International Journal of Information and Communication Technology Research

http://www.esjournals.org

An Intelligent Intrusion Detection System Based On Expectation Maximization Algorithm in Wireless Sensor Networks

[1]Hamed Khanbabapour ,[2]Hamid Mirvaziri
Department of Computer, Science and Research Branch, Islamic Azad University of Kerman, Kerman, Iran
Department of Computer Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

ABSTRACT

In recent years, wireless technology has enjoyed a tremendous rise in popularity and its usage opens new fields of applications. Due to their possible deployment in remote locations for civil, educational, scientific, military purposes and security which includes intrusion detection and intrusion prevention are of utmost importance.Thewireless sensor networks have problems on detecting and preventing malicious nodes which always bring destructive threats and compromise multiple sensor nodes. Therefore, sensor networks need to support an authentication service for sensor identity and message transmission. Furthermore, intrusion detection and prevention schemes are always integrated in sensor security appliances so that they can enhance network security by discovering malicious or compromised nodes.
In this paper, we propose a protocol for intrusion detection in clustered wireless sensor networks (WSNs). Detection system is implemented based on expectation maximization algorithm. We implemented the system in network simulator "GloMoSim". Also we have rigorously evaluated the performance of our proposed solution by performing a variety of experiments and have found our solution approach to be promising.

Keywords: *Intrusion Detection System(IDS),Expectation Maximization (EM),Wireless Sensor Network (WSN)*

1. INTRODUCTION

Wireless sensor networks (WSNs) have become animportant technology, combining sensing technology, embedded computing, distributed information processing, and wireless communication technology [1].WSNs have broad applications, such as medical monitoring, environment pollution monitoring, forest firemonitoring, target tracking, combat field reconnaissance,and military command and control, and so on. Data collectionis the basic objective of in these applications. Data collectioncapability of a sensor network depends on itssensing coverage and network connectivity. Sensor nodesare often powered by batteries, and it is often difficult or impossible to recharge the deployed nodes.Great effortshave been devoted to minimizing the energyconsumption and extending the lifetime of thenetwork [1].

In recent years, Intrusion Detection Systems find tremendous importance in the field of detection technology and is used as a countermeasure to preserve data integrity and system availability during an intrusion. When an intruder attempts to break into an information system or performs an action not legally allowed, we refer to this activity as an intrusion. Intruders can be divided into two groups, external and internal. The former refers to those who do not have authorized access to the system and who attack by using various penetration techniques. The latter refers to those with access permission who wish to perform unauthorized activities. Intrusion techniques may include exploiting software bugs and system misconfigurations, password cracking, sniffing unsecured traffic, or exploiting the design flaw of specific protocols [2].Network Security consists of the provisions made in an underlined computer network infrastructure and policies adopted by the Network Administrator to protect the network and networkaccessible resources from unauthorized access, consistent and continuous monitoring and measurement of its effectiveness combined together.Computer forensics is a branch of forensic science pertaining to legal evidence found in computers and digital storage mediums.Computer forensics is also known as digital forensics. The goal of computer forensics is to explain the current state of digital artifacts. The term digital artifact may include a computer system, a storage medium (such as hard disk or CDROM), an electronic document (and email message or JPEG message) or even a sequence of packets moving over a computer network [3], [4].

Intrusion Detection Systems (IDS) are systems for detecting intrusions and reporting them accurately to theproper authority. There are generally two accepted categories of intrusion detection techniques: misuse detection and anomaly detection. Misuse detection refers to techniques that characterize known methods to penetrate a system. These penetrations are characterized as a 'pattern' or a 'signature' that the IDS look for.IDS can also be divided into two groups depending on the place where they look for intrusive behavior: Network-based IDS and Host-basedIDS. The former refers to systems that identify intrusions by monitoring traffic through network devices. Host-based IDS monitor file and process activities related to a software environment associated with a specific host. The architecture combines a number of different approaches to the intrusion detection problem, and includes different techniques based on Artificial Intelligence (AI) to help identify intrusive behavior. It uses both anomaly detection and misuse detection techniques and is both a network-based and host-based system [5].

AI is a machine learning method that is concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data such as traffic through network devices. Expectation Maximization (EM) is a method for finding maximum likelihood or maximum a posteriori (MAP) estimates of parameters of data, depending on unobserved latent variables. Latent variables are variables that are not directly observed but are rather inferred from other variables that are observed (directly measured.) EM is an iterative method which alternates between performing an expectation (E) step, which computes the expectation of the log-likelihood, evaluated using the current estimate for the latent variables, and maximization (M) step, which computes parameters maximizing the expected loglikelihoodfound on the E step. These parameter-estimates are then used to determine the distribution of the latent variables in the next E step [6].

A Bayesian Classification (BC) is a simple probabilistic classification technique based on applying Bayes' theorem with strong independence assumptions and is referred to as "independent feature model" [7].In fact, intrusion detection can be regarded as a classification problem, namely; identifying normal and other types of intrusive behavior. Hence, the key point here is to choose an effective classification approach to build accurate intrusion detection models. A number of machine learning algorithms have been applied to intrusion detection to learn intrusion rules or build normal usage patterns.

AI strategies were developed aiming at finding solutions to a broad class of problems, named complex problems which could not be resolved by traditional methods. Some of the methods presented constitute hybrid approaches, where AI techniques, like neural networks, fuzzy systems and most influentially evolutionary strategies, were combined leading to the emergence of new computational paradigms. There are many AI approaches that use one or the other above-mentioned model to solve the intrusion detection problem like: data mining techniques, fuzzy rule learning, neural networks, genetic algorithms, learning automata based techniques and so on.

