ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 12, December 2015

Survey on Network Intrusion Detection Using Ant Colony Networks and SVM Classification

Sandeep Ranode, Prof.Ramesh G. Patole

Master of Engineering Student, Dept. of Computer Engineering, G.H.Raisoni College of Engineering, Wagholi, Pune, Maharastra, India

Assistant Professor, Dept. of Computer Engineering, G.H.Raisoni College Of Engineering, Wagholi, Pune, Maharastra, India

ABSTRACT: Intrusion detection systems classify computer behavior into two main categories: normal and abnormal activities. In order to achieve the categorization, Intrusion detection.In this paper, we introduce a new machine-learning-based data classification algorithm that is applied to network intrusion detection. The basic task is to classify network activities (in the network log as connection records) as normal or abnormal while minimizing misclassification. Although different classification models have been developed for network intrusion detection, each of them has its strengths and weaknesses, including the most commonly applied Support Vector Machine (SVM) method and the Clustering based on Self-Organized Ant Colony Network (CSOACN). Our new approach combines the SVM method with CSOACNs to take the advantages of both while avoiding their weaknesses. Our algorithm is implemented and evaluated using a standard benchmark KDD99 data set. Experiments show that CSVAC (Combining Support Vectors with Ant Colony) outperforms SVM alone or CSOACN alone in terms of both classification rate and run-time efficiency.. Proposed approach we can be implement with both well-known dataset i.e. KDD cup 1999 as well as on individual real time network dataset. IDS will be evaluate with the detection rate, detection speed, false alarm rate and attack types.

I.INTRODUCTION

In today's information system management, large-scale dataclustering and classification have become increasingly important and a challenging area. Although various tools and methods have been proposed, few are sufficient and efficient enough for real applications due to the exponential growing-in-size and high dimensional data inputs. As a particular application area, Intrusion Detection Systems(IDSs) are designed to defend computer systems from various cyber attacks and computer viruses. IDSs build effective classification models or patterns to distinguish normal behaviors from abnormal behaviors that are represented by network data.

Intrusion detection systems classify computer behavior into two main categories: normal and abnormal activities. In order to achieve the categorization, Intrusion detection.In this paper, we introduce a new machine-learning-based data classification algorithm that is applied to network intrusion detection.

II.RELATED WORK

Issues related to intrusion detection can be categorized into twobroad areas: (1) network security and intrusion detection models, and (2) intrusion detection methods and algorithms based on artificial intelligence (mostly machine learning) techniques. In this section we shall briefly review some related work in the second area, and leave area (1) to the next section, when we discuss the background of IDSs. Intrusion detection as a classification problem has been studied for decades using machine learning techniques, including traditional classification methods (single classifier) such as K-NearestNeighbor (K-NN), Support Vector Machines (SVMs), Decision Trees (DTs), Bayesian, Self-Organized Maps (SOMs), Artificial Neural Networks(ANNs), Generic Algorithms (GAs), and Fuzzy Logic, as wellas hybrid classifiers that combine multiple machine learning techniquesto improve the performance of the classifier. A review of using these approaches was given , which also included statistics of the use of these techniques reported in 55 research articles during the period 2000–2007. The review indicates that SVM and K-NN were the most commonly used techniques while the use of a hybrid increased significantly after 2004 and became mainstream. Another more recent review provided a thorough survey of intrusion detection using computational intelligence. It presented the

DOI: 10.15680/IJIRCCE.2015. 0312128

details of the classification algorithms and swarm intelligence methods to solve problems using the decentralized agents. Most recently, IDS was introduced by integrating On Line Analytical Processing (OLAP) tools and data mining techniques. It is shown that the association of the two fields produces a good solution to deal with defects of IDSs such as low detection accuracy and high false alarm rate. As stated, as one of the swarm intelligence approaches, Ant Colony Optimization (ACO),has been applied in many fields to solve optimization problems, but its application to the intrusion detection domain is limited. Several methods were reported using ACO for intrusion detection. For example, an ant classifier was proposed that used more than one colony of ants to find solutions in multiclass classification problem. Another ant-based clustering algorithm applied to detect intrusions in a network presented inshowed that the performance was comparable to some traditional classification methods like SVM, DT, and GA 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. The work presented also used ACO for intrusion detection in a distributed network. The basic ingredient of their ACO algorithm was a heuristic for probabilistically constructing solutions. All these ACO-based intrusion detection approaches are single classifiers as categorized. Hybrid intrusion detection approaches involving SVM have been studied in the past, such as the one reported that uses the Dynamically Growing Self-Organizing Tree (DGSOT) algorithm for clustering to help in finding the most qualified points to train the SVM classifier. It starts with an initial training set and expands the set gradually so that the training time for the SVM classifier is significantly reduced. Another hybrid intrusion detection approach was recently reported that combines hierarchical clustering and SVM.

III.PROPOSED ALGORITHM

Support Vector Machines (SVMs) have been widely acceptedas a powerful data classification method. On the other hand, the Self-Organized Ant Colony Network (CSOACN) has been shown to be efficient in data clustering. Our work aims to develop an algorithm that combines the logic of both methods to produce a high performance IDS. One challenge of developing IDSs is to realize real-time detection in high-speed networks. In this paper we are using 2 algorithms:

1) **Support vector machines (SVM)**
2) **Combination of Support vector machines (SVM) and Clustering based on a self-organizing ant colony network (CSOACN)**

IV.PSEUDO CODE

1) SUPPORT VECTOR MACHINES (SVM):

```
Algorithm Steps:
Input: A training set with each data point labeled as positive
or negative (class labels).
Output: A classifier.
1) Begin
2) Randomly select data points from each class.
3) Generate a SVM classifier.
4)While more points to add to training set do
5) Find support vectors among the selected points;
6) Apply CSOACN clustering around the support vectors;
7) Add the points in the clusters to the training set;
8) Retrain the SVM classifier using the updated trainingset;
9)End
10)End
```

2) Combination of Support vector machines (SVM) and Clustering based on a self-organizing ant colony network (CSOACN):

```
Algorithm Steps:
Input: A training data set.
Input: N – number of training iterations.
Input: RR – detection rate threshold.
Output: SVM and CSOACN Classifiers.
1)Begin
2) Normalize the data;
3) Let r be the detection rate, initially 0;
4)While r <RR do
5) for k = 1, ・ ・ ・ , N do
6) SVM training phase;
7) Ant clustering phase;
8)End
9) Construct classifiers;
10)do testing to update r;
11)End
12)End
```

V.ACKNOWLEDGEMENT

We would like to take this opportunity to express my sincere gratitude to my Project Guide Ramesh G. Patole (Assistant Professor, Computer Engineering Department) for his encouragement, guidance, and insight throughout the research and in the preparation of this dissertation. He truly exemplifies the merit of technical excellence and academic wisdom.

VI.CONCLUSION AND FUTURE WORK

We will work with this research HIDS as well NIDS. We found lots of existing work in HIDS basically we try to increase the detection rate NIDS with different attack using proposed algorithms.

As future work, we are considering integrating the privacy preserving OLAP with the proposed framework in order to improve the effectiveness and the flexibility of IDS system. We also plan to enhance the CSVAC algorithm to generate more SVM classifiers to handle multiclass cases and find ways to convert a nonlinear classification problem to a linear one by applying the recently proposed Maximum Information Coefficient (MIC) method. For further performance analysis, comparisons with other existing algorithms such as those of will be conducted by applying the KDD99 data set but with different distributions and also using other standard benchmark data sets.

REFERENCES

[1] W. Lee, S.J. Stolfo, K.W. Mok, A data mining framework for building intrusiondetection models, in: Proceedings of IEEE Symposium on Security and Privacy, 1999, pp. 120–132.

[2] W. Lee, S.J. Stolfo, K.W. Mok, Mining audit data to build intrusion detection models, in: Proceedings of the 4thInternational Conference on Knowledge Discovery and Data Mining, AAAI Press, 1998, pp. 66–72.

[3] T. Zhang, R. Ramakrishnan, M. Livny, BIRCH: an efficient data clustering method for very large databases, in: Proceedings of SIGMOD, ACM, 1996, pp. 103–114.

[4] L. Khan, M. Awad, B. Thuraisingham, A new intrusion detection system using support vector machines and hierarchical clustering, The VLDB Journal 16 (2007) 507–521.

[5] X. Xu, Adaptive intrusion detection based on machine learning: feature extraction, classifier construction and sequential pattern prediction, Information Assurance and Security 4 (2006) 237–246.
[6] J.X. Huang, J. Miao, Ben He, High performance query expansion using adaptive co-training, Information Processing & Management 49 (2) (2013) 441–453.
[7] Y. Liu, X. Yu, J.X. Huang, A. An, Combining integrated sampling with SVM ensembles for learning from imbalanced datasets, Information Processing &Management 47 (4) (2011) 617–631.
[8] V. Vapnik, The Nature of Statistical Learning Theory, Springer, 1999.