
Intrusion detection of web sites with data mining techniques – A survey

K.P.S.D. Kumarapathirana* and C.I. Kithulgoda

Department of Computer Science, Faculty of Science, University of Ruhuna

Growth of computer intrusions highlights the incontestable importance of intrusion detection. Most common attacks on web sites are cross site scripting (XSS), SQL injection, denial of service (DoS) and session hijacking. As the usage of computer networks increased dramatically, networked computers are exposed to intrusions such as unauthorized access, bandwidth theft and DoS than ever before. Intrusion detection systems employ traditional signature based methods or data mining based methods as the basic technique of detecting intrusions. In this paper, we primarily focus on data mining based method, more specifically, anomaly detection. The usage of data mining functions such as preprocessing, association rule mining, classification and clustering in the domain of anomaly detection of web sites are discussed. Several data preparation techniques can be used to improve the performance of analysis and this process is known as preprocessing. Labeled preprocessed log data are taken to recognize classes and to generate rule associations from the frequent patterns. Importantly, previously unknown intrusions are detected by clustering. It is observed that clustering based anomaly detection techniques rely on an assumption which differentiate anomalies from normal data. As the output of our survey, comprehensive intrusion detection system for a web site was modeled. The proposed system employs several agents namely, data collector, preprocessor and detector.

Key words: Anomaly Detection, intrusion, intrusion Detection Systems, data Mining

*samantha@dcs.ruh.ac.lk