The aim of using data mining techniques over system audit data is to eliminate, as much as possible, the manual and ad-hoc elements from the process of building an intrusion detection system. Intrusion detection is done as a data analysis process from data-centric point of view. Here anomaly detection is about finding the normal usage patterns from the audit data, whereas misuse detection is about encoding and matching the intrusion patterns using the audit data [8].

The behaviors in computer network are hard to predict. This prediction process may generate false alarms in many anomaly based intrusion detection systems. But if
we use fuzzy logic, the false alarm rate in determining intrusive activities can be reduced; a set of fuzzy rules can be used to define the normal and abnormal behavior in a computer network, and fuzzy inference algorithm can be applied over such rules to determine when an intrusion is in progress. The main problem with this approach is to generate good fuzzy classifiers to detect intrusions. This approach deals with the fuzzy classifiers using genetic algorithms that can detect anomalies and some specific intrusions. The main idea is to evolve two rules, one for the normal class and other for the abnormal class using a profile data set with information related to the computernetwork during the normal behavior and during intrusive (abnormal) behavior [9],[10].

The use of genetic algorithms (GA) to detect malicious computer behavior is a novel approach to the computer network intrusion detection problem presented in designing IDS. A genetic algorithm is a method of AI problem-solving based on the theory of Darwinian evolution applied to mathematical models. The GA designed for an experiment for attacking intrusion detection problem explained in [11] promoted a high detection rate of malicious behavior and a low false positive rate of normal behavior classified as malicious.

There are many studies in field of intrusion detection which we have briefly represented them. In [12], the authors have designed an Intrusion response (IR) system cooperating with IDS using mobile agents distributed throughout the network, based on stigmergic properties. In [13], the authors introduced a self-organized ant colony based intrusion detection system (ANTIDS) to detect intrusions and compares its performance with linear genetic programming (LGP) [14], Support vector machines (SVM) [15] and Decision Trees (DT) [16]. Other works have made use of Multiple Adaptive Regression Splines (MARS) [17]. In [18], the authors have compared various data mining algorithms for detecting network intrusions. The authors have used Naïve Bayes algorithm in building a network intrusion detection model [19]. In [20], the authors proposed Bayesian Belief network (BBN) with genetic and simulated annealing local search in order to build an efficient network intrusion detection model. The authors have compared various ensemble algorithms in detecting the intrusion detection in [21]. Modeling intrusion detection system using hybrid intelligent systems is proposed in [22]. In this, DT and SVM are combined as a hierarchical hybrid intelligent system model (DT_SVM) and an ensemble approach combining the base classifiers. In [23], the authors propose support vector learning approach to classify network requests. The authors in [24] used K-Means and DBSCAN to demonstrate how cluster analysis can be used to effectively identify groups of traffic that are similar using only transport layer statistics. The authors propose hierarchical Gaussian Mixture Model (HGMM) a novel type of Gaussian Mixture which detects network based attacks as anomalies using statistical processing classification in [25]. In [26], the authors use automated feature weighting for network anomaly detection. They conclude that their proposed method not only increases the detection rate but also reduces false alarm rate as well.
In this paper the authors have proposed an approach to intrusion detection by Expectation Maximization algorithm. In this algorithm, we have used of voting process and gathering information about thenodal behaviors in clustered wireless network.

The organization of the paper is as follows. After the introduction and related works in section 1, section 2 talks about the proposed algorithm and the principle of our Solutions for intrusion detection. This is followed by the evaluations of the proposed approach through simulation experiments which are presented in section 3. Finally, conclusions are presented in section 4.

2. PROPOSED ALGORITHM

In the proposed algorithm, we intend to detect abnormal behaviors in the network. The detection of misbehaving nodes consist of two important stages: The first stage is the detection of suspect nodes and second stage is to determine the correction of the suspicion about the suspect node. We have used of EM algorithm for second stage. We define misbehaving nodes as those that have aberrations in data exchange patterns.In our proposed method, each node monitors and measures the behavior of its neighbors and the other nodes which are located in communication path in one cluster. In other words, each node measures the functionality of its inter cluster neighbors which that node wants to communicate withthem. Then the node saves these information in SNB table.

The next stage in detection of misbehaving nodes is performed based on voting process. Therefore, when one node suspects to other node in same cluster, then it sends a request message for voting to the cluster-head. The cluster-head allow for voting from the inter-cluster nodes based on the score of nodal behavior (SNB). Afterward, the cluster-head decides based on the majority of collected votes' response. So, if many of participating nodes in the voting process give a positive vote to suspect node, then the cluster-head adds suspect node to the black list and asks of all nodes in its cluster to avoid of it in further communications. Otherwise, the suspect nod is a well-behaving node and the other node which give the negative votes to it, should update their SNB values for the suspect node.

In our algorithm, each node consists of the following three main components: data collection module, local aggregation and fusion module (LAFM) and intrusion response module.

A. Data Collection Module

The functionality of the data collection module is to collect security related data via monitoring local activities and behaviors of neighbor nodes.In this scheme, the information about the functionality of nodes in forwarding packets is sent to the cluster-head. Cluster-head decides based on two stage E and M according to EM algorithm.

Moreover, for using EM algorithm and constructing the complete-data sample (X), incomplete-data space (Y) anda parameter (θ), we define a bucket as a specific count of packets that are transmitted between two nodes [27]. Data collection module is formed of three modules: monitor module, packet trade module and data transmission quality (DTQ) module. In the following we explain detailsof these three modules.

