An Evaluation of Effective Intrusion DoS Detection and Prevention System Based on SVM Classifier for WSN

Prabakaran K¹, Kumaratharan N², Suresh G³ and Epsiba P⁴

¹ Abdul Hakeem College of Engineering and Technology, Tamil Nadu, India
² Department of ECE, Sri Venkateshvara College of Engineering, India
³ Department of ECE, Sri Indu College of Engineering and Technology, India
⁴ Department of ECE, Pallavi Engineering College, India
¹prabuswamy@yahoo.com, ²kumaratharan@rediffmail.com, ³geosuresh@gmail.com, ⁴epsisuresh@gmail.com

Abstract. The trending technology of Wireless Sensor Networks (WSNs) are broadly used in this recent earth for a greatest rate of purpose for the better worth of person living however owing to their energy constrains as well as utilization and problems as mentioned above, these WSN has their restriction in their intention as well as their practicality. The wireless sensor networks face main threat challenges such as Denial of Service (DoS) attack, flooding attack, replay attack. In this approach, we proposed an Intrusion Detection (ID) technique using Support Vector Machine (SVM). In the ID algorithm, we employ feature vector to constitute nodes. Such components consist of routing message number, varies destinations in sending and receiving routing packets, and similar source node. The similar source node consists of similar characteristics; the node of the similar category has the same features. Therefore, irregular nodes will be notable. The measurement result shows an introduced scheme attain higher rate of accuracy. Here, Ant colony Optimization (ACO) technique to choose best route between entire adjacent feasible routes from source to receiver node for data communication. The proposed approach enhances the network lifespan as well as discovers intrusion nodes in WSN.

Keywords: Ant colony Optimization, Denial of Service Attack, Intrusion Detection, Support Vector Machine, Wireless Sensor Networks.

1. Introduction

Generally, WSNs consisting of quantitatively great amount of dispersed sensor nodes interconnected together to examine and determine the status of real atmospheric conditions. The evaluated variables are communicated to various distributed end users through Wireless Network (WN). WSNs are miniaturized, power consumption is considerably good and have been extensively established for real time applications as compared to the existing networks. Even then, it faces lot of storage constraints, message delivery distance, processing complexity and energy harvesting. Conversely, embedded system plays a major role in the past decade. These are very small in size, cheap, consumes less power and they had many sensors which works on various environment [3]. A typical WSN consists of various sensor components. These components combined to a simple small device and this is implemented into various application scenarios [1] [2]. These advance specifications make them to be flexible when compared to
the ordinary traditional networks. The applications of WSNs are abundant: they are extremely scalable and a high density range, due to their cheap rate and their smaller size [15], [8]. Normally, all the sensor network protocols are designed with their own organizing capacity. Hence these sensor network has an effective maintenance and they have forbearance against communication failure and the topology could change based on the nodes that could break down, or be mobile or there can be depletion of energy. Moreover, sensor nodes could be implemented in the rough environment to maintain the given task without any man power.

Moreover, supplementary applications are inadequate due to some disadvantages of WSN. For instance, the restricted giving out capability and minimum rate of data in the nodes of the sensor couldn’t warranty the maximum performance in some situations particularly for instantaneous applications [16]. When the communication range is very small, then there could be maximum energy consumption and network incompetence, since the requirement of multi-hop communications is often used for the transportation of data [17]. The high drive restriction also direct to bad inspection: the development of data processing, as the energy gets drained rapidly and makes the network ineffective for the particular job. Low amount of energy consumed maintains multi-hop communications work for a maximum duration.

WSN has become most favoured networking solution, though; it’s tremendously incomplete resources face serious threats. In addition, the nodes of sensors are very sensitive to some attacks which could be caused due to its lack of tamper-resistance. Several threats which occur in WSN are DoS attack, differential power attack, node balancing attack, Scheme-dependent attack as well as imaging attack [7], [11]. DoS (Denial of Service) attack is typically anticipated attack of opponent for wiping out the sensor network. It limits or eliminates the anticipated performance of the sensor network than expected. DoS attack intrudes the efficiency of the target networks by consuming the resources, destructing infrastructure configuration, and the network components.

