Intrusion Detection System based on SVM and Bee Colony

Monika Gupta
Department of Information Technology
SATI College
Vidisha, M.P. (India)

S. K. Shrivastava, Ph. D
Department of Information Technology
SATI College
Vidisha, M.P. (India)

ABSTRACT
An intrusion detection system (IDS) is an active process or device that analyzes system and network activity for unauthorized entry. Nowadays many intrusion detection systems are developed based on many different machine learning techniques. Some of the models are based on single classifying techniques while some models are based on combining different classifying techniques, such as hybrid or ensemble techniques. The basic task is to classify network activities (in the network log as connection records) as normal or abnormal while minimizing misclassification. Even if different classification models have been developed for network intrusion detection, each classification technique has its vitality and vulnerability. The machine learning based SVM method is a good choice for learning with little volume of data. Whenever new information is added into a system, updating of the old model is required immediately to ensure that the system is properly protected. As retraining may take weeks, or even months, it is impractical to retrain the new model on all available data. Thus, a mechanism is needed to generate an adaptive model that can be updated by cooperation of the old model with the new information. We can take advantage of the clustering based on Bee Colony in updating the models. We propose a new approach of combining SVM and Bee Colony to achieve high quality performance of Intrusion Detection System. Our algorithm is implemented and evaluated using a standard benchmark KDD99 data set. In this paper experimental result shows that SVM with Bee colony achieves an average accuracy is 88.46%.

Keywords
Intrusion Detection System (IDS), Data Classification, Machine Learning, Support Vector Machine (SVM), Bee Colony.

1. INTRODUCTION
Intrusion detection is the detecting of actions that attempt to compromise the integrity, confidentiality or availability of natural resources on the network. In case of an intrusion, an IDS (Intrusion Detection System) detects it as soon as possible and takes appropriate action. The composite of facts, such as the rapid growth of the Internet, the big financial possibilities opening up in electronic trade and the lack of truly protected systems, makes IDS an important front edge in research orientation of network security [1].

Although many different IDSs have been developed, their detection schemes generally fall into one of two categories: anomaly detection or misuse detection. Anomaly detector looks for behavior that deviates from normal use of system whereas misuse detectors look for behavior that matches a known attack scenario. The normal behavior is based on deficiency of innocuous factors and sumptuous variables. Therefore, the selection of features to monitor is the main issue in anomaly detection. The approach of misuse detection is to model abnormal system behaviors at first and define any other behaviors as normal behavior. Namely, known intrusion attacks are depicted in the form of pattern or signature, activity that match those attack scenarios can be detected and different suitable actions will be further taken for different types of intrusions. The main issue in misuse detection systems is the pattern recognition and signature depiction of the pertinent attack, which should encompass all possible variations but should not match non-intrusive activities [1].

Nowadays Intrusion Detection System is developed based on data mining techniques in real time environment. Data mining techniques can be dissimilated by their different model functions and representations, preference criterions, and algorithms. There are some important things that contribute for an Intrusion detection implementation using data mining: Firstly, Removing normal activity from alarm data for focusing real attacks; Second, Identifying false alarms and “awful” sensor signatures and; Third Finding abnormal activity that uncover a real attack. Classification is data mining technique which is used to predict association for data instances. Classification techniques evaluate and classify the data into known classes. Each data sample is marked with a known class label. Also, these methods are used to learn a model using the training set data sample. This model is used to separate the data samples as anomalous behaviour data or the normal behaviour data [1].

The following section represents some of the most popular classification algorithms for implementing Intrusion Detection System such as: SVM (Support Vector Machine, see [2]) and Clustering based on Bee Colony [13] are approaches in data mining to generate classifiers. Both methods have been applied in intrusion detection to separate the normal and abnormal network connecting records. From the machine learning point of view, the process of SVM is in supervised learning [3] whereas the clustering based on Bee colony algorithm is unsupervised learning [4].

One of the promising technique is support vector machine (SVM) [10], whose mathematical foundations (Khan, Awad, & Thuraisingham, 2007; Yu, Yang, Han, & Li, 2003) have provided fulfilling results. SVM separates data into multiple classes (at least two) by a hyperplane and simultaneously minimizes the empirical classification error and maximizes the geometric margin. Thus, it is also known as maximum margin classifiers.