1) Data Transmission Quality

We use the Data Transmission Quality (DTQ) function proposed by Tao Li et al. [14] to measure a node's communication quality. Each node calculates DTQ value for its neighbor which the node is trying to transmit data through them. DTQ calculates as follows.

$$DTQ=k\times\frac{D\times STB()}{E\times P()} \quad (1)$$

In the equation (1), E is the total cost for transmitting a bucket. D is the total packets that are transmitted successfully. P() is the probability of successful transmission of information when all nodes work together normally. STB() shows the stability of nodal behavior and calculates as equation (2).

$$STB=\left[\frac{S(d,u)}{L(d,u)}\right]^{\alpha}=\left(\frac{\sum_{j=0}^{N}\left(\frac{d_j}{u_j}\right)}{\sum_{i=0}^{M}\left(\frac{d_i}{u_i}\right)}\right)^{\alpha} \quad (2)$$

In this equation, M is the total number of buckets that one node transmits them through all paths. N is the number of buckets which one node sends them on a specific path. Also, d_i and u_i present the number of bytes that have been sent successfully and the number of bytes that have been trying to send them respectively. So, S(d,u) shows the sum of successful transfer rateof N last packets and L(d,u) shows the sum of transfer rate of M sent packets from the past. Finally, STB() can be calculated from equation (3).

$$STB()=\left(\frac{\text{Total ACKed messages for the last N messages}}{\text{Total ACKed messages for the last M messages}}\right)^{\alpha} \quad (3)$$

2) Monitor Module

In routing protocols (such as DSR) the routing information is defined in the source node. In the other words, the source node puts these information in the header of the packets and sends them to the destination node. Due to the nature of wireless sensor networks, if node A is located in the radio range of node N, it can sense the sent or received traffic by the node N. Then there is the possibility to listen to the next hop transmission (Figure 1). Hence, we can use this feature of wireless sensor networks to monitor the neighbors in forwarding packets.

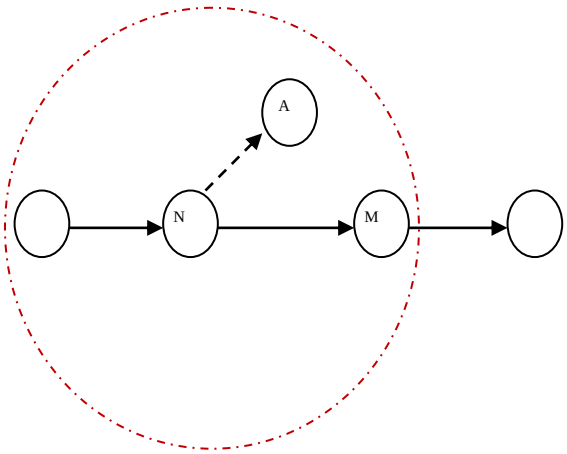

Fig. 3. Monitoring neighbor nodes in promiscuous mode

If the overheard packet from node N can be match with the buffered packet in A, then node Asure thatN has transmitted the packet to the next hop. Then A deletes the packet form its buffer.

Moreover, If this adaption is not done after a certain period, then the monitor module reduces the rating I of node N. We have defined the rating as follows.

$$R=\frac{\sum \text{packet}_{\text{forwarded}}}{\sum \text{packet}_{\text{actual}_{\text{resived}}} - \sum \text{packet}_{\text{destination}}} \quad (4)$$

the packets and also accepts the packets as the destination node.

B. Local Aggregation and Fusion Module (LAFM)

The output of LAFM module is a table called SNB. Each node has a SNB table in which each entry of this table maintains the SNB values of neighbor nodes in same cluster. This module calculates the SNB function of each node based on three functions DTQ, R and CP as follows.

$$SNB_i=\times\mu DTQ_i\times\delta+R_i+£\times CP_i \quad (6)$$

In this equation, the coefficients μ, δand η are the weights of functions DTQ_i, R_i and CP_i respectively.

Each SNB table value is updated at the end of each bucket. Then this table is sent to the detection engine module. The detection engine module decides based on the EM algorithm.

C. Detection Engine Module

After the detection engine received the updated SNB table from the LAFM module, the detection engine identifies the misbehaving nodes. In most intrusion detection approaches based on anomaly is defined a

In the equation (4), $\sum packet_{actual\_received}$ is the total number of packets which the node has received. $\sum packet_{destination}$ is the total number of packets which the node is destination of those packets. $\sum packet_{forwarded}$ is the total number of packets that the node has forwarded them as the intermediate node.

3) Packet trade module

In this subsection, we have introduced a mechanism for cooperating nodes together. This mechanism not only can detect misbehaving nodes, but also creates the motivation for cooperating nodes together. In this mechanism, the rate of cooperation (CP) calculates as equation (5).

$$\begin{aligned} CP &= \lambda \sum \text{packet}_{\text{forwarded}} \\ &+ \gamma \sum \text{packet}_{\text{destination}} \\ &- \varphi \sum \text{packet}_{\text{originated}} \end{aligned} \quad (5)$$

In this equation $\sum \text{packet}_{\text{forwarded}}$is the total number of packets that the node has forwarded them as the intermediate node. $\sum \text{Packet}_{\text{destination}}$is the total number of packets which the node is destination of those packets. $\sum \text{Packet}_{\text{originated}}$is total number of packets which the node creates and transmits them as source node.

According to the equation 5, forthe CP value be positive, one node has to cooperate and forward threshold to distinct normal and abnormal behaviors. So, to determine threshold and to classify the nodes' behavior, we have used of EM algorithm.The EM algorithm uses a maximum likelihood type parameter estimation approach and is ideally suited for cases in which the available data set is incomplete. It is an iterative procedure which under given certain conditions converges to values at a local or global maximum of the likelihood function. EM is a general technique of estimating features of a given data set applicable when analyzed data are incomplete or have missing values. It is an iterative algorithm, where there are two steps for each iteration– expectation I step and maximization (M) step.

