Comparison of Intrusion Detection Systems for Low-Powered Devices

Pratyay Amrit (140953430)

Under the guidance of

Ms. Ipsita Upasana,
Assistant Professor
Department of I & CT
MIT, Manipal

Contents

- Introduction
- Problem Definition
- Objective
- Scope
- Methodology
- Work Done
- Remaining Work

Introduction

- Security in low powered devices
- Rise of IoT and WSN
- Intrusion Detection System (IDS)

Problem Definition

- Why IDS for low powered devices?
- Signature based IDS
- Anomaly based IDS
- Problems with low powered devices

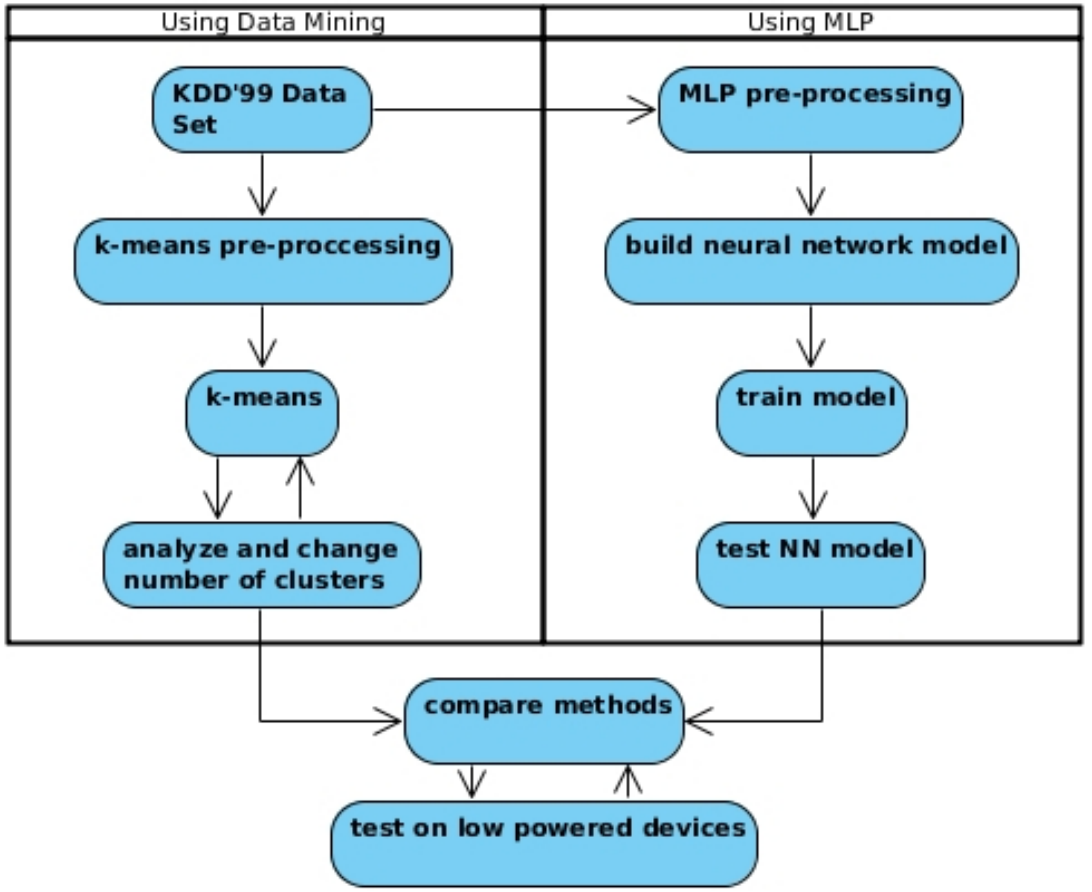
Objectives

- To find and choose a dataset that resembles a network dump as closely as possible while conforming with modern trends to make sure the IDS is well trained for novel attacks.
- To develop IDSs based on two very popular classification algorithms, K-Means and Multilayer Perceptron.
- To analyze and compare the two algorithms based on similar parameters.
- To compare the effects of decreased computational power by testing the algorithms in low powered devices.

Scope

- Cyber Security service providers
- People working in IoT/WSN
- Further research

Methodology



Methodology: Dataset

- KDD’99 vs. UNSW-NB15
- 9 types of attacks:
  - Fuzzers
  - Analysis
  - Backdoor
  - DoS
  - Exploit
  - Generic
  - Reconnaissance
  - Shell Code
  - Worm

Methodology: Dataset

- 49 features divided into 6 categories:
- Flow Features(5), Basic Features(13), Content Features(8), Time Features(10), Additional Generated Features(11), Labeled Features(2)

Methodology: Pre-processing

- Import data (python pandas)
- Fill empty values
- Transform nominal to numeric
- Add feature names to Pandas DataFrame

Methodology: K-Means

- Unsupervised clustering algorithm.
- To reduce computational load, 2 clusters. (attack/normal)

Methodology: Multilayer Perceptron

- A class of Artificial Neural Networks, with at least 3 layers and non-linear activation functions (except input nodes)
- Learning: Backpropogation.
- Optimization: Binary weights.

Methodology: Performance Parameters

- True Positive (TP): correctly classified attacks.
- True Negative (TN): correctly classified normal.
- False Positive (FP): incorrectly classified attacks.
- False Negative (FN): incorrectly classified normal.
- $$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Work Done: Overview

- Week 1: Finding research material and datasets.
- Week 2 – Week 5: Learning python for data science.
- Week 6: Started work on project.
- Week 7: Change dataset.
- Week 8: Application and performance testing.

Work Done: Week 1

- Finding research papers
- Problems with selected papers
- KDD’99 dataset

Work Done: Week 2 - 5

- Learning python for data science
- IPython basics
- Numpy
- Pandas
- Scikit-learn
- matplotlib

Work Done: Week 6

- Process KDD’99 with pandas
- Removing duplicates
- Nominal to numeric
- Apply k-means

Work Done: Week 7

- Problems with KDD’99
- UNSW-NB15
- UNSW-NB15 vs. KDD’99
- Started working on UNSW-NB15

Work Done: Week 8

- Processing UNSW-NB15 [Figure 2 (report)]
- Applied k-means
- Analysis using performance parameters [Figure 3 (report)]
- TP: 43,350
  TN: 327,461
  FP: 319,790
  FN: 9,399
- Accuracy: 52.97%

Remaining Work

- Implement MLP
- Analyze and compare on regular device
- Analyze and compare on low powered device

References

- N. Moustafa and J. Slay, “The significant features of the unsw-nb15 and the kdd99 data sets for network intrusion detection systems,” in 2015 4th International Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS), Nov 2015, pp. 25–31.
- S. Duque and M. N. bin Omar, “Using data mining algorithms for developing a model for intrusion detection system (ids),” Procedia Computer Science, vol. 61, pp. 46 – 51, 2015, san Jose, CA November 2-4, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050915029750
- P. V. S. Alpao, J. R. I. Pedrasa, and R. Atienza, “Multilayer perceptron with binary weights and activations for intrusion detection of cyber-physical systems,” in TENCON 2017 - 2017 IEEE Region 10 Conference, Nov 2017, pp. 2825–2829.
- Python for Data Science Handbook. Available: https://github.com/jakevdp/PythonDataScienceHandbook
- Scikit-learn documentation