Although SVM have shown good results in data classification, they are not favorable for large-scale dataset because the training complexity is very dependent on the amount of data in the training set. When new information is added into a system, updating the old model is necessary immediately to ensure that the system is properly protected. As retraining may take weeks, or even months, it is impractical to retrain the new model on all available data. Thus, a mechanism is needed to generate an adaptive model that can be updated by

cooperation of the old model with the new information. We can take advantage of the clustering based on Bee Colony in updating the models. Clustering in intrusion detection is used to resolve the multiple classification problems [8].

In this research, we present a framework that aims to combine the two algorithm for intrusion detection and therefore, to reach a excelling system performance. A new algorithm CSVBC, that combines the SVM and Clustering based on Bee Colony to minimize the training dataset while allowing new data points to be added to the training set dynamically.

The next sections in paper are organized as follow. In Section 2, we have discussed about related work of various existing algorithm. In section 3 and 4 we briefly introduce SVM and clustering based on Bee Colony Network. The innovation proposed is described in section 5. Section 6 explains the experimental interpretation about proposed algorithm. Lastly, conclusion of research is summarized in section 7.

2. RELATED WORK

The many intrusion detection system based on machine learning techniques are summarized-Shi-Jinn Hornc *et al.* proposed a SVM-based intrusion detection system based on a hierarchical clustering algorithm [10] to preprocess the KDD Cup 1999 dataset before SVM training. The hierarchical clustering algorithm was used to provide a high Quality, abstracted, and reduced dataset for the SVM training, instead of the originally enormous dataset. Thus, the system could greatly shorten the training time and also achieve better detection performance in the resultant SVM classifier. The purpose of using the hierarchical clustering algorithm is to provide the SVM classifier with fewer but higher quality training data that may reduce the training time and improve the performance of the classifier. The term ‘‘high quality” means that the abstracted data points in the reduced dataset must represent all data points in the original dataset, and the number of abstracted data points cannot be too small or too large.

Levent Koce *et al.* [7] explained the need to apply data mining methods to network events to classify network attack events. They summarized the results of earlier studies and explored the earlier models on the performance improvement of the Naïve Bayes model in data mining .They proposed the HNB model as a solution to the intrusion detection problem. An extended version of the Naïve Bayesian classifier is the hidden Naïve Bayes (HNB) classifier, which relaxes the conditional independence assumption imposed in the Naïve Bayesian model. The HNB model is based on the idea of creating a hidden parent for each attribute; the influences from all of the other attributes can be easily combined through conditional mutual information by estimating the parameters from the training data.

Monther Aldwairi *et al* proposed an approach to improve the accuracy and speed of IDSs by harnessing the advantages and properties of bee’s environment [11]. In addition, it reduces the computation complexity and time by applying feature selection technique while making sure the feature reduction does not affect the detection accuracy. Using artificial bee colony in intrusion detection can be benefitted from the advantages of bee colony in its environment in terms of self-organization and self-structuring. They optimized and adapted ABC to work for Anomaly Intrusion Detection.

Reda M. *et al.* proposed two data-mining techniques [6] which are used in misuse, anomaly and hybrid detection. First off, the random forests algorithm is used as a data mining classification algorithm into a misuse detection method to build intrusion patterns from a balanced training dataset, and to seperate the captured network connections to the main types of intrusions due to the built patterns. The main drawback of the misuse detection method is that it cannot detect novel intrusions that are not trained on earlier. Secondly, the k-means algorithm which is used as a data mining clustering algorithm into a unsupervised anomaly detection method to partition the captured network connections into a specified number of clusters, and then observe the anomalous clusters depending on their features. The main drawback of the anomaly detection method is the high false positive rate. They proposed that a hybrid framework can strengthen the advantages of both misuse and anomaly detection by increasing the detection rate and decreasing the false positive rate

Ant Colony Optimization (ACO) [14] has been applied in many fields to solve optimization problems, but its appliance to the intrusion detection domain is limited. Several methods were reported using ACO for intrusion detection. In [15,16], the authors evaluated the basic ant-based clustering algorithms and proposed several improvement strategies to overcome the limitations of these clustering algorithms that would not perform well on clustering large and high-dimensional network data.

Qinglei Zhang *et al.* proposed new algorithm for intrusion detection system by combining two existing machine learning methods [8] (i.e. SVM and CSOACN). Based on the new algorithm, IDS can be developed. SVM and CSOACN are two interactive phases that are taken multiple times. Support Vector Machines (SVMs) have been widely accepted as a powerful data classification method. On the other hand, the CSOACN has been shown to be efficient in data clustering.