The following notation is useful for the further discussion about the EM algorithm:

X – set of all available features of all observed samples;
Y – set of all unknown features of all observed samples;
θ^i – estimate of distribution parameters at the i-th stage of the algorithm iteration;
θ – variable for the new estimate describing the (full) distribution.

The aim of the EM algorithm is to maximize the expected log-likelihood of the model parameters set θgiven the joint distribution of the observed data X and the missing dataY. For a set of independent samples X, drawn from a single distribution described with the set of parameters θ, the likelihood function can be thought of as a function of the parameters θwhere data X is fixed, i.e. it givesthe likelihood L of the data X given distribution parameters θ.

$$L(\theta|X)=p(X|\theta)=\prod_{n=1}^{N} p(X_n|\theta) \quad (7)$$

In order to find the set of parameters θ that maximizes the likelihood, the usual approach is to avoid direct likelihood maximization and to try to maximize the logarithm of the likelihood function.

In the proposed algorithm, the set of observed samples has defined as follows.

$$X = \{SNB_1, SNB_2, \ldots, SNB_n\} \quad (8)$$

As we have mentioned before, X represents the set of N observed samples from this mixture. Each sample X from this set is a multi-dimensional authentication data vector X = [x_1, x_2, …,x_D] where D represents dimension of the sample. Moreover, SNB_i represents the observed data set of all available samples. In the other words, SNB_i shows the data of neighbor nodes of misbehaving node. Nevertheless, there are some nodes in the cluster which can participate in the voting processbut according to the long-term bucket statistics, these nodes are not in the neighboring of suspect node.The gathered statistics for table that are less than the threshold, then it realizes that there may be one or some misbehaving nodes in its cluster. So, EM algorithm will detect misbehaving nodes in a cluster based on received DTQ and SNB values.

To calculate posterior probability we must know the prior probability $P(\omega_k)$ which specifies the initial probability – an observed feature vector belongs to a class $\omega_k$as well as probability density function $P(X|\omega_k)$ as follows.

$$P(w_k|X)=\frac{P(X|w_k)P(w_k)}{P(X)} \quad (11)$$

This is the probability for data X given the class ω_k also known as class-conditional probability. P(X) represents the sum of posterior probabilities across all classes:

$$P(X)=\sum_{k=1}^{K} P(X|w_k)P(w_k) \quad (12)$$

Its value is therefore same for all posterior probabilities considered. As such it can be neglected during $P(\omega_k|X)$the classification and the classification rule can therefore be simplified to the assignment of vector X to a class ω_k if:

SNB values of these neighbor nodes create set of all unknown features (Y).

We have adopted a finite mixture model for our IDS which assume that the authentication data arises from two groups with known and unknown parameters. The following probabilistic model is used to describe this situation.

$$P(X|\theta)=\sum_{k=1}^{K} \alpha^k p_k(X|\theta) \quad (9)$$

where the unknown parameters $\theta = (\alpha_1, \ldots, \alpha_k, \theta_1, \ldots, \theta_k)$ are such that $\sum_{k=1}^{K} p_k$& each p_k is a probability density function parameterized by θ_k. Thus, Kcomponent densities are mixed together using K mixing coefficientsα_k. The task is to classify the data vector to K different classes adopting the Bayes classification rule to assign the particular sample X to a class with the highest posterior probability $P(\omega_k|X)$. for detection the classes, we have defined the a threshold for each cluster as follows.

$$Th=\tau\times\frac{1}{|N_{SNB}|}\sum_{i\in N_{SNB}} S_i \quad (10)$$

In equation 10, N_{SNB}shows all listed nodes in the SNB table of the cluster-head node. So $|N_{SNB}|$ is the number of the nodes, and S_i shows the value of SNB related to the node i.

The values more than Th show the suspect node is a well behaving node in the cluster. But, if the detection engine module finds one or some values inthe SNB

$$P(w_k|X)>P(w_j|X) \text{ for } j=1, 2, \ldots, K \& j\neq K \quad (13)$$

It is known that this classification rule yields a classifier with the minimum probability of error.

The major problem in applying the Bayesian classifier is the estimation of class-conditional probability density function $P(X|\omega_k)$ as parameters of that class need to be known or if not known estimated from the measured data set using some of the available parameter estimation methods.

The aim of the EM algorithm is to maximize the expected likelihood of the model parameters set θ given the joint distribution of the observed data X and the missing dataY. For a set of independent samples X, drawn from a single distribution described with the set of parameters θ, the likelihood function can be thought of as a function of the parameters θwhere data X is fixed, i.e. it givesthe likelihood L of the data X given distribution parameters θ.

$$L(\theta|X)=p(X|\theta)=\prod_{n=1}^{N} p(X_n|\theta) \quad (14)$$

The unobserved part of data set Y for our system can be considered as knowledge of which component k produced each sample X_n. In order to find the set of parameters θ that maximizes the likelihood, we have considered a binary vector $Y_n=\{y_{n,1}, \ldots, y_{n,k}\}$ for each sample of X_n where $y_{n,k}=1$ if the sample was produced by the component k or zero otherwise. The likelihood function to be maximized by the EM algorithm considers the joint distribution of the observed data X and the missing data set Y and it can be shown that:

$$L(\theta \mid X, Y) = \sum_{n=1}^{N}\sum_{k=1}^{K} y_{n,k}\alpha_k (p(X_n \mid \theta_k)) \quad (15)$$

In the E-step, the EM algorithm calculates the conditional expectation of the complete data likelihood function,

given X and the current estimate $\theta=\theta^i$ of the parameters. It therefore evaluates the function:

$$Q(\theta|\theta^i)=L(X, w|\theta) \quad (16)$$

As we have mentioned before W represents the conditional expectation of Y with respect to set of observed data and parameter θ , i.e., W = E[Y| X, θ]. Elements of W are defined as:

$$w_{n,k}=P\left(y_{n,k}=1 \middle| X_n, \theta^i\right) \quad (17)$$

i.e. those are the probabilities that particular data sample X_n is produced by component k.

