ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 6, Issue 2, pp: (191-193), Month: April - June 2018, Available at: www.researchpublish.com

Building an Intrusion Detection System Using Filter-Based Feature Selection Algorithm

[1]Mrs. Shilpa S G, [2]Vivek YS, [3]Suhas B V, [4]Veerendra Kumar C G

[1]Asst. Professor, Dept. Computer Science and Engineering, Karnataka, India
[2, 3, 4] Dept. Computer Science and Engineering, DBIT, Bengaluru Karnataka, India

Abstract: **Redundant and irrelevant features in data have caused a long -term problem in network traffic classification. These features not only slow down the process of classification but also prevent the classifier from making accurate decisions, especially while copying the big data. In this paper we propose a mutual information based algorithm that analytically selects the optimal feature for classification. The mutual information based feature selection can handle both linear and non linear dependent data features. An Intrusion Detection Systems (IDS) named Least Square Vector Machine Based IDS(LSSVM-IDS),is build using the features selected by our proposed feature selection algorithms.**

Keywords: **Intrusion detection, feature selection, mutual information, linear correlation coefficient, least square support vector machine.**

I. INTRODUCTION

Despite increasing awareness of network security, the existing solutions remain incapable of fully protecting internet applications and computer networks against the threats. Developing effective and adaptive security approaches, therefore has become more critical than ever before. The traditional security techniques, as the first line of security defence, such as user authentication, firewall, and data encryption, are insufficient to fully cover the entire landscape of the network security while facing challenges from ever evolving intrusion skills and techniques. Hence, another line of security defence is highly recommended, such as Intrusion Detection System(IDS).Recently, an IDS alongside with antivirus software has become an important complement to the security infrastructure of most organizations. The combination of these two lines provides a more comprehensive defence against those threats and enhances network security.

The key contribution of this paper are as follows:

1) This work proposes a new filter-based feature selection method in which mutual information is introduced to evaluate the dependence between features and output classes. The most relevant features are retained and used to construct classifiers for respective classes.

2) We conduct complete experiments on two well known IDS datasets in addition to the dataset used. This is important in evaluating the performance of IDS since KDD dataset is outdated and doesn't contain most noval attack patterns.

3) Different from the detection framework proposed in that designs only for binary classifications, we design our proposed framework to consider multiclass classification problems. Recently, a forward feature selection algorithm using the mutual information method to measure the relation among feature. The optimal feature set was the used to train the LS-SVM classifier and to build the ideas. The hierarchical clustering algorithm was used to provide the classifier with fewer and higher quality training data to improve the classifying performance of the classifier.

II. EXISTING SYSTEM

- A significant amount of research has been conducted to develop intelligent intrusion detection techniqueswhich help achieve better network securityBagged boosting-based on C5 decision trees and kernel Miner are the earliest attempts to build intrusion detection schemes.
- Mukkamala et. al. investigated the possibility of assembling various learning methodsincluding Aritficial Neural Networks(ANN),SVM's and Multivariate Adaptive Regression Splines(MARS) to detect intrusions.

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 6, Issue 2, pp: (191-193), Month: April - June 2018, Available at: www.researchpublish.com

III. PROPOSED SYSTEM

- We have proposed a hybrid feature selection algorithm (HFSA).HFSA contains two phases.
- The upper phase conducts the preliminary search to eliminate the redundancy and irrelevant features from the original data. This helps the wrapper method to decrease the searching range from the entire original space to the pre selected features.
- The work proposes a new filter-based feature selection method, which theoretical analysis of mutual information is introduced to evaluate the dependence between feature and output classes.
- WE conduct complete experiments on two well known IDS datasets in addition to the dataset used. This is very important in evaluating the performance of IDS since KDD dataset is outdated and does not contain most novel attack patterns in it.

