Journal of Computer Engineering and Information Technology

Opinion Article A SCITECHNOL JOURNAL

A Survey of Supervised Machine Learning's Algorithms for Intrusion

Michael Mangan*

Department of Computer Engineering, University of Bangalore, Bangalore, Karnataka, India.

***Corresponding Author:** Michael Mangan, Department of Computer Engineering, University of Bangalore, Karnataka, India. E-mail: manganchael@gmail.com

Received date: 12 May, 2022, Manuscript No. JCEIT-22-61570;

Editor assigned date: 13 May, 2022; PreQC No. JCEIT-22-61570(PQ);

Reviewed date: 30 May, 2022, QC No. JCEIT-22-61570;

Revised date: 10 June, 2022, Manuscript No. JCEIT-22-61570(R);

Published date: 29 June, 2022, DOI: 10.4172/jceit.1000234.

Description

As network attacks have increased in number and severity over the past few years, intrusion detection system (IDS) is increasingly becoming a critical component to secure the network. Due to large volumes of security audit data as well as complex and dynamic properties of intrusion behaviors, optimizing performance of IDS becomes an important open problem that is receiving more and more attention from the research community. This system compares the performance of Intrusion Detection System (IDS) Classifiers using various feature reduction techniques. To enhance the learning capabilities and reduce the computational intensity of competitive learning comparing the performance of the algorithms is performed respectively, different dimension reduction techniques have been proposed. These include classifying and clustering Algorithms Naïve Bayes, Simple k mean, Decision tree and Linear Discriminate Analysis, and Independent Component Analysis. Many Intrusion Detection Systems are based on neural networks. However, they are computationally very demanding. This system provides a review on current trends in intrusion detection together with a study on technologies implemented by some researchers in this research area. We try to build as system which create clusters from its input data by labelling clusters as normal or anomalous data instances and finally used these cluster to classify unseen network data instances as either normal or anomalous.

Both training and testing was done using different subset of data which is very popular and widely used intrusion attack dataset. Key words are Supervised Machine Learning, Unsupervised Machine Learning, Network Intrusion Detection and Network Security. Intrusions can be defined as actions that attempt to bypass the security mechanisms of computer systems. Intrusions may take many forms: attackers accessing a system through the Internet or insider attackers authorized users attempting to gain and misuse non-authorized privileges. So, we say that intrusions are any set of actions that threaten the integrity, availability or confidentiality of a network resource. Intrusion detection is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of intrusions. Intrusion detection systems (IDS) raise the alarm when possible intrusions occur. He concludes that the evaluation of human integration is necessary to reduce classification errors. His experimental results showed that achieves similar or even better results, compared to just an RBF. Network Security maintenance is one of the major safety concerns for neutralizing any unwanted activities. It is not only for protecting data and network privacy issues but also for avoiding any hazardous situations. From January through June 2010 Microsoft security intelligence report shows that the infection trends are still increasing on average around the world at a higher rate. For decades, Network security is one of the major issues and different types of developed systems are being implemented.

Support Vector Machine

Network intrusion is an unauthorized activity over the network that steals any important and classified data. Also sometimes it's the reason of unavailability of network services. The unexpected anomaly occurs frequently and a great loss to internet cyber world in terms of data security, the safety of potential information's etc. There are several types of method proposed for network intrusion detection. The anomaly network intrusion detection is a major part of network security. A lot of research into artificial neural networks (ANNs) has been undertaken. In artificial neural networks and support vector machine (SVM) algorithms were applied to intrusion Wireless sensor network (WSN) consists of sensor nodes. WSN suffers from several attacks, intrusion and security vulnerabilities. Intrusion detection system (IDS) is one of the essential security mechanisms against attacks in WSN. In this system present a comparative evaluation of the most performing detection techniques in IDS for WSNs, the analyzes and comparisons of the approaches are represented technically. Attacks in WSN also are presented and classified into several criteria. To implement and measure the performance of detection techniques we prepare our dataset, based on five steps, after normalizing our dataset, we determined normal class and 4 types of attacks and used the most relevant attributes for the classification process.

A lot of research work has been done in the field. RBSs (Rule Based Systems) are well suited for event correlation to perform misuse detection. However, other techniques are better suited for anomaly detection, such as statistical methods and clustering. The ability to facilitate anomaly detection is one of the benefits that have motivated much research on machine learning for intrusion detection. In the last decade, an increasing amount of research on machine learning for misuse detection can also be observed in this review. The application of techniques such as Artificial Neural Networks (ANNs) to misuse detection offers some desirable flexibility in the detection process compared with conventional variations of learned attacks can be detected. The in- flexibility of RBSs, due to operating with "crisp" rules, has been considered one of their main drawbacks .However, this observation is no longer entirely accurate, since researchers have proposed several applications. In any database system query optimization is an important component. Optimizer designing that takes less time for searching and provides the most optimal query execution plan has been a challenge for research community in the last decade. The cost of executing a plan depends heavily on selectivity which keeps changing frequently and thus static compilation provides an inconsistent performance.

Database Management

Adaptive method has depends on the external method and it has the several methods and optimization is an excellent method of generating optimal plans quite consistently. This proposes new hybrid

All articles published in Journal of Computer Engineering and Information Technology are the property of SciTechnol and is protected by copyright laws. Copyright © 2022, SciTechnol, All Rights Reserved.

architecture for query optimization which combines features of adaptive query processing and also reduces the search space for re optimization using reduced plan diagrams and cost diagrams. This hybrid architecture is bound to give more efficient performance as compared to any other optimization technique along with increased robustness and a substantial increase in consistency the selection of the most optimal execution plan. Optimization Query is a non-trivial task for every commercial database management systems. It is that step in query processing that controls how much time will be consumed for executing a query. Since SQL query is declarative in nature, no information regarding execution sequence is provided by the user. Thus finding out the optimal sequence of execution becomes the overhead of database query optimizers. Query processing is a multistep process. A simple query written in a declarative language such as SQL is first converted to an equivalent relational algebra expression and is then converted into a query tree. The task of a query optimizer is to analyze all the query execution plans and select the optimal plan for executing the query. The selection for best plan is done by applying some rules heuristics optimization and by using some cost functions.

The number of query execution plans for a given plan or the plan cardinality may be huge which makes it impossible to analyze each and every plan for optimality. If much time is spent in just searching the best plan, its execution won't prove to be much beneficial. Thus a trade-off between searching time and execution time is necessary. Because of this most query optimizers put efforts to search for the near optimal instead of searching for the most optimal plan. Another problem faced by optimizers is the selectivity of base relations. The choice for best plan is made on the basis of some complex cost functions whose major parameters are the selectivity of base relation. Since the selectivity keeps on changing frequently the cost calculation becomes wrong and a sub optimal plan may get selected. Thus a dynamic calculation of selectivity is required for getting correct value of cost functions. Static compilation of selectivity is highly prone to errors.