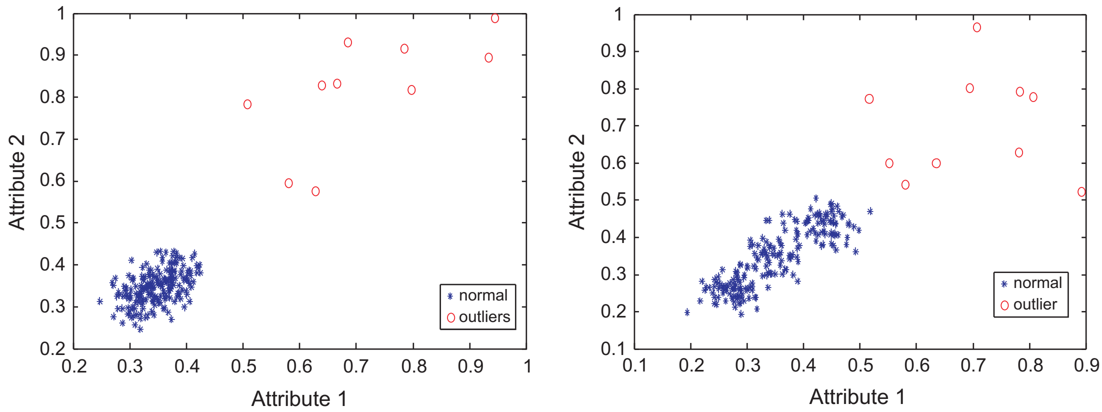


Fig. 5. (left) Data plot for a single node with spherical data distribution and (right) data plot for a single node with ellipsoidal data distribution.

We also examine the effect of the regularization parameter υ for EBOD, SBOD, EOOD, EAOD and SAOD in the input space. As representing the fraction of data vectors that can be outliers, the parameter υ with the larger value results in better detection rate, however, it can lead to higher false alarm rate. The value of υ is usually chosen based on a priori knowledge of the data characteristics and its normal behavior [20]. In the case of having no a priori knowledge about the outliers ratio, the parameter υ can be used to evaluate the robustness of the techniques. It indicates that a robust technique can achieve high accuracy rate while keeping a false alarm rate low regardless of the increase or decrease of the parameter υ. In the simulation, we have varied υ between 0.02 and 0.08 in intervals of 0.01. A receiver operating characteristics (ROC) curve is usually used to represent the trade-off between the detection rate and the false alarm rate. The larger the area under the ROC curve, the better the accuracy of the technique. Furthermore, after finding a robust outlier detection techniques, a relatively reliable υ can be determined when the outlier detection technique archives the best trade-off between the detection rate and the false alarm rate, which can be illustrated in the area under the ROC.

Fig. 8 shows the detection rate and the false alarm rate obtained by our online techniques EOOD, EAOD, SAOD as well as EBOD and SBOD offline techniques in the input space for the first synthetic data with ellipsoidal data distribution. We can see when data vectors have ellipsoidal data distribution, our ellipsoidal SVM-based techniques EOOD and EAOD achieve better detection accuracy and lower false alarm compared with spherical SVM-based SBOD and SAOD in presence of different υ parameters. Furthermore, our EOOD and EAOD perform better than EBOD.

Fig. 6. The small cluster of the Grand St. Bernard deployment [8].