IV. SYSTEM ARCHITECTURE

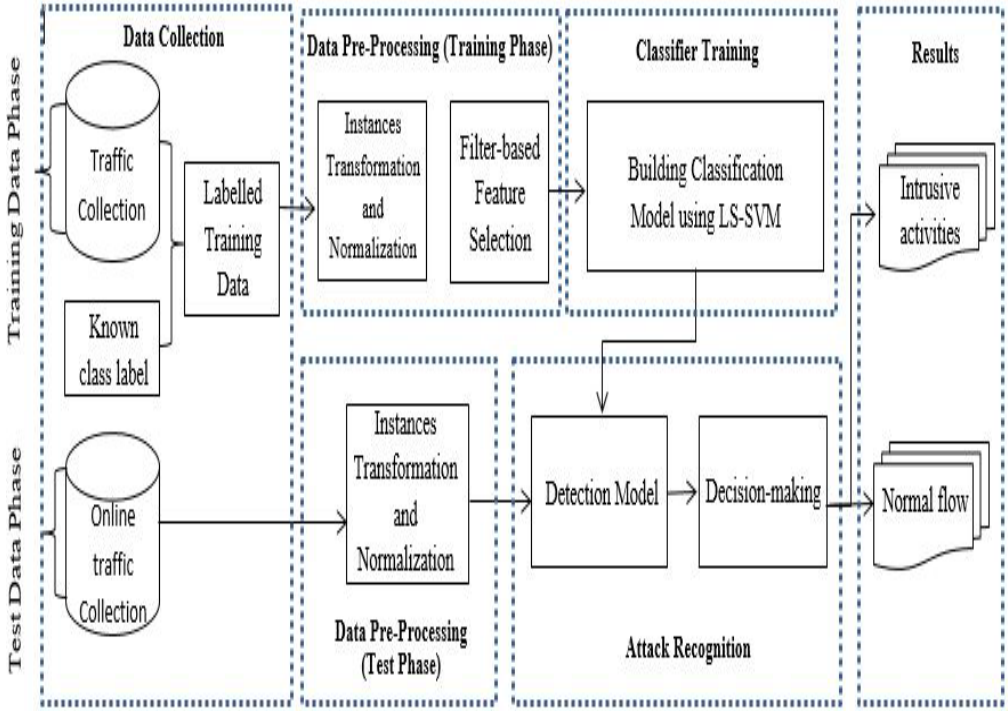

Fig 1: System architecture

V. DISADVANTAGES OF EXISTING SYSTEM

- Existing solutions remain incapable of fully protecting internet applications and computer networks against the threats from ever-advancing cyber attack techniques such as DoS attack and computer malware.
- Current network traffic data, which are often huge in size, present a major challenge to IDSs. These “big data” slow down the entire detection process and may lead to unsatisfactory classification accuracy due to the computational difficulties in handling such data.

VI. CONCLUSION

- Recent studies have shown that two main components are essential to build an IDS. They are a robust classification method and an efficient feature selection algorithm.
- In this paper, a supervised filter-based feature selection algorithm has been proposed, namely Flexible Mutual Information Feature Selection (FMIFS). FMIFS is an improvement over MIFS and MMIFS. FMIFS suggests a modification to Battiti’s algorithm to reduce the redundancy among features.

- FMIFS eliminates the redundancy parameter b required in MIFS and MMIFS. This is desirable in practice since there is no specific procedure or guideline to select the best value for this parameter.
- FMIFS is then combined with the LSSVM method to build an IDS. LSSVM is a least square version of SVM that works with equality constraints instead of inequality constraints in the formulation designed to solve a set of linear equations for classification problems rather than a quadratic programming problem.

Although the proposed feature selection algorithm FMIFS has shown encouraging performance, it could be further enhanced by optimising the search strategy. In addition, the impact of the unbalanced sample distribution on an IDS needs to be given a careful consideration in our future studies

REFERENCES

[1] Abhishek Mitra, Kenil Dinesh Patel, Prof.Indumathy, Dinesh. K. "Android Based Smart Parking Reservation" , International Journal of Innovative Research in Computer and Communication Engineering ,Vol. 4, Issue 9, September 2016.

[2] Shinde Smita N., Shinde Komal V., NagpureRashmilaD. ,Tupkar Avanti S., Prof. Ankoshe M. S. " An Android Application for Parking Management and Dissemination System ",*International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 3, March 2015.*

[3] SupriyaGatalwar, Radhika Agnihotri, Nitesh Gujarathi, AtmeshBehere."ParkSmart: Android Application for Parking System", IJCSN International Journal of Computer Science and Network, Volume 5, Issue 1,Febraury 2016.

[4] Lalitha Iyer , Manali Tare, Renu Yadav, HetalAmrutia, Android Application for Vehicle Parking System: "Park Me",IJIACS International Journal of Innovations & Advancement in Computer Science, Volume 3, Issue 3,May 2014.

[5] M. Ataur Rehman, M.M. Rashid, A. Musa, A. Farhana and N. Farhana, "Automatic parking management and parking fee collection based On Number Plate Recognition", International Journal of Machine Learning and Computing, vol. 2, no. 2, pp. 93-98, 2012.

[6] Patrick Sebastian, Hamada R.H. Al-Absi, Justin Dinesh Daniel Devraj and Yap VooiVoon, "Vision based automated parking System", 10 International conference on Information Science, Signal Processing and their Applications (ISSPA 2010), no. 1, pp. 757-760, 2010.