The M-step of the EM algorithm performs for estimating distribution parameters for K-component. This is performed by calculating initial probabilities of α' in the network.The E- and M-steps of the algorithm are repeated until a convergence criterion is met. The convergence criterion is usually chosen so that EM is interrupted when change in values evaluated at each algorithm iteration falls below a certain threshold value. The possible criteria in proposed algorithm is calculated in each step as equation (18).

$$\left|\Theta^{i+1}-\theta^i\right|<\varepsilon \quad (18)$$

When the condition in equation (18) is met and the information of any nodes in the cluster are received in the cluster-head, then the cluster-head will find a misbehaving node in its cluster.

D. Response Module

According to the results of detection engine module, if the node (m) is a well-behaving thenshouldbeacquitted. So, the intrusion response module realizes that the node m is a well-behaving one theother nodes that can use of node m for communications. Similarly, if the node m is misbehaving one then should be penalized.Moreover, a misbehaving node (m) is added to the black list. Thenthe detection engine module eliminates a misbehaving node from SNB and routing table. Thereafter, the intrusion response module prevents the other nodes from cooperating with the blacklisted node.

3. EVALUATION

In this section, we have implemented the proposed algorithm by Glomosim simulator [27][28], a scalable discrete event simulator developed by UCLA. Moreover, we have evaluated the attack impacts on the proposed algorithm and also measured the energy consumption of protocol. Finally, we have evaluated the results.

Simulation Settings

We have used the Distributed Coordination Function (DCF) of IEEE 802.15.4 for distributed wireless sensor networks as the MAC layer protocol. The routing protocol is Dynamic Source Routing (DSR) protocol. The network area size is 2000*2000 (in m2). The mobility model is the random waypoint model. The minimum speed is 5 m/s, and the maximum speed is 20 m/s. The number of nodes varies from 10 to 100 nodes. Radio bandwidth is 250000(in bps). Initial energy level of each node is 5(mW) and radio transmit power is 10 (in dBm). Various source-destination pairs are selected randomly to generate Constant Bit Rate (CBR) traffic as the background traffic. The size of all data packets is set to 512 bytes. The duration of each simulation is 1800 seconds.

Simulated Attacks

In this article, we have simulated four attacks to evaluate the functionality of our IDLAP which are Flooding, Black-hole, Gray-hole and Denial of Service (DoS).

- Flooding attack [28]: In this attack, the misbehaving node pumps a great deal of useless and garbage packets to the network. In this way, it corrodes the resources of the network such as bandwidth and energy.
- Black-hole attack [29]: In this attack, a misbehaving node uses the routing protocol to advertise itself as having the shortest path to the node whose packets it wants to intercept. The attacker will then receive the traffic which is destined for other nodes, and then it can drop or modify the packets.
- Gray-hole (selective forwarding): This is a special kind of black-hole attack. Contrary to the black-hole attack, the gray-hole attack is used to attack a traffic based on a probability function or is used to attack specific nodes' packets, but not all the received packets.
- Denial of Service (DoS) attack [30]: In this attack, the misbehaving node prevents other nodes from

cooperating by depleting the resources of the network. When the resources of the node, especially the energy, is depleted and the resources reach a specific threshold, the node prevents itself from cooperating with other nodes so that it can increase its life time.

Performance Metrics

- Detection Ratio: It is defined as the percentage of IDS capability in detecting the misbehaving nodes. And it is resulted from dividing the accurate detections into all detections.
- False Positive Ratio: It is defined as the percentage of decisions in which well behaving nodes are flagged as misbehaving ones inaccurately.
- Average Energy Consumption: It is gained from dividing the sum of nodes' energy consumption to total number of nodes.

Simulation Results

In the first simulation, we have considered the relation between the detection rate and number of the nodes. These results is showed in figure 4.

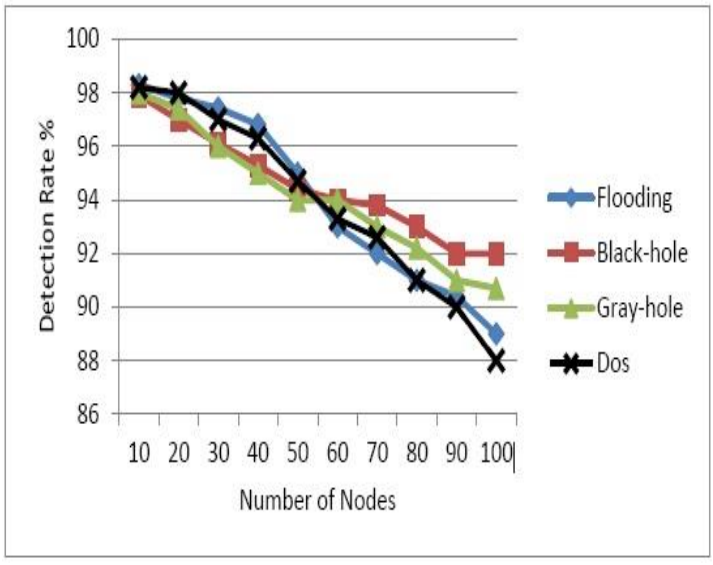

Fig. 4. Detection rate vs. number of nodes

In detecting Gray-hole, Back-hole, DoS and the Flooding attack, the detection rate has decreased with increasing number of nodes in the network; the reason is that increasing number of nodes in the network makes to increase number of clusters. So, when the number of clusters increases, inter-cluster communications will increase too. Due to the increasing inter-cluster communications, this is possible that the cluster-head makes mistake to detect malicious nodes. However, these measurements are lower for Flooding and DoS attacks than the others. In these two attacks, proposed method has had a good stability, and its detection rate has been over 88%.