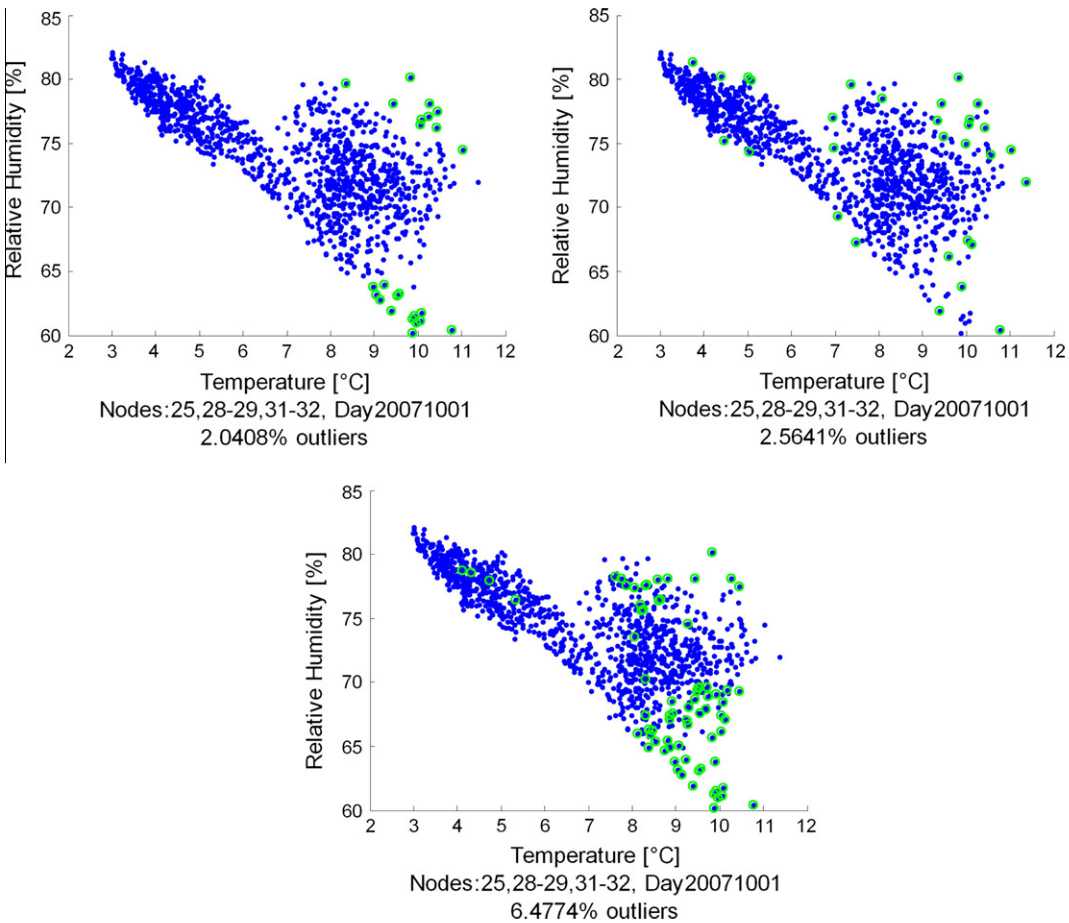


Fig. 7. (left) Plot for labeled data based on Mahalanobis distance, (right) Plot for labeled data based on density and (lower) plot for labeled data based on running average.

Please cite this article in press as: Y. Zhang et al., Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.11.001