1.1 Architecture of WSN

As a reliable technology, the WSN has been employed widely in other technologies such as Internet of things. Numerous associated uses have been performed. The function of the WSN is gaining reliability and fashionable in network. WSN is considered as most significant technology in the 21st century, WSN is considered as a vital component in linking the rationality entropy as well as the prevailing world. Though, the exposed working principle of WSN and entropy media as well as the deprived organization of surroundings possess many threats to the security as well as it is extremely limiting the claim of WSNs.

![Figure 1. Architecture of WSNs](image)

The trending technology of Wireless Sensor Networks (WSNs) are broadly used in this recent earth for a greatest rate of purpose for the better worth of person living however owing to their energy constrains as well as utilization and problems as mentioned above, these WSN has their restriction in their intention as well as their practicality. In spite of its accuracy and energy constrain analysis from real exploitation
can suffer restrictions, since the data acquired by the sensors from a wide area is very tough to obtain. On the other hand, analytical/theoretical models obtained from the mathematical calculation offers huge exploration capacity but it can simplify through idealized hypotheses which could cause trouble during the practical implementation.

2. Related work

WSN works in several environments which could consist of effective and limited processing sensor nodes that is low in cost, extremely small and consumes less power. This wireless sensor networks differs from the traditional network due to its wide range application specifications. The applications of WSNs are enormous: they are extremely scalable and have a high density range, extremely cheap in cost and their smaller size. In [13] proposed a technique on the basic principles of least cut set and less cluster head was carried out in a distributed intrusion detection module, with the help of signature based detection technique. Detection of black hole by centralized IDS procedure and the cluster-based wireless sensor networks that works by selective forwarding has been scheduled [6]. In [15] proposed partially distributed IDS that work on minimum power and less memory occupation. Detection carried by anomaly is also called as behavioral based detection in certain cases. Irregularities can be separated into three classes called point, appropriate as well as finally collective unusual person [12]. The WSN networks can be used as an alert system for the flood detection and it was suggested at Lancaster University [14] by a research. They made a protocol and did its functionality by a platform composed of powerful GridStix sensor [5] which has certain modules like GPRS to monitor the flood arrival and estimation in UK. This system was implemented at river ribble in the Yorkshire Dales which covered about 1 km. The system was deployed with 13 modes and the implementation is carried out as a real-time. The attacks in DoS based approaches affects the most important threads in OSI layer [4], [9] as described. More often these impairments are caused due to the system weakness or protocol defencelessness.

3. Support vector machine

In this context, SVMs [5] [1] are one of the supervised machine learning approach which employed for analysis. For given finite set of training samples, the SVM marks them for categories and thus it builds the training model which later assigns new test samples into the relevant category. It is binary linear and non-probabilistic classifier. In SVM model the training/test samples are represented as dot/points in the space and they are mapped in such a way that the clear gap among the categories appears which separates the samples. New test samples/examples are later mapped on the Using the kernel trick SVMs are also be able to perform non-linear categorization in addition to linear categorization. SVMs mechanically chart the remarks on towering multidimensional quality distances. In usual SVM defends to build anxious aircraft in any gap as well as this can be engaged for any jobs for example failure, forecast. SVM hyper plane is the only responsible for good separation of training data with the largest distance to its nearby data. This approach is generally known as functional margin. The state of the rule is if margin is larger than the generalization error of classifier will be lower. The characteristics of SVM are established to be exceedingly valuable for higher dimensional planes. It is highly preferable in less data samples provided as compared to the more dimensions. SVMs are flexible, memory efficient and which improves efficiency in critical decision making even more than one kennel method.

The important intend of rating the nodes is to precisely predict the objective class from the data in every container. The classifier engine procedure admits two treads:

1. Classifier construction: This approach is to construct a taking stage. The word that developed from tuition sure information examples is built through the categorization methods. Individuality instance which is collected of the training cluster is observed as a cluster. These tuples can as well as remarked to as data details.

