Comparison of Intrusion Detection Systems based on Machine Learning and Data Mining Algorithms for Low-Powered Devices

Project synopsis submitted

to

MANIPAL ACADEMY OF HIGHER EDUCATION

For Partial Fulfillment of the Requirement for the

Award of the Degree

of

Bachelor of Technology

in

Computer and Communication Engineering

by

Pratyay Amrit
Reg. No. 140953430

Under the guidance of

Ms. Ipsita Upasana
Assistant Professor
Department of I & CT
Manipal Institute of Technology
Manipal, India

MANIPAL INSTITUTE OF TECHNOLOGY
MANIPAL
(A constituent unit of MAHE, Manipal)

Jan 2018

1 Introduction

Security in low-powered devices has always been a major concern as their limited computational capacity prohibits the use of sophisticated algorithms. An Intrusion Detection System(IDS) provides a first-line monitoring service for the devices to help them protect themselves against anomalous data. Several methods have been proposed to implement such a system. This work attempts to analyze and compare the performance of some of these methods under similar scenarios. the KDD'99 data-set is used to implement and compare these systems.

2 Problem Definition

Technology is growing rapidly every day, and so are incidents concerned with cyber security. Intrusion Detection Systems are responsible for monitoring the network and detecting a data packet that may pose a risk for the network. Primitive IDSs were built to detect attacks on the basis of their signature. Every time a new attack was discovered, it's signature was recorded to protect the system from similar future attacks. This, however, could not prevent the system from new attacks. With the advent of Machine Learning and Data Mining techniques, these IDSs could now be taught how an attack looks in the network to perform detection of new attacks. This however, significantly increased the false positives in detection. Another problem came up as IoT began expanding exponentially. Massive networks of small, low-powered devices have taken over the majority of space in today's world. These small devices, however, lack the ability to execute complicated machine learning algorithms. This work attempts to assess the performance of these algorithms in such low powered devices.

3 Objective

- Implement IDS using data mining and neural network
- Analyze and compare the two
- Test on low powered devices
- Scale down the algorithms to increase performance by decreasing complexity
- Re-run tests to understand the pros and cons of scaling down

4 Scope