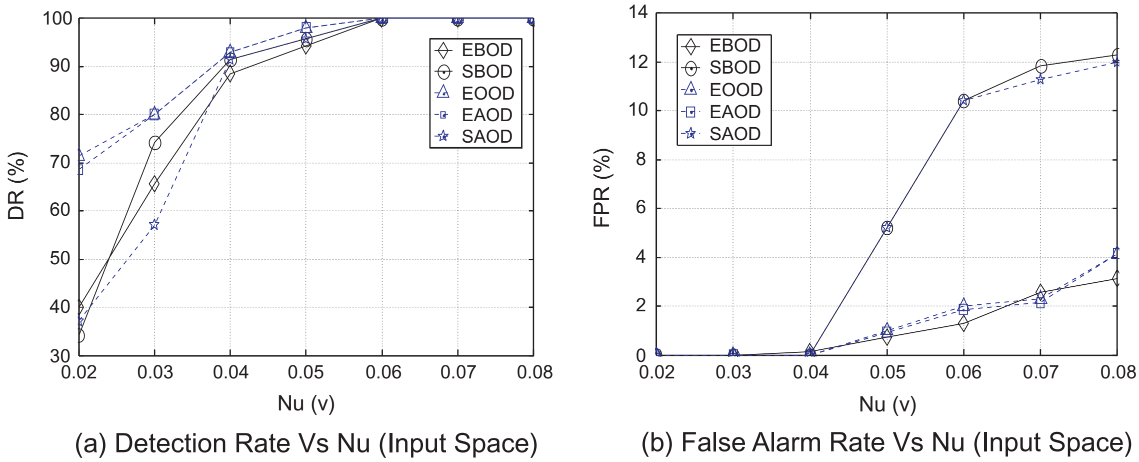


Fig. 8. (a) Detection rate in the input space for synthetic data of ellipsoidal data distribution and (b) false alarm rate in the input space for synthetic data of ellipsoidal data distribution.

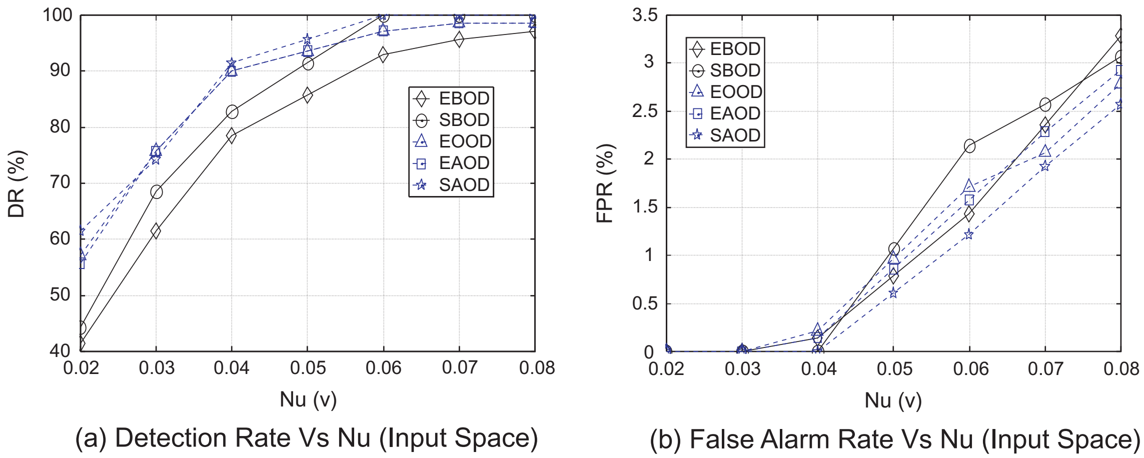


Fig. 9. (a) Detection rate in the input space for synthetic data of spherical data distribution and (b) false alarm rate in the input space for synthetic data of spherical data distribution.