3. SUPPORT VECTOR MACHINE

To introduce the new algorithm, it is necessary to give a brief review of SVM.

3.1 Linear SVM trained on separable data

In an SVM [8], a data point is viewed as a vector in the d-dimensional feature space. Assume that all data points belong to either class A or class B. Each training data point xi can be labelled by yi based on (1):

$$y_i = \begin{cases} -1 & X \in \text{classA}, \\ 1 & X \in \text{classB}. \end{cases}$$

Thus, the training data set can be denoted as

$$D = \{ (xi, yi) | \ i = 1, 2, 3 \ldots, N \}.$$

Data points with label 1 and −1 are referred to as positive and negative points, respectively. In the linear separable case, there are distinct hyperplanes which might separate the positive from the negative points. The algorithm generally looks for the largest margin separating hyperplane, where the ‘‘margin’’ of a separating hyperplane is defined to be the sum of the distances from the hyperplane to the closest positive and negative points.

In order to compute the margin of a separating hyperplane H, consider the hyperplanes H1 and H2 that contain the closest positive training points and the closest negative training points to H, respectively:

$$H: \ w.x - b = 0, \ x \in R^d,$$

$$H1: w.x - b = 1, \ x \in R^d,$$

$$H2: w.x - b = -1, x \in R^d,$$

where w is the normal to H and b is the distance from H to the origin. Obviously, H, H1, and H2 are equally distant. In addition,

$$w.xi - b \geq 1 \quad \text{for } yi = 1,$$

$$w.xi - b \leq -1 \quad \text{for } yi = -1.$$

The margin of the separating hyperplane H is equal to the distance between H1 andH2, which is $2/\|w\|$. Finding the maximum margin can be accomplished by minimizing $\|w\|$. It is clear that the separating hyperplane H will not change if the training points which are neither on H1 nor on H2 are removed from the training data set. Thus the points that lie on H1 or H2 are critical elements in the training data set, called support vectors.

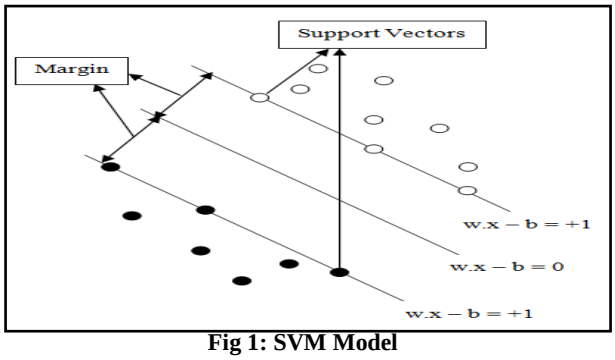

Fig 1: SVM Model

3.2 Non-Linearly Separable Cases

In the case of the data points not being linearly separable [10], a slack variable can be added to allow for some errors to be made, and the aim is to minimize the number and effect of the errors. The new optimization problem becomes

minimize $\| w \|^2/2 + C\xi_i$

subject to $C_i(w \,.\, \xi i - b) \geq 1 - \xi i$, for all $1 \leq i \leq n$.

where ξi is a slack variable that measures the degree of misclassification in the data point xi, and C is a constant. By introducing this slack variable, the classifier can minimize the impact of errors by finding the most successful decision boundary. Other approaches exist for the separation of nonlinearly separable data, such as the use of a kernel function to create a nonlinear decision boundary. A kernel function takes a data set and transforms it into a higher dimension through the use of some function (common ones include radial basis functions, Gaussian functions, and sigmoidal functions). The transformed data set may become linearly separable in the higher dimension, despite of the fact that it remains inseparable in its original space. The result is a decision boundary which appears nonlinear in the original dimension, but which is linear in the higher dimension, and which is capable of separating the two classes. While one SVM is capable of separating only two distinct classes, the combination of multiple SVMs can successfully separate more than two classes. This allows for more specific classification of attacks, such as ‘‘DoS’’ and ‘‘User to Root’’, instead of simply ‘‘attack’’ and ‘‘legitimate’’.

Training support vector machines requires the solution of a very large quadratic programming (QP) optimization. Many different algorithms have been proposed and new algorithms are still being studied. A commonly used method called Sequential Minimal Optimization (SMO) for solving the QP problem is used in our research. The basic idea of SMO is to break down the problem into two-dimensional sub problems that may be solved analytically.