The results of the simulation are shown in figure 5 which is false positive versus the number of nodes.In networks with high densities, the false positive rate is high and this is so natural. In Flooding attack with high densities, false positive rate is higher than the others. In this case, the amount of traffic rate is so high in the network and it is not possible to identify whether the high rate of traffic is due to flooding attack or normal situation of the network. In Flooding attacks, the energy level of each node decreases significantly and the available data set of observed sample decreases. But using of EM algorithm in proposed method causes to the false positive rate be uniform in all attacks.

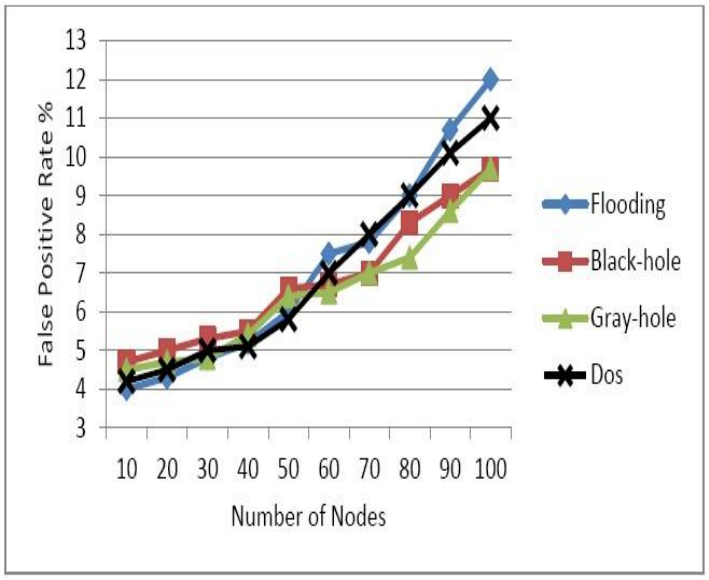

Fig. 5. False positive vs. number of nodes

Next, we have investigated the effect of misbehaving nodes percentage as parameters on the performance of proposed method.The number of nodes in these simulations is 100.In figure 6, results of the simulation show the detection rate with variation in the percentage of misbehaving nodes. As it can be seen, most of the misbehaving nodes have been detected successfully. In Flooding and DoS attacks with high percentages of misbehaving nodes, the detection rate decreases. This happens due to different reasons and the most important reason is the process of detection, because the EM algorithm performs based on the received data samples in a specific path. But in Black-hole and Gray-hole attacks, detection rate initially decreases and then increases because the set of all available features of all observed samples are complete.

International Journal of Information and Communication Technology Research

http://www.esjournals.org

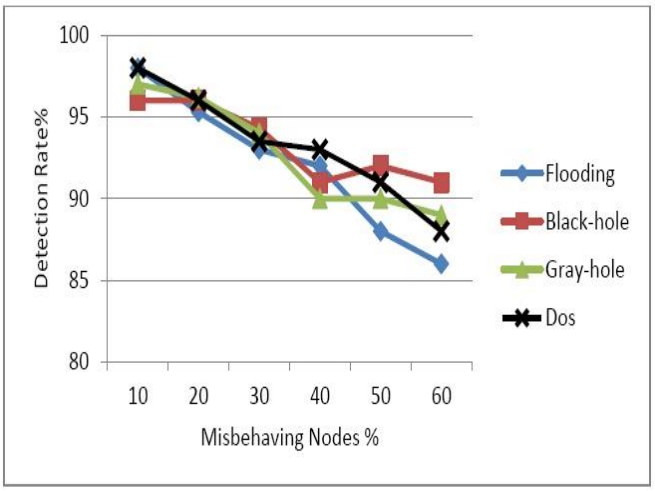

Fig. 6. Detection rate vs. percentage of misbehaving nodes

The results of the simulation in figure 7 show false positive rate with variation in the percentage of the misbehaving nodes.As it can be seen in this diagram, with increasing the percentage of misbehaving nodes the false positive rate increases as well. This is predictable because the detection rate decreases as it has been illustrated in figure 6. Also the deduction of the detection rate causes the false positive to increase. As the quantity of misbehaving nodes increases, we can have more stable false positive rates.

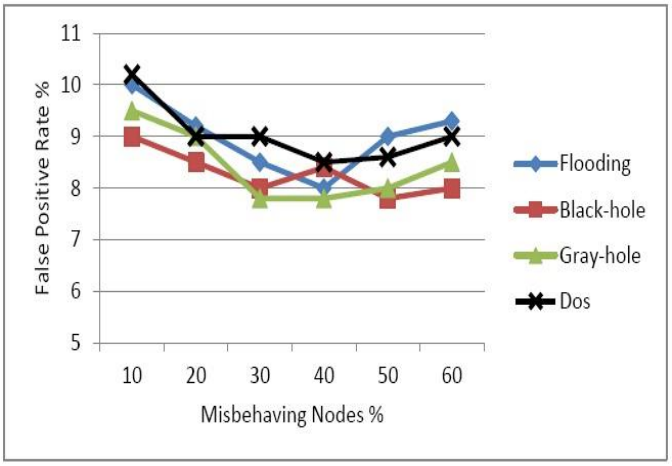

Fig. 7.False positive rate vs. percentage of misbehaving nodes

In the next simulation, we have discussed detection rate with variation in the number of clusters. As it is exhibited in figure 8, the best result is gained when the number of clusters is 15. Because when there are a few numbers of clusters, there are a few number of cluster-head nodes, correspondingly. So, the confirmation of the misbehaving nodes detection can be difficult. But when the number of clusters increases, detecting the misbehaving nodes can be easier. In greater numbers (more than 15) we face with lower number of nodes in each cluster; therefore, this is difficult to detect misbehaving nodes. In fact, the data set of observed samples in EM algorithm is reduced. So the number of clusters should not exceed a specific criterion that is 15 in this example.