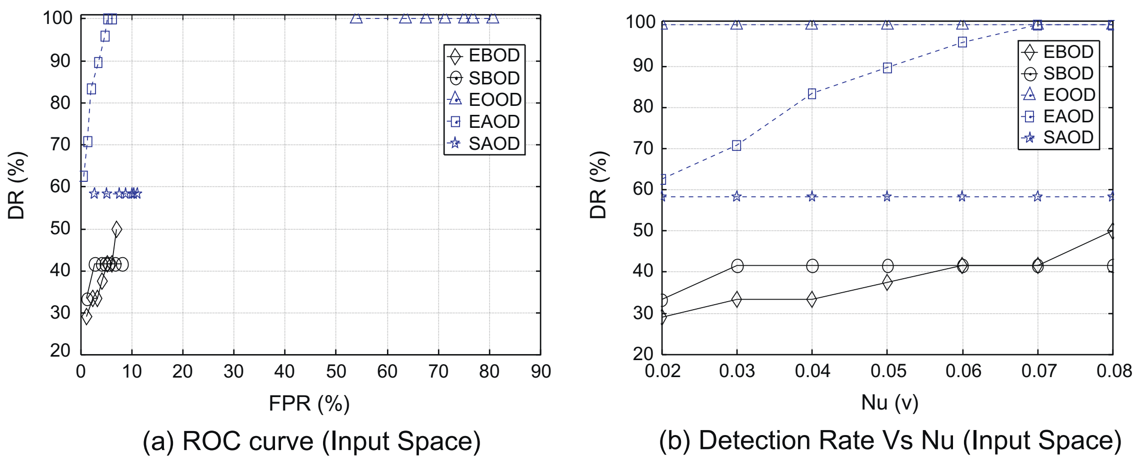
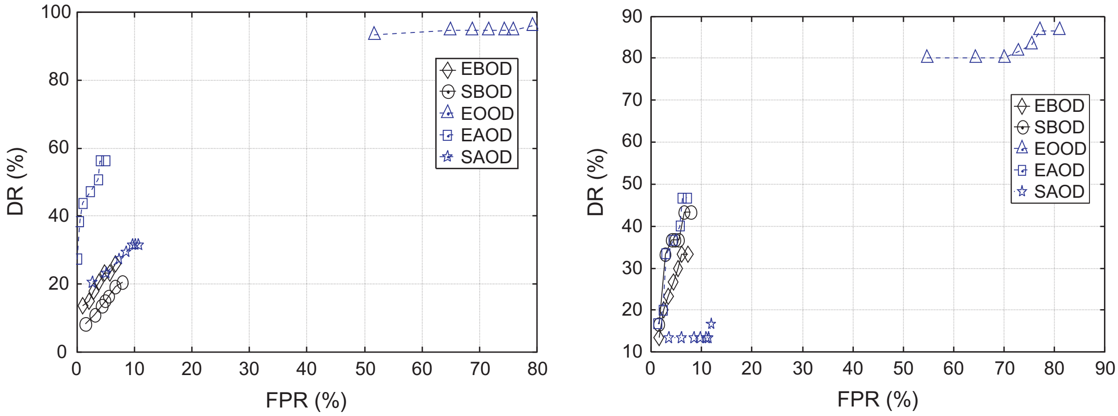


Fig. 10. (a) ROC curve in the input space for labeled data based on Mahalanobis distance and (b) detection rate in the input space for labeled data based on Mahalanobis distance.

Please cite this article in press as: Y. Zhang et al., Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.11.001



(a) ROC curve (Input Space) (b) ROC curve (Input Space)

Fig. 11. (a) ROC curve in the input space for labeled data based on running average and (b) ROC curve in the input space for labeled data based on density.

Fig. 9 shows the detection rate and the false alarm rate obtained by our online techniques EOOD, EAOD, SAOD as well as EBOD and SBOD offline techniques in the input space for the second synthetic data with spherical data distribution. We can clearly see when data vectors have spherical data distribution, our spherical SVM-based technique SAOD achieves better detection accuracy and lower false alarm compared with ellipsoidal SVM-based EOOD, EAOD and EBOD. Moreover, our SAOD performs better than SBOD. Furthermore, our EOOD and EAOD have better detection accuracy than EBOD.

Fig. 10 shows ROC curve and the detection rate obtained by our online techniques EOOD, EAOD, SAOD as well as EBOD and SBOD offline technique in the input space for the real dataset labeled by Mahalanobis distance. It can be seen that EOOD generates the highest detection rate and highest false alarm due to the fact that it does not update the normal profile, while the data distribution has changed. EAOD achieves the best accuracy with the highest detection accuracy and the lowest false alarm.

Fig. 11 shows ROC curve obtained by our online techniques EOOD, EAOD, SAOD as well as EBOD and SBOD offline techniques in the input space for the real dataset labeled by running average and density labeling techniques. The results show that accuracy of our EAOD is better than other techniques using both labeling techniques although it has not achieved good detection accuracy. EOOD still has the highest false alarm rate while generating high detection rate.

7. Performance evaluation

This section further analyzes accuracy results achieved in our simulation. The complexity of those techniques is also compared in terms of communication overhead, computation and memory complexity.