2. Utilisation of Classifier – The training classifier rendered applying the data set will categorize test data preset objects.
Significance Analysis: Information might contain several distinct properties. Correlation analysis will recognize two specified properties are connected.

Data alteration as well as diminishing: The subsequent procedure assists in data communication:

Normalization: This communication concerns scaling each value which will build them fall inside a precise series. This approach is mostly utilized in the discovering stage while the neural networks otherwise these procedures developing evolution are utilized. Simplification: It is a main idea of communication. At this point we can employ the hierarchy idea. If the preparation nodes of n points is given and the method of it is \((-x\rightarrow 1, y_1)\) .... \((-x\rightarrow n, y_n)\) and here \(y_i\) is 1, this indicates the category to that the point \(-x_i\) is present. Each \(-x_i\) is a real vector of P-dimension. Our interest is to discover the “extreme- border hyper plane” which classifies the GoS/points since the group of points \(-x_i\) here \(y_i = 1\). It is understandable that H1 usually cannot distinguish the classes. While, H2 divide them with a small margin, on the other hand H3 depart them with the highest margin. The hyper plane is defined as the cluster nodes \(\rightarrow x\) fulfilling. The SVM consists the sample on the margin. The offset of the hyper plane from the origin with normal vector \(\rightarrow w\) is determined by the parameter \(\rightarrow w\). Hard-margin two parallel hyper planes which separate two classes of data can be selected, when the training data sets are linearly separable components. So, the distance between them is as possible as large. The “margin” is nothing but the area bounded by two hyper-planes. And the maximum margin hyper plane lies between these planes. Where, the hyper planes are in and \(\rightarrow w\) is the region distance of the two hyper planes, by minimizing \(w\) we can maximize the distance between the planes. By adding the constraint: for each either or , if \(y_i = -1\), hence the data points can be prevented from falling into the margin.

![Figure 2. Categorization based on linear SVM](image)

According to the constraints/conditions each and every data point should be the actual side of the region of two hyper planes. The rewritten equation is as observes:

\[
y_i (\rightarrow w \ast \rightarrow x - b) \geq 1, \text{ for all } i \leq \text{I} \leq n
\]

To get the optimization problem we can put all these together: "Minimize \(\rightarrow w\) subject to \(\rightarrow w \ast \rightarrow x - b \geq 1\) for \(i = 1, 2, 3., N\). Completely the maximum margin hyper plane is firmed by over right arrow(\(x_i\)) which lies nearer to the region. Here, the \(\rightarrow x\) is known support vectors.
Soft-Margin is the loss function with the equation $\max(0, 1 - y_i(-\mathbf{w}^* \cdot \mathbf{x} - b))$. Especially, this was introduced to enhance Support Vector Machine where the data are non-linearly separable. If the constraint in (1) is satisfied, then this function value is zero. The function with minimization is as below:

$$ h\text{igh}(0,1 - y_i (-\mathbf{w}^* \cdot \mathbf{x} - b)) + \lambda - \mathbf{w} $$

Here the parameter $\lambda$ finds the trade-off between increment the values of margin-size and making sure that the $\rightarrow x$ is fall/separated with the actual side of it. Therefore, the soft margin Support Vector Machine may behave equally same as hard margin Support Vector Machine for enough small values of $\lambda$ in the case when linearly classifiable test data are available (Yang, 2016).

4. Proposed intrusion detection

In the ID algorithm, we employ feature vector to characterize nodes. Such components consist of routing message number, varies destinations in sending and receiving routing packets, and similar source node. The similar source node consists of similar characteristics; the node of the similar case has the similar features. Therefore, irregular nodes will be notable. The feature weights and SVM parameter chromosomes are formed based on optimal feature subset. The best optimal chromosome is fixed by analyzing the chromosome with the greatest categorization precision of the chromosome; the best attributes as well as the characteristic weights are decrypted. The information is partitioned addicted to k sub-sets with the equal size as initial, next, ... k sub- parts, as well as the residual k-1 sub-parts as training data sets and for categorization the sub portions are combined k prediction results, where k=10 is chosen to evaluate the performance and to authenticate the robustness of the proposed SVM classifier [10].

