Abstract

The last few years have seen a rapid growth in internet, which comes along with an increased need of building secure systems and networks. However, it is becoming increasingly harder to keep systems secure as more and more security threats emerge due to the increasing attack surfaces. This is because of the growing increase in the number of connected network devices and services. Traditional security measures do not adapt well to the ever-changing network architecture as well as the ever-increasing number of attack types. To address this ugly scenario, network intrusion detection systems (NIDS) can be deployed in existing data networks to detect and prevent attacks without affecting their operations. Together with traditional firewalls, they form the next-generation firewall that provides more security and more awareness against attacks. This thesis explores several supervised machine learning approaches for detecting intrusion in a typical NIDS. A comparison of approaches is made to identify promising and robust candidates for further study. At the same time the main challenges associated with supervised machine learning approaches are investigated. Particular attention is paid to mitigating the problem of small sample size classes to leverage imbalanced datasets.