ECE 417 - Final Project Proposal
Network Intrusion Detection System

Evangelou Sotirios, Kalais Konstantinos, Chatziefremidis Eleftherios

March 2019

Problem and its significance to modern Cybersecurity The so-called fourth industrial revolution has changed the modern life as we knew it. Besides personal computers, the huge network we call Internet now includes, mobile phones, tablets, IOT devices and many more people have network access and knowledge that was considered arcane 20 years ago. The possibilities and opportunities are limitless; unfortunately, so too are the risks and chances of malicious intrusions. **Cybersecurity** is the practice of protecting systems, networks, and programs from digital attacks, and the subprocess of the identification of these attacks is called **Intrusion Detection**. Intrusion Detection Systems (IDS) that monitor network traffic and anomalies are called NIDS (network intrusion detection systems) and often use an approach called **anomaly-based detection** which detects deviations from normal traffic behaviours, usually using **Machine Learning**.

The data The data we decided to use are provided from this link: *UNSW dataset*. The raw network packets of the UNSW-NB 15 dataset was created by the IXIA PerfectStorm tool in the Cyber Range Lab of the Australian Centre for Cyber Security (ACCS) for generating a hybrid of real modern normal activities and synthetic contemporary attack behaviours. The dataset contains label values indicating malicious/normal traffic, thus the learning model we will utilise is supervised.

The model The model will be supervised. We conducted our own research on papers regarding Machine Learning techniques on IDS systems. Following [1], we were led to the use of either Multilayer Perceptron Neural Networks (MLP), or a Supported Vector Machines model (SVM). According to [2], the SVMs and Artificial Neural Network solutions provide very accurate results, with SVMs showing slightly better results, due to their superior properties of fast training, scalability and generalization capability. So, these are the methods we will research and choose the most suitable and accurate for our problem.

The model's evaluation Features we suppose to be irrelevant will be dropped from the dataset in the early stages. Then, we will split the dataset into two subsets, one for the training phase and one for the testing phase. The testing data will be evaluated by different metrics, including Mean Squared Error and percentile classification accuracy of the predictions. Lastly, we will define a metric for the false positive errors, as we deem them to be more important than false negatives for the nature of our problem.

Anticipated challenges The first and most crucial is the choice of the model from the suggested models mentioned in the model section. One other challenge will be to choose by trial and error the correct hyperparameters of the model and other metrics in order to train/test/evaluate our model in a suitable and correct way.

The promise The goal of the project is to create an intrusion detection system that works well with real life traffic, and can detect the attacks introduced by the dataset at a high percentage. We aspire to have a reliable and usable model that can be used both by an individual or even a company.

References

[1] Mahdi Zamani, *Machine Learning Techniques for Intrusion Detection*, University of New Mexico, 2013.

[2] Srinivas Mukkamala, Guadalupe Janoski, Andrew Sung, *Intrusion Detection: Support Vector Machines and Neural Networks* , New Mexico Institute of Mining and Technology, 2002