4. CLUSTERING BASED ON BEE COLONY

In order to cluster data using bee colony algorithm [12].This model offers us a chance to apply bee colony optimization algorithm for the optimal clustering of a collection of data.

4.1 Representation of Solutions

In order to apply BCO to solve clustering problem, This model have used floating point arrays to encode cluster centers. The assignment matrix has the properties that each data must be assigned exactly to one cluster. An assignment that represents *K* nonempty clusters is a legal assignment. In this model, each food source discovered by each bee is a candidate solution and corresponds to a set of *K* centroids. Let us denote it by a finite set of pre-selected stages, where *K* is the number of stages. By *B*, we represent the number of bees to participate in the search process and by *I* the total number of iterations.

At each forward pass, bees are supposed to visit a single stage. All bees are located in the hive at the beginning of the search process. Every artificial bee allocates some of the data to the corresponding cluster with special probabilities in each stage, and in this way constructs a solution of the problem incrementally. Bees are combining solution components to the current partial solution until they visit all of the *K* stages. The search process is composed of iterations. The first iteration is completed when bees create feasible solutions. The best obesrved solution during the first iteration is saved, and then the second iteration begins. In every iteration of algorithm, for each cluster (stage), all the bees leave the hive to allocate some of the data to that cluster with special probabilities and come back to the hive to see the work of other bees until that time they decide whether to continue its way or select one of the other bees’ solution and continue on that way.

4.2 Evaluation of solutions

A key characteristic of most clustering algorithms is that they use a global criterion function whose optimization drives the entire clustering process. For those clustering algorithms, the clustering problem can be stated as computing a clustering solution such that the value of a particular objective function is optimized. Our objective function is to minimize intra-cluster similarity while maximizing the inter cluster similarity.

Fitness value of every solution is measured by equation

$$f = \sum_{j=1}^{k}\left(\sum_{i=1}^{n_j} D(d_{ij}, C_j)\right) \qquad (1)$$

A food source represents a possible solution to the problem. The quantity of existing sources of pollen, nectar in the areas that is explored by the bees corresponds to the quality of the solution represented by that food source. Bees search for food sources in a way that minimize the ration f where f is the proportional to the nectar amount of food sources identified by bees. In this problem, the goal is to find the minimum of the objective function.

5. PROPOSED ALGORITHM

We now introduce the new machine learning algorithm, named Combining Support Vector with Bee Colony, that is based on a combination of the two algorithms discussed above (i.e., SVM and Clustering based on Bee Colony). In this algorithm, SVM and Clustering based on Bee Colony are two interactive phases which are taken multiple times. SVM is

used to find support vectors and to generate hyperplane that separates normal and abnormal data while a Clustering based on Bee Colony is used to find data added to SVM training set and to finally generate models for normal data as well as for each class of abnormal data.

5.1 Intrusion Detection based on combination of SVM and Clustering based on Bee Colony

We now describe the combination of the two algorithms (SVM and Clustering based on Bee Colony). As mentioned before, the input data in intrusion detection are the network connecting records, which can be viewed as data points in SVM and objects in Clustering based on Bee Colony. The training process of the new method is given in Algorithm1. Algorithm1 shows the process of the new CSVBC (Combining Support Vectors with Bee Colony) algorithm. SVM and Clustering based on Bee Colony are two interactive phases that are taken multiple times. The distinctive phases are described as the following.

SVM training phase (line 5): The training data for each SVM training are objects (viewed as data points in this phase) selected by the last bee clustering phase or initially selected, randomly. In order to let the first generated hyperplane correctly separate normal data and different types of abnormal data, initially one data point is randomly selected from each type of data. After this phase, the support vectors among the selected data points will be found and marked for the next ant clustering phase.

Algorithm1: Training in CSVBC

Input: A training data set.

Input: N – number of training iterations.

Input: RR – accuracy threshold.

Output: SVM Classifier

1 begin

2 Let r be the accuracy, initially 0;

3 while r < RR do

4 for k = 1, • • •, N do

5 SVM training phase;

6 Find support vectors among the selected data points;

7 Apply Bee clustering around the support vectors;

8 Add the points of the clusters to the training set;

9 end

10 Constructing classifiers;

11 Do testing to update r;

12 end

13 end

Bee clustering phase (line 7): The clustering process is divided into several clustering periods by clustering around certain objects each time. Those certain objects for each clustering period are support vectors (viewed as objects in this phase) marked by the last SVM training phase. After this phase, all objects within the neighborhoods of the support vectors will be selected for the next SVM training phase.

