Chapter 12
Entropy-Based Feature Selection for Network Intrusion Detection Systems

Sellappan Devaraju
https://orcid.org/0000-0003-3116-4772
VIT Bhopal University, Bhopal, India

Sundaram Jawahar
https://orcid.org/0000-0002-8101-8725
Christ (Deemed), Ghaziabad Campus, India

Srinivasan Ramakrishnan
https://orcid.org/0000-0002-8224-4812
Dr. Mahalingam College of Engineering and Technology, Pollachi, India

Dheresh Soni
VIT Bhopal University, Bhopal, India

Alagappan Somasundaram
Sri Krishna Arts and Science College, Coimbatore, India

ABSTRACT

A network intrusion detection system (NIDS) has a significant role in an industry or organization to protect their data. NIDS should be more reliable to manage huge traffic over the networks to detect the emerging attacks. In this chapter, novel entropy-based feature selection is proposed to select the important features of intrusion detection system. Feature selection reduces the computational time and improves detection rates. In entropy, within-class entropies and between-class entropies are computed for the various classes of intrusion in the KDD dataset. Based on computed entropy values, features are ranked and selected. Radial basis neural network (RBNN) is employed as a classifier. Performances of the proposed entropy-based feature selection algorithm are evaluated using the 10% dataset for training and two other datasets for testing. The proposed system shows significant improvement in the detection rate, reduces the false positive rate (FPR), and also reduces the computational time.

DOI: 10.4018/978-1-6684-3991-3.ch012

INTRODUCTION

The Network Intrusion Detection (NIDS) System is a dependable and secure system that monitors for network vulnerabilities. The flaw will take advantage of a flaw in information assurance. On a daily basis, it is critical to lessen vulnerability from numerous enterprises. The internet is widely utilised for a variety of purposes, including commerce, education, games, entertainment, and other related activities. As a result, any organization's Network Intrusion Detection System (NIDS) is critical in protecting its data from misbehaviors. Despite the fact that every firm uses firewalls and other security measures to protect data, many intruders remain undetected. As a result, information must be better protected. Signature-based and anomaly-based IDS are the two most common types of NIDS. (i)A signature-based NIDS detects an intrusion by comparing it to previously detected intrusions. In the log files, there are signatures. (ii) The anomaly-based NIDS monitors system action and categorises it as either attack or normal (Gupta, 2010). Network-based IDS and host-based IDS are two types of IDS. When the system can converse with each other via the networks, and the network-based IDS identifies misbehaviour. (ii) If there is any misbehaviour, the host-based IDS monitors and analyses the single computer system (Devaraju, 2013; Nie, 2009).

The intrusion is detected using a signature-based or misuse-based intrusion detection system that evaluates previous signatures in log files (Ashara, 2012; Devaraju, 2019; Mansour, 2010; Suseela, 2005). Signature-based assaults rely on the knowledge gathered from previous strikes. Attack signatures, which are sets of rules that uniquely identify attacks, represent this information. Because they have superior accuracy and lower false positive rates, knowledge-based techniques are relatively straightforward for the administrator to sustain the attacks. When users detect an intrusion and compare it to the signatures log files, signature-based assaults are portrayed as known attacks. The log file contains a list of known assaults that have been detected on a computer system or network. Furthermore, the signature-based attack lacks the potential to to detect all types of attacks, particularly new attacks and those involving privilege misuse (Devaraju, 2019; Gang, 2010; Nor, 2008).

Unknown attacks are intrusion detection based on anomalies; these attacks are detected by the network and distinguished from conventional attacks. They can detect attempts to exploit novel and unexpected attacks, which gives them an advantage over signature-based attacks. However, anomaly-based techniques have their own set of drawbacks, including a high false positive rate due to a lack of training data and anomalous behaviour. Signature-based techniques are ideally suited for intrusion detection for these reasons (Devaraju, 2019; Shih-Wei, 2012; Mei, 2011).

Network-based, host-based, and application protocol-based intrusion detection systems are the three types (Jawahar, 2020; Arman, 2009; Yousef, 2014; Pablo, 2012). Misuse or anomaly-based attacks are employed in network-based assaults. The interconnectedness of computer systems is used to detect network-based assaults. For intrusion detection, audit patterns are formed over networks, however extracting the essential information from the audit patterns of the networks is difficult. When two computers are linked together, the attack is sent from one to the other via networking equipment. The assault detection rate and false positive rate are also influenced by large volumes of audit patterns (Jawahar, 2020; Gaik-Yee, 2013; Selvakani, 2011).

Host-based breaches are easy to detect and prevent because they originate from a specific computer system. When some external devices are connected to the computer system, incursions occur. Floppy discs, compact discs, pen drives, and other similar devices are used. The audit patterns generated by the computer system, which comprises system log files and error log files, have been evaluated by the