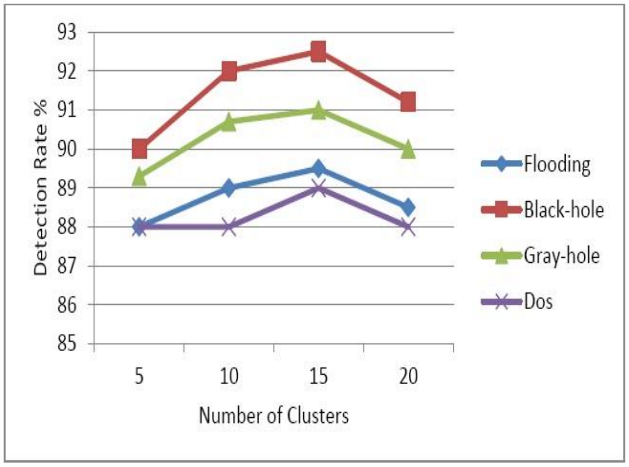

Fig. 8.Detection rate vs. number of clusters

In figure 9, the results of simulation show false positive with variation in the number of clusters.In the figure 9, the rate of false positive is high at the first, and the best results occur when there are 20 zones approximately. As the number of zones increases, the number of observed sample set reduces. Therefore, not only the detection is more difficult, but also the number of unknown features of all observed samples decreases. In fact, by reducing the number of nodes in each cluster, the set of all known features of observed samples decreases.

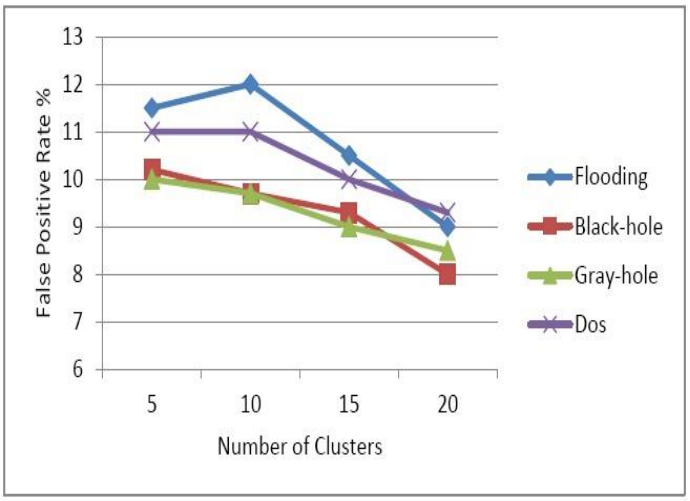

Fig. 9. False positive rate vs. number of clusters

The results of energy consumption are shown with variation in number of nodes in figure 10. The Flooding and Dos attacksare simulated because They have a lot of control messages in the network.It is clear from Figure 10 that when there are 100 nodes, the maximum energy consumption is 212 (in mWhr) in flooding attack. The energy level of each node is very good criterion for the life of the network. It is clear that the energy level of nodes will not be zero and this indicates that the network lifetime is acceptable. So the network will not be

partitioned and will continue to operate. This means the network's lifetime will increase

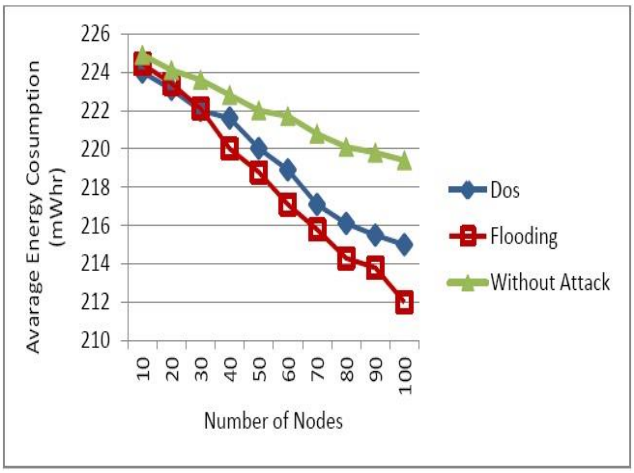

Fig. 10. Average energy consumption in the network vs. number of nodes

4. CONCLUSIONS

In this paper, we have proposed a protocol for the intrusion detection problemfor clustered wireless sensor networks, in which the detection of malicious nodes were organized by cluster-heads using expectation maximization algorithm. To design this protocol, the collected statistics by the cluster-heads is used for EM algorithm to verify the collected data set and so the algorithm can detect misbehaving nodes. We have carried out a simulation and demonstrated its effectiveness.

REFERENCES

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey onsensor networks, IEEE Communications Magazine (August) (2002)102–114.

[2] N. Bezroukov, "Intrusion Detection (general issues)." Softpanorama: Open Source Software Educational Society. 19 July 2003.