Constructing classifiers (line 10): A binary classifier (i.e. SVM classifier) for intrusion detection can be constructed by the result generated in SVM training phase. Namely, all data can be separated into two cases. Although Clustering based on Bee Colony focuses on the boundary clustering, the other data points meanwhile are also clustered and the classifier for each class can be generated.

Classifier modification (line 4-10): The classifiers can be modified gradually by repeating the three steps: SVM training phase, bee clustering phase and constructing the classifier. The repeating loops will terminate when both of the two classifiers (i.e. SVM classifier and Clustering based on Bee Colony classifier) obtain a satisfying accuracy upon the entire training data. Furthermore, as it is time consuming to construct the classifiers and test the entire training data, it is more reasonable to repeat the SVM training phase and the bee clustering phase for a certain number of loops (denoted as *N* in algorithm1) before constructing and modifying the classifiers. The value of *N* is determined by experiments. The best value of *N* is the smallest number of loops that are needed to construct and test the classifiers one time before obtaining the satisfying accuracy.

Figure 2 shows the process of the new CSVBC (Combining Support Vectors with Bee Colony) algorithm. SVM and Clustering based on Bee Colony are two interactive phases that are taken multiple times.

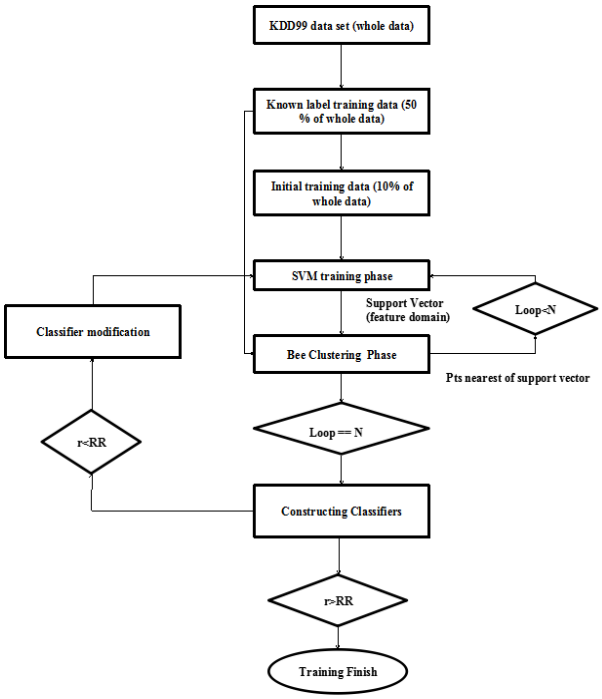

Fig 2: Flow graph of the CSVBC Algorithm

6. EXPERIMENTAL RESULTS

We selected the KDD99 data set [51] for training and testing of the newly developed IDS. As a version of the 1998 DARPA data set [17], KDD99 was first used in The Third International Knowledge Discovery and Data Mining Tools Competition, and is now considered as a standard benchmark for evaluation of data mining based IDSs. In KDD99, the data records of attacks fall into four main categories [17]: denial of service (dos), remote to local (r2l), user to root (u2r), and probing. In order to distinguish normal connections from attacks, each network connection data item in KDD99 is described by 41 higher-level features defined by Stolfo et al.

[18]. More details on feature name, description, and type can be found. Furthermore, each data item of KDD99 is labeled as either normal or an attack with one specific attack type. There are total 23 types of attacks and all of them belong to the four main attack categories (DoS, R2L, U2R, Probe).

One advantage of the new IDS is the small volume of training data. This was achieved by the selection of the SVM (see Section 3). Therefore, instead of using the whole KDD99, only a small part of it was used in the training experiments. It is known that SVM achieves the best detecting rate when the amounts of training data from the two classes to be distinguished are balanced. Let D denote the training data set. The distribution of each class in D is shown in Table 1. To compare the system performance under the two different modes (CSVAC and CSVBC), the classifiers generated by them should be tested separately by the same testing data set. As the common comparison results do not depend on the amount of testing data and the distribution of each data class, we choose the amount of testing data to be about 10 times the amount of training data.