**Figure 3.** Categorization based on Hard SVM.

**Figure 4.** The representation of SVM ID technique
We employed two important parameters, namely, limit value and score. The score represents total values of most adjacent nodes. Limit value represents the limit or threshold employed for classifying abnormal nodes from normal nodes. The RREQ (Route Request Packet), entails the target identity of all specified nodes of the network is spread. Thus, the SVM detection method categorizes from where irregular nodes transmit RREQ messages extra repeatedly than normal nodes. By connecting the incidence to transmit RREQ messages of every node, we can categorize as well as notice the irregular nodes. The characteristic collection marks the node with the incidence of RREQ messages. For the intention of categorization, it is essential to utilize a best quantity of characteristics for machine learning simulating. Afterwards a vast quantity of character elevate the calculation necessitates; it turns additional disputing to outline precise result borders inside a big multidimensional gap. This represents that a best separation of characters’ desires to be certain for the intention of machine learning. The achievement of a categorization method mostly based on the characters particular as well as the information they offer for their function in the replica. Throughout the categorization period, the rules of the incidence of RREQ messages arriving through the nodes are separated as whichever irregular otherwise regular nodes on the foundation of the character particular as well as employ assurance evaluates to specify which a flooding attack is happened.

The boundary value honestly impresses the discovery achieves. The apposite boundary value ought to create the error rate of recognition algorithm as small. For the unnamed allocation representation of the network, best possible boundary value is not received however during examination as well as information; we can get the choice of the appropriate boundary value.

**Proposed Energy Efficient Routing**

In this strategy, we concern Ant Colony Optimization (ACO) procedure that detects the shortest route among source to receiver. Here, an ant tries to determine the smallest distance to face from nest to food source. Before route seeking we have occupied several initialization factors as traces. The n is the surrounding phenomenon which illustrates the lesser quantity of pheromones in the communication route. The conversion possibility is utilize for an ant a to development from sender node x to receiver node r that denotes the routing information. The possibility decision rule.
\[ P_{\omega}^r \left[ \lambda_{\omega} \right] \left[ \beta_{\omega} \right] \frac{1}{\sum \left[ \lambda_{\omega} \right] \left[ \beta_{\omega} \right]} \quad \text{if } r \in L^\omega \]

Here, \( \lambda_{\omega} \) the quantity of pheromone as well as \( \beta_{\omega} \) is the heuristic function. The \( \gamma \) and \( \rho \) are the constant factor utilize to control the manipulate of the \( \lambda_{\omega} \) and \( \beta_{\omega} \). In ACO technique ants restrain the memory \( L^a \) that has the information about nodes which previously inflicted or not. The heuristic function \( \beta_{\omega} \) is the mutual of Euclidean distance \( d_{xr} \) among x as well as r receiver node. After closing of ant tours the quantity of pheromones updated consequently to the

\[ \lambda_{\omega} (t+ N) = (1- \theta) \cdot \lambda_{\omega} (t) + \Delta \lambda_{\omega} \]

Here, \( \theta \) denotes the local pheromone destroy factor, \( t \) represents the previous pheromone update time, pheromone updates at period of time \( t+N \). \( \Delta \lambda_{\omega} \) denotes the extra pheromone amount at the \( t + N \) point. This iteration procedure is repetitive till ants determine the finest potential route from source to receiver. The lifespan of the network computed through the energy utilization of the sensor node at best route as well a whole energy utilization of the ant \( r \).

\[ \text{Energy} = \sum_{x=1}^{k} \sum_{r=1}^{k} e_{\omega}^{r} W_{\omega}^{r} P_{\omega}^{r} \]

Here \( k \) is the amount of the network as well as \( e_{\omega}^{r} \) is the energy utilization among source to receiver. \( W_{\omega}^{r} \) is the available energy receiver node. If ant a has travel edge \((x,r)\) next the value \( P_{\omega}^{r} \) is 1 else 0.

5. Evaluation

So as to measure the function; we engaged the GAINZ nodes as well as terminal device ready on wired network card to obtain transmit blacklist as well as traffic. The terminal device is applied to rule manage method as well as check the irregular nodes, also reply to attacks. The Tiny operating system and AVR Studio integrated expansion surrounding as illustrated in figure 4. The serial transaction helper is applied to exchange the manage data message

![Figure 6. Representation of Test Scenario](image)

The flooding attack occurrence, amount of flooding attack node as well as regular node as attack attributes are arbitrarily particular. There is over flow in one node as well as flooding attack frequencies
are observed. The first case situation is accomplished under varies flooding attack frequencies on the equivalent amount of flooding attack nodes as well as regular nodes as well as irregular nodes. We selected three regular nodes and one flooding attack node.

Figure 7. Flooding attack frequency varies

Now, when there are 60 RREQ packets every second, there is change flow of 35 byte/s and when flooding occurs with 150 RREQ packets every second, the flow changes to 34 Bytes/s. We also observed that the number of normal nodes rises when the flow declines. The reason is that many nodes participate in progressing RREQ packets. Hence, with large number of nodes the bigger the effect in the network. We also evaluated and compared the measurable parameters to authenticate the function of the introduced method with Na"ive Bayes. The amount of True Positive(TP), False Positives(FP), False Negatives(FN), as well as True Negatives(TN) is offered via uncertainty matrix that contains two rowing as well as two columns. Such allows extra complete examination than insufficient fraction of precision. For every rowing in the uncertainty matrix means an investigational class, every column means an accepted class as well as every cell corresponds the amount of samplings in the association of those two classes. The uncertainty matrix is as given below:

Table 1. uncertainty matrix

|       | Yes            | No            |
|-------|----------------|---------------|
| Yes   | True Positive (TP) | False Negative (FN) |
| No    | False Positive (FP) | True Negative (TN) |

The true positives as well as true negatives categorizations are measured as precise categorizations which dwell alongside the mainly significant slanting line in the uncertainty matrix. The model errors signify the residual areas. We can obtain function attributes from the uncertainty matrix. At the end, the characteristic weights as well as measurable attributes of SVM are all together optimized based on feature subset. The experimental results exhibit that the proposed scheme speeds up the convergence, enhances the TP rate, reduces the error rate, as well as shrinks the categorization period. While comparing other schemes, SVM depend ID algorithm increase the detection rate and the FP as well as
FN rates are lowered. We can decide the precision in the %age of whole amount of guess is positive by employing equation 3:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\%$$  (3)

Where, True Positive (TP) rate is the %age of positiveness which is appropriately recognized, as determined employing the equation:

$$\text{Recall} = \frac{(T_P)}{(T_P + F_N)} \times 100\%$$  (4)

The Precision is the %age of the expected positiveness which is precise, as obtained by employing the equation:

$$\text{Precision} = \frac{(F_P)}{(F_P + T_N)} \times 100\%$$  (5)

Detection time: In order to recognize an attack in packet during detection, we employed detection time, and a short detection time is considered as better detection time. The assessment of all the classifiers applying the process of 10 folds annoyed establishment with attain attribute as well as lacking score characteristic is specified in the table below.

| Table 2. Comparison of classifiers employing the method of 10 folds cross validation |
|----------------------------------|----------------------------------|----------------------------------|
| Evaluation Criteria              | Classifiers                      | Classifiers                      |
|                                  | (Without Score Feature)          | (With Score Feature)             |
| Accuracy (%)                    | Naïve Bayes                      | Decision Tree                    | PC                   |
|                                 | 92.53                           | 95.52                           | 97.01                |
|                                 | Sensitivity (%)                 | 87.87                           | 96.96                | 96.96                |
|                                 | Specificity (%)                 | 97.05                           | 94.11                | 97.05                | 74.66                | 79.10                | 82.08                |
|                                 | MSE                             | 0.0139                          | 0.0108               | 0.009                | 0.0126               | 0.0174               | 0.0093               |

Following ten-fold cross confirmation is discover the greatest function in every classifier is 92.53% for Naïve Bayes, 95.52% for decision tree as well as 97.01% for Proposed Classifier (PC). It is discover that amongst classifiers, Naïve Bayes as well as Decision tree, the introduced classifier executes better with the categorization precision of 97% when tested with 10 folds cross confirmation. Furthermore, the model as well enhances categorization quality in terms of both sensitivity as well as particularity employing a better feature vector.
6. Conclusion
The WSN endure from several security disputes. In this paper, we introduced an ID technique established SVM. In the ID technique, we employ feature SVM vector to represent nodes. These components consist of routing message number, varies destinations in sending and receiving routing packets, and similar source node. The similar source node consists of similar characteristics; the node of the identical variety has the similar characteristics. Here, Ant colony Optimization (ACO) technique to choose best route between entire adjacent feasible routes from source to receiver node for data communication. The proposed approach enhances the network lifespan as well as discovers intrusion nodes in WSN. Therefore, irregular nodes will be notable. Hence, the proposed scheme achieved higher rate of accuracy from the evaluation results.

References
[1] Lim B and Uddin MS 2005 Statistical-based SYN-flooding detection using programmable network processor 3rd Int. Conf. on Information Technology and Applications (ICITA’05) 2 465-470
[2] Schuba CL, Krsul IV, Kuhn MG, Spafford EH, Sundaram A and Zamboni D 1997 Analysis of a Denial of Service Attack on TCP IEEE Symp. on Security and Privacy 208-223
[3] Karlof C and Wagner D 2003 Secure routing in wireless sensor networks: Attacks and countermeasures Ad hoc networks 1 293-315
[4] Cho EJ, Hong CS, Lee S and Jeon S 2013 A partially distributed intrusion detection system for wireless sensor networks Sensors 13 15863-15879
[5] Stavrou E and Pitsillides A 2010 A survey on secure multipath routing protocols in WSNs Computer Networks 54 2215-2238
[6] Butun I, Morgera SD and Sankar R 2013 A survey of intrusion detection systems in wireless sensor networks IEEE communications surveys & tutorials 16 266-282
[7] Akyildiz IF and Stuntebeck EP 2006 Wireless underground sensor networks: Research challenges Ad Hoc Networks 4 669-686
[8] Duraipandian M and Palanisamy C 2014 An intelligent agent based defense architecture for ddos attacks Int. Conf. on Electronics and Communication Systems (ICECS) 1-7
[9] Yi P, Zhu T, Zhang Q, Wu Y and Li J 2012 Green firewall: An energy-efficient intrusion prevention mechanism in wireless sensor network IEEE global communications conference (GLOBECOM) 3037-3042
[10] Yang Q, Fu H and Zhu T 2016 An optimization method for parameters of SVM in network intrusion detection system Int. Conf. on Distributed Computing in Sensor Systems (DCOSS) 136-142
[11] Mitchell R and Chen R 2014 A survey of intrusion detection in wireless network applications Computer Communications 42 1-23
[12] Malhotra S, Bali V and Paliwal KK 2017 Genetic programming and K-nearest neighbour classifier based intrusion detection model 7th Int. Conf. on Cloud Computing, Data Science & Engineering-Confluence 42-46
[13] Chandola V, Banerjee A and Kumar V, 2009 Anomaly detection: A survey ACM computing surveys (CSUR) 41 1-58
[14] Geng X and Whinston AB 2000 Defeating distributed denial of service attacks It Professional 2 36-42
[15] Yu Y, Li K, Zhou W and Li P 2012 Trust mechanisms in wireless sensor networks: Attack analysis and countermeasures Journal of Network and computer Applications 35 867-880
[16] Hemanth Kumar and Ramesh GP 2020 Energy efficiency and Data packet security for wireless sensor networks using African Buffalo Optimization IJCA 13 944-954
[17] Shahada SAA, Hreiji SM and Shamsudheen S 2019 IOT Based Garbage Clearance Alert System with Gps Location Using Arduino International Journal of MC Square Scientific Research 11 1-8