[3] T. Crothers, "Implementing Intrusion Detection Systems", ISBN: 0-7645-4949-9, John Wiley & Sons, Inc. 2003.

[4] S. Northcutt & Judy Novak, "Network Intrusion Detection", 3/e, ISBN: 9780735712652, Sams Publishing 2005.

[5] E. Guillen, D. Padilla, Y. Colorado, "Weakness and Strength Analysis over Network-based Intrusion Detection and Prevention System", 2009, IEEE Latin American Conference on Communications, 2009. LATINCOM '09, ISBN 978-1-4244-4387-1.

[6] N. Foukia, S. Hassas, S. Fenet and J. Hulaas, "An Intrusion response scheme: Tracking the source using the sigmery paradigm", in Proc. Of security of mobile multi agent system workshop (SEMAS-2002).Italy, July 16, 2002.

[7] R. Hanson and J. Stutz "Bayesian Classification Theory" Technical Report FIA-90-12-7-01. Artificial Intelligence Research Branch NASA Ames Research Center, Mail Stop 244-17, Moffet Field, CA 94035, USA.

[8] W. Lee and S. Stolfo, "Data mining approaches for intrusion detection," in Proceedings of the 7th USENIX security symposium, (San Antonio, TX), 1998.

[9] J.Gomez, F.Gonzalez and D.Dasgupta, "An Immuno-Fuzzy Approach to Anomaly Detection". To appear in the Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZIEEE) May 25-28, 2003.

[10] J.Gomez and D.Dasgupta, "Evolving Fuzzy Classifiers For Intrusion Detection". To appear in the Proceedings of the 2002 IEEE Workshop on Information Assurance, June 2002.

[11] A. Chittur "Model Generation for An Intrusion Detection System using Genetic Algorithms" Ossining High School, NY NOV, 27, 2001.

[12] N. Foukia, S. Hassas, S. Fenet and J. Hulaas, "An Intrusion response scheme:Tracking the source using the sigmery paradigm", in Proc. Of security of mobile multi agent system workshop (SEMAS-2002).Italy, July16, 2002.

[13] V. Ramos and Ajith Abraham, "ANTIDS-Self organised Ant based clustering model for intrusion detection system. WSTST, 2005, pp.103-112.

[14] M. Brameier and W. Banzhaf, "A comparison of linear genetic programming and neural network in medical data mining", in IEEE Transaction on Evolutionary computation.5 (1), pp.17-26, 2001.

[15] V.N. Vapnik, The Nature of Statistical Learning Theory. Springer, 1995.

[16] J. Denebourg ,et al. "The dynamic of collective sorting Robot like ants and Ant like Robots", in 1th conf. on simulation of Adaptive behavior ;from animals to animats,cambridg,MA,MITPress,pp.356-365,1991.

[17] S. Mukkamala, A.H. Sung, A. Abraham, V. Ramos, "Intrusion Detection Systems using Adaptive Regression Splines", in ICEIS-04, 6th Int. Conf. on Enterprise Information Systems, to appear at

Kluwer Academic Press, 2005, Porto, 14-17 April 2004.

[18] M. Panda and M. R. Patra, "A comparative study of data mining algorithms for network intrusion detection", proc. of ICETET, India, 2008.pp.504-507. IEEE Xplore.

[19] M. Panda and M. R. Patra, "Network intrusion detection using Naïve Bayes", International journal of computer science and network security, vol.7, No.12, 2007, pp.258-263.

[20] M. Panda and M. R. Patra. "Bayesian belief Network using genetic local search for network intrusion", International journal of secure digital information age.Vol.1, issue.1, June 2009. In Press.

[21] M. Panda and M. R. Patra, "Network intrusion detection using boosting support vector classifiers", In 2009 IEEE Intl.Advance computing Conference, ptaila, Punjab, pp.926-931. IEEE Press.USA.

[22] S. Peddabachigari, A. Abraham, C. Grosan and J. Thomas, "Modelling intrusion detection system using hybrid intelligent systems", Journal of Network and computer applications.Elsevier, 30(1), pp.114-132, 2007.

[23] J. Mill and A. Inoue, "Support vector classifier and network intrusion detection", In proc. of 2004 IEEE international conference on fuzzy system, WA, USA, 2004, Vol.1, pp.407-410. ISBN: 1098-7584.

[24] J. Erman, M. Arlitt and A. Mahanti, "Traffic classification using clustering algorithms", in SIGCOMM-06 workshops, sept.11-15, 2006, Pisa, Italy.pp.281-286.ACM Press.

[25] M. Bahrololum and M. Khaleghi, "Anomaly intrusion detection system using hierarchical gausian mixture model", International journal of computer science and network security, vol.8, no.8, pp.264-271, August 2008.

[26] D. Tran, W. Ma and Dharmendra Sharma, "Automatedfeature weighting for network anomaly detection", in International journal of computer science and network security, Vol.8, no.2, pp.173-178, Feb.2008.

[27] K. Kumar, "Intrusion Detection in Mobile Adhoc Networks", Master's Thesis, Advisor Dr. MansoorAlam. The University of Toledo, December 2009.

[28] C. Hongsong, F. Zhongchuan and et al,s "Using Network Processor to Establish Security Agent for AODV Routing Protocol", Journal of Computing and Information Technology – CIT 15, pp. 61–70, 2007.

[29] F. Anjum and P. Mouchtaris, "The Handbook of Security for Wireless Ad Hoc Networks", (Chapter 1), CRC, Press LLC, 2007.

[30] C. Rong and E. Çayırcı, "Security Attacks in Ad Hoc, Sensor and Mesh Networks, in Book Security in Wireless Ad Hoc and Sensor Networks", (Chapter 8), CRC Press LLC, 2009.