Table 1 Class distributions in network connection records

Class	Number of network connection records			
	KDD99 data set	**Known label training data (50 % of whole data)**	**Initial training data D (10% of known label training data)**	**Testing set T**
Dos	377	189	19	377
Normal	185	93	10	185
Probe	185	93	10	185
Remote to login	234	117	12	234
User to root	19	10	1	19
Total	1000	502	62	1000

Let T be the testing data used for the comparison of the two algorithms. The class distributions of T are also shown in Table1. In order to further evaluate the effectiveness of the new combination algorithm of CSVBC we have compared the performance of our algorithm with the KDD99 data set.

We evaluate the performance of CSVBC using the criteria of Accuracy, Precision, Recall, F-measure. False positive is an event signalling an alarm (intrusion) when no attack has taken place while false negative representing the case of a failure to detect an actual attack. The experiment has repeated five times with different data records selected from the KDD99 data for training. A quantitative evaluation of the proposed technique is performed using the following criteria. These criteria are calculated from the confusion matrix in table 2, and defined as follows. This confusion matrix has created for four main attack categories (DOS, R2L, U2R, Probe) and normal data.

Table 2: Confusion matrix for U2R

Predicted Class (Result class)			
		Positive	**Negative**
Actual Class	Positive	TP (True positive) =1	FN (False Negative) =18
	Negative	FP (False positive) = 0	TN (True Negative) =981

1) Accuracy = $\frac{TP+TN}{TP+TN+FP+FN}$

2) Precision= $\frac{TP}{TP+FP}$

3) Recall= $\frac{TP}{Tp+FN}$

4) F-measure= $\frac{2*P*R}{P+R}$

We have reported about the performance of the proposed Intrusion Detection based on SVM with Clustering based on Bee Colony method. Four quantities for measuring the performance of the proposed method are considered with respect to accuracy, precision, recall and F-measure.

In below tables 2, 3 the performance of both the approaches is listed based on four parameters. In table 2 and 3 the KDD99 data item (DOS, Normal, Probe, Remote to login and User to root) are taken and the proposed approach is applied on test dataset and seen that the proposed method is high accuracy than the previous method.

Table 2 SVM with ant colony method results when applied on test dataset (T)

Class	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
DOS	76.70	63.14	91.78	74.81
Normal	88.60	79.83	51.35	62.50
Probe	92.40	94.31	62.70	75.32
R2l	90.90	84.21	75.21	79.46
U2R	98.20	100.00	5.26	100.0
Average	85.54	77.62	73.40	0.7248

In table 2 shows the result when SVM with ant colony method is applied onto test dataset (T) and it is seen that in SVM method with ant clustering, average accuracy is 85.54%.

Table 3 SVM with Bee Colony method results when applied on test dataset (T)

Class	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
DOS	83.50	72.84	89.66	80.38
Normal	88.20	71.34	60.54	65.50
Probe	94.10	94.37	72.43	81.96

R2l	91.40	81.36	82.05	81.70
U2R	98.20	100.00	5.26	100.0
Average	88.46	79.05	77.70	76.89

In proposed method same test dataset (T) applied and it is seen that, in proposed method technique, average accuracy is 88.46% as it can be seen in Table 3 and precision is 79.05% , recall is 77.70% and f-measure 76.89%. Fig 3 shows the comparison graph for Accuracy.

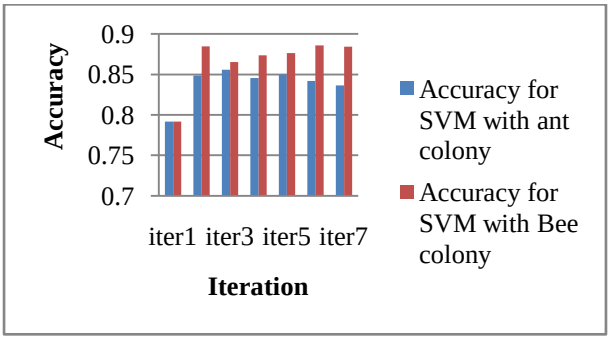

Fig 3: Comparison for Accuracy

7. CONCLUSION

In this research we introduce the new machine learning algorithm, named Combining Support Vector machine with Bee Colony that is based on a combination of the two algorithms discussed above (i.e. SVM and Clustering based on Bee colony). SVM is suitable for the time intense case that only one binary classifier is required by training upon a few amount of labeled data. Namely, it only needs to differentiate the normal data from abnormal one. On the other hand, the Clustering based on Bee colony is suitable for the preciseness intensive case and can solve multiclass problems upon both label and unlabeled data. In this algorithm, SVM and Clustering based on Bee colony are two interactive phases which are taken multiple times. The experiment result indicates that the performance of the CSVBC is better than CSVAC in terms of accuracy. The accuracy of CSVBC is 88.46%.