7.1. Accuracy analysis

For the data vectors which have ellipsoidal data distribution, EOOD and EAOD perform better than spherical SVM-based SBOD and SAOD. The good results of EOOD and EAOD stem from taking into account the correlation of data attributes and having better understanding of multivariate nature of data distribution. On the contrary, spherical SVM-based techniques ignore the correlation between data attributes and use a spherical boundary to fit the data. This results in low detection rate and high false alarm rate in case of non-spherical data distribution. Moreover, among ellipsoidal SVM-based techniques, our EOOD and EAOD perform better than EBOD due to the fact that they exchange essential ellipsoidal information, e.g., median and covariance matrix, with neighboring nodes for reliable outlier detection.

For the data vectors which have spherical data distribution, SAOD achieves better detection accuracy and lower false alarm compared with ellipsoidal SVM-based EOOD, EAOD and EBOD. This is because spherical SVM-based techniques assume that data vectors are distributed around the center of mass in an ideal spherical shape. Although ellipsoidal SVM-based techniques take into account correlation of data attribute, they do not perform as good as spherical SVM-based techniques for spherical data distribution. Moreover, among spherical SVM-based techniques, our SAOD performs better than SBOD since SAOD alleviates the influence of outliers by using median and median absolute deviation (MAD) and also exchanges these spherical information with neighboring nodes for reliable outlier detection.

For the real data, EAOD achieves the best accuracy with the highest detection accuracy and the lowest false alarm since it considers the correlation of data attributes as well as use of ellipsoidal information (median, covariance matrix) from neighboring nodes. On the contrary, EOOD generates the highest false alarm due to the fact that it does not update the normal profile, while the data distribution has changed. Therefore, EAOD performs best compared to other SVM-based techniques.

Furthermore, one should note that hyperellipsodial SVMs used in this paper generally suit to the multivariate data vectors with the correlated data attributes. The direction of the formed ellipsoid reveals the multivariate nature of data distribution trend as well as the strength of the correlation between data attributes. For the data vectors with

Please cite this article in press as: Y. Zhang et al., Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.11.001

Table 3
Complexity analysis of five outlier detection techniques for each sensor node.

Techniques	Communication complexity	Computational complexity	Memory complexity
EBOD	–	$O(kmd^2)$	$O(md + m^2)$
SBOD	$O(d)$	$O(kmp)$	$O(md + m^2)$
EOOD	$O(d^2)$	$O(md^2)$	$O(md)$
EAOD	$O(d^2)$	$O(md^2)$	$O(md)$
SAOD	$O(d)$	$O(mp)$	$O(md)$

no strong data attribute correlation or with a data distribution of a skewed shape, a spherical SVM or a hyperplane-based SVM could be used to characterize the kind of data sets. Therefore, a good understand of data distribution and correlation among data attributes is essential to model the normal behavior of data vectors and design a suitable outlier detection technique.

7.2. Complexity analysis

The communication complexity of our distributed techniques depends on the transmission of local hyper-ellipsoid radius information as well as the median and covariance matrix parameters. The communication overhead in EOOD for each node is $O(d^2)$, where d is the dimension of observations. Each node only transmits its local hyper-ellipsoid radius information as well as the median and covariance matrix once at the initial training phase. EAOD requires no update of radius information during online outlier detection and only possibly communicates the updated median, covariance matrix and radius information with nodes at the end of a sliding time window. The maximum communication overhead of EAOD for each node is approximately equal to $O(d^2)$.

The computational complexity in EOOD is related to computation of the median, the covariance matrix, the linear optimization function and the distance between every new observation and the origin. The computational complexity of our techniques mainly depends on solving a linear optimization problem, which is represented as $O(p)$, as well as computing covariance matrix, which is represented as $O(md^2)$. Hence, the maximum computational complexity of each node in EOOD and EAOD is $O(md^2)$, where m is the number of new observations to be classified. EBOD still needs to compute kernel matrix and the transformation of centered kernel function (especially for RBF kernel function), whose complexity is represented by $O(k)$. Thus the maximum computational complexity of EBOD for each node is $O(kmd^2)$.

The memory complexity of our techniques is mainly related to keeping observations of the size of sliding window in memory and is represented as $O(md)$, where d is the dimension of observations and m is the number of new observations to be classified. Overhead of storing other parameters such as covariance matrix with a complexity of $O(d^2)$ is negligible since $m > d$. Hence the maximum memory complexity of each node for our techniques is $O(md)$. Due to the fact that EBOD needs to keep $m \times m$ kernel function, its memory complexity of each node is $O(md + m^2)$. Table 3 summarizes these complexities.

8. Conclusion

In this paper we propose two distribute and online outlier detection techniques based on hyper-ellipsoid one-class SVM. We take into account data attribution correlation to precisely detect outliers. To cope with the problem of generating high false alarm rate, we also propose an updating strategy to incorporate new arrived observations and update the modeled hyper-ellipsoid SVM for more reliable outlier detection and detect changes of the normal behavior of sensor data. We compare performance of these two hyper-ellipsoid SVM-based techniques with our previously proposed quarter-sphere SVM-based technique as well as two existing batch SVM-based techniques using both synthetic and real datasets as well as different labeling techniques. Simulation results show that our EAOD achieves better detection accuracy and lower false alarm. It implies that understanding data distribution and correlation among data attributes is essential to design a suitable outlier detection technique. Our future research includes testing our EAOD using real datasets with variant data distributions and implementing EAOD on wireless sensor nodes in real-life.

Acknowledgment

This work is supported by the EU's Seventh Framework Programme in the context of the SENSEI and GENSEI projects.

References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, IEEE Communications Magazine 40 (2002) 102–114.
[2] C.F. Garcia-Hernandez, P.H. Ibarguengoytia-Gonzalez, J. Garcia-Hernandez, J.A. Perez-Diaz, Wireless sensor networks and applications: a survey, International Journal of Computer Science and Network Security 7 (2007) 264–273.
[3] T. Arampatzis, J. Lygeros, S. Manesis, A survey of applications of wireless sensors and wireless sensor networks, in: Proceedings of the 13rd Mediterranean Conference on Control and Automation, Limassol, Cyprus 2005, pp. 719–724.
[4] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: a survey, ACM Computing Surveys 41 (2009) 1–58.
[5] Y. Zhang, N. Meratnia, P.J.M. Havinga, Outlier detection techniques for wireless sensor network: a survey, IEEE Communications Surveys & Tutorials 12 (2010) 159–170.
[6] P. Kelly, An Algorithm for Merging Hyperellipsodial Clusters, Technical report, Los Alamos National Laboratory, 1994.
[7] D. Wang, D.S. Yeung, E.C.C. Tsang, Structured one-class classification, IEEE Transactions on System, Man and Cybernetics 36 (2006) 1283–1295.
[8] F. Ingelrest, G. Barrenetxea, G. Schaefer, M. Vetterli, O. Couach, M. Parlange, SensorScope: application-specific sensor network for environmental monitoring, ACM Transactions on Sensor Networks 6 (2010) 1–32.
[9] S. Rajasegarar, C. Leckie, M. Palaniswami, J.C. Bezdek, Quarter sphere based distributed anomaly detection in wireless sensor networks, in: Proceedings of IEEE International Conference on Communications, Glasgow, 2007, pp. 3864–3869.
[10] S. Rajasegarar, C. Leckie, M. Palaniswami, CESVM: Centered hyperellipsoidal support vector machine based anomaly detection, in: Proceedings of IEEE International Conference on Communications, Beijing, China, 2008, pp. 1610–1614.
[11] B. Scholkopf, J. Platt, J. Shawe-Taylor, A.J. Smola, R.C. Williamson, Estimating the support of a high dimensional distribution, Journal of Neural Computation 13 (2001) 1443–1471.
[12] D.M.J. Tax, R.P.W. Duin, Support vector data description, Journal of Machine Learning 54 (2004) 45–56.
[13] C. Campbell, K.P. Bennett, A linear programming approach to novelty detection, Advances in Neural Information Processing Systems 14 (2001) 395–401.

[14] P. Laskov, C. Schafer, I. Kotenko, Intrusion detection in unlabeled data with quarter sphere support vector machines, Detection of Intrusions and Malware & Vulnerability Assessment (2004) 71–82.
[15] V.N. Vapnik, Statistical Learning Theory, John Wiley and Sons, 1998.
[16] G.H. Golub, C.F.V. Loan, Matrix Computations, John Hopkins, 1996.
[17] P.N. Tan, Knowledge discovery from sensor data, Sensors (2006).
[18] J. Han, M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, San Francisco, 2006.
[19] K. Varmuza, P. Filzmoser, Introduction to Multivariate Statistical Analysis in Chemometrics, CRC Press, 2009.
[20] M. Davy, F. Desobry, A. Gretton, C. Doncarli, An online support vector machine for abnormal events detection, Journal of Signal Processing 8 (2006) 52–57.
[21] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, John Wiley and Sons, 2001.
[22] Y. Zhang, N. Meratnia, P.J.M. Havinga, Ensuring high sensor data quality through use of online outlier detection techniques, International Journal of Sensor Networks 7 (2010) 141–151.
[23] S. Rajasegarar, C. Leckie, J.C. Bezdek, M. Palaniswami, Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks, IEEE Transactions on Information Forensics and Security 5 (2010) 518–533.
[24] S. Subramaniam, T. Palpanas, D. Papadopoulos, V. Kalogerakiand, D. Gunopulos, Online outlier detection in sensor data using non-parametric models, Journal of Very Large Data Bases (2006).
[25] Y. Zhang, Observing the Unobservable – Distributed Online Outlier Detection in Wireless Sensor Networks, PhD thesis, University of Twente, The Netherlands, 2010.
[26] P.N. Tan, M. Steinback, V. Kumar, Introduction to Data Mining, Addison Wesley, 2006.

Yang Zhang is a postdoctoral researcher in the Pervasive Systems Group at the University of Twente in the Netherlands. He received the B.Sc. and M.Sc. degrees in Computer Science and Technology from the University of Jiangsu, China, in 2002 and 2004. Afterwards he obtained the second M.Sc. degree in Telematics and his Ph.D. degree in the area of outlier detection for wireless sensor networks from the University of Twente in 2006 and 2010. He has participated in a few European funded projects (EYES, e-SENSE, SENSEI) and is currently involved in EU GENESI project. His research interested include distributed data processing, outlier detection and event detection in sensor networks.

Nirvana Meratnia is assistant professor in the Pervasive Systems Group at the University of Twente. After receiving her Ph.D. in 2005 on moving object data management, she joined the Computer Architecture for Embedded Systems (CAES) group and later on the Pervasive Systems (PS) group, as a researcher. Her research interests are in the area of distributed data management/data mining and reasoning in wireless sensor networks, ambient intelligence, context-awareness, smart and collaborative objects. She has been involved in a number of EU (i.e., e-SENSE, CoBIs, and Embedded WiSeNts) and Dutch (i.e., Smart Surroundings) projects focusing on ambient intelligence, collaborative smart objects, and distributed data processing and reasoning in wireless sensor networks.

Paul J.M. Havinga is full professor and chair of the Pervasive Systems research group at the Computer Science department at the University of Twente in the Netherlands. He received his Ph.D. at the University of Twente on the thesis entitled ''Mobile Multimedia Systems'' in 2000, and was awarded with the 'DOW Dissertation Energy Award' for this work. He has a broad background in various aspects of communication systems: on wireless communication, on chiparea network architectures for handheld devices, on ATM network switching, mobile multimedia systems, QoS over wireless networks, reconfigurable computing, and on interconnection architectures for multiprocessor systems. His research themes have focused on wireless sensor networks, large-scale distributed systems, and energy-efficient wireless communication.

Please cite this article in press as: Y. Zhang et al., Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine, Ad Hoc Netw. (2012), http://dx.doi.org/10.1016/j.adhoc.2012.11.001