Chapter 32
Big Data Analytics With Machine Learning and Deep Learning Methods for Detection of Anomalies in Network Traffic

Valliammal Narayan
Avinashilingam Institute for Home Science and Higher Education for Women, India

Shanmugapriya D.
Avinashilingam Institute for Home Science and Higher Education for Women, India

ABSTRACT

Information is vital for any organization to communicate through any network. The growth of internet utilization and the web users increased the cyber threats. Cyber-attacks in the network change the traffic flow of each system. Anomaly detection techniques have been developed for different types of cyber-attack or anomaly strategies. Conventional ADS protect information transferred through the network or cyber attackers. The stable prevention of anomalies by machine and deep-learning algorithms are applied for cyber-security. Big data solutions handle voluminous data in a short span of time. Big data management is the organization and manipulation of huge volumes of structured data, semi-structured data and unstructured data, but it does not handle a data imbalance problem during the training process. Big data-based machine and deep-learning algorithms for anomaly detection involve the classification of decision boundary between normal traffic flow and anomaly traffic flow. The performance of anomaly detection is efficiently increased by different algorithms.

INTRODUCTION

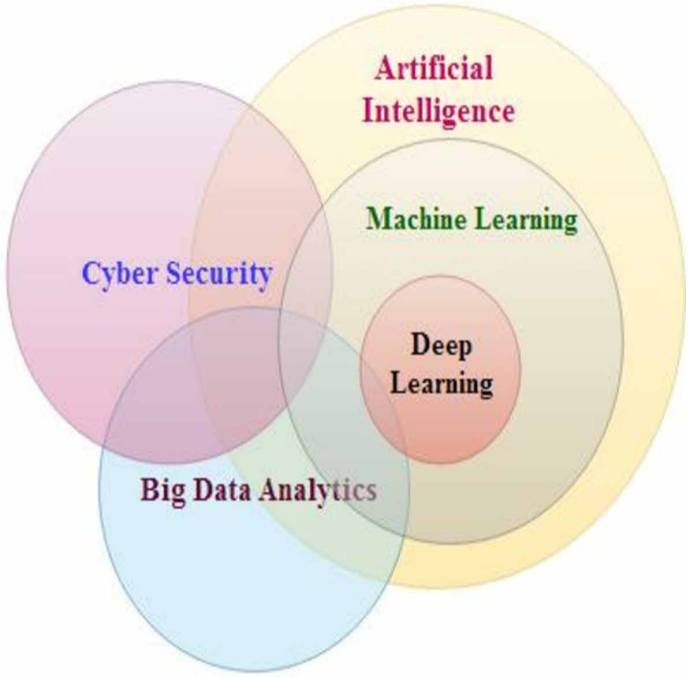
Over the past decades, the significance of cyber-security has increased and developed as a general branch of an individual life that is associated with a computer or a mobile phone. When a person submits his/her information via online, it becomes susceptible to cyber-attacks or cyber-crimes like hijacking or

DOI: 10.4018/978-1-6684-3662-2.ch032

unauthorized access, injection of virus, malware, etc. As a result, authorized access via web services is offered by cyber-security. This chapter summarizes the significance of cyber-security, how it can be developed and the considered key points during the selection of a cyber-security service provider.

The cyber world is expanding rapidly day by day and more and more people are getting connected to this world, resulting in generation of a large amount of data called Big Data. Big data is large in both quantity and quality and can be efficiently used to analyze certain patterns and behaviour anomalies which can help us prevent or be prepared for the thread or any upcoming attack. This proactive and analytical approach will help us greatly reduce the rate of Cyber Crimes and also get the knowledge out of that data which was not previously observable. Big Data analytics using machine learning techniques have a major and evolving role to play in cyber security (M.D. Anto Praveena, 2017) as in Figure 1 The cyber security problems can now impact every aspect of modern society, from hospitals, banks, and telecoms to governments and individuals.

Figure 1. Overview of the Big Data Analytics for Cyber Security



The battle against cyber security breaches is fought along the four dimensions of Prevention, Preparation, Detection, and Response. Over the last decade, the security industry seems to have largely given up on Prevention, but that is a topic for another day. It is in the dimensions of Preparation and Detection that Big Data Analytics capabilities are being used to identify anomalous patterns and to connect the dots across diverse systems and data sets. The data may be categorized into transaction and interaction data, entity data, systems operations data, reference data, and activity logs data. Big Data analytics using artificial intelligence techniques will self-learn normal patterns by observing a data flows under normal operations (Sebestyen.G, 2017).