As future work, we are considering integrating the privacy preserving with the proposed framework in order to improve the effectiveness and the flexibility of IDS system. We also plan to enhance the CSVBC algorithm to generate more SVM classifiers to handle multiclass cases and find ways to convert a nonlinear classification problem to a linear one. The performance of IDS system can improved in future using many machine learning algorithm.

8. REFERENCES

[1] Qinglei Zhang, Wenying Feng. 2009. Network Intrusion Detection by Support Vectors and Ant Colony. Proceedings of the 2009 International Workshop on Information Security and Application (IWISA 2009), pp 639-642.

[2] J. C. Burges and Christopher. 1998. A tutorial on support vector machines for pattern recognition. DataMining and Knowledge Discovery 2, PP. 121-167.

[3] S. Kotsiantis. 2007. Supervised machine learning: A Review of classification techniques. Informatics Journal 31, PP.249-268.

[4] R. O. Duda, P. E. Hart and D. G. Stock. 2001. Unsupervised Learning and Clustering (2nd edition). wiley, New York, ISBN 0-471-05669-3, P.571.

[5] Ashis Pradhan,. 2012. SUPPORT VECTOR MACHINE-A Survey. International Journal of Emerging echnology and Advanced Engineering. vol. 2, Issue 8, pp 82-85.

[6] Reda M. Elbasiony, Elsayed A. Sallam, Tarek E. Eltobely, Mahmoud M. Fahmy. 2013. A hybrid network intrusion detection framework based on random forests and weighted k-means. Ain Shams Engineering Journal.

[7] Levent Koc, Thomas A. Mazzuchi, Shahram Sarkani. 2012. A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier. Expert Systems with Applications 39. pp 13492–13500.

[8] Wenying Feng, Qinglei Zhang, Gongzhu Hu, Jimmy Xiangji Huang. 2014. Mining network data for intrusion detection through combining SVMs with ant colony networks, Future Generation Computer Systems.

[9] Jaskiran Kaur, Inderpal Singh. (2013, June). A Survey on Ant Colony Optimization. International Journel Of Compute r Science & Engineering Technology (Ijcset). Issn: 2229-3345 Vol. 4 No.

[10] S. J. Horng, M. Y. Su, Y. H. Chen, T. W. Kao, R.J. Chen, J.L. Lai, C.D. Perkasa. 2011. A novel intrusion detection system based on hierarchical clustering and support vector machines. Expert Systems with Applications. 38, 306–313

[11] Monther Aldwairi, Yaser Khamayseh and Mohammad Al-Masri. 2012. Application of artificial bee colony for intrusion detection systems. Security And Communication Networks Security Comm. Networks.

[12] Abolfazl Toroghi Haghighat, "Data Clustering Using Bee Colony Optimization" ICCGI 2012: The Seventh International Multi-Conference on Computing in the Global Information Technology.

[13] S.X. Wu, W. Banzhaf. 2010. The use of computational intelligence in intrusion detection systems: a review. Applied Soft Computing 10 (2010) 1–35.

[14] C.-H. Tsang, S. Kwong. 2006. Ant colony clustering and feature extraction for anomaly intrusion detection. In: warm Intelligence in Data Mining, in: Studies in Computational Intelligence. vol. 34, Springer, , pp. 101–123.

[15] S. Janakiraman, V. Vasudevan. 2011. ACO based distributed intrusion detection system. Journal of Digital Content Technology and its Applications 3 (1).

[16] Lincoln Laboratory, MIT, Intrusion detection attacksdatabase,2009.http://www.ll.mit.edu/mission/communications/ist/corpora/ideval/docs/attackDB.html.

[17] S.J. Stolfo, W. Fan, W. Lee, A. Prodromidis, P.K. Chan, 2000 Cost-based modeling and evaluation for data mining with application to fraud and intrusion detection: results from the jam project 2, , pp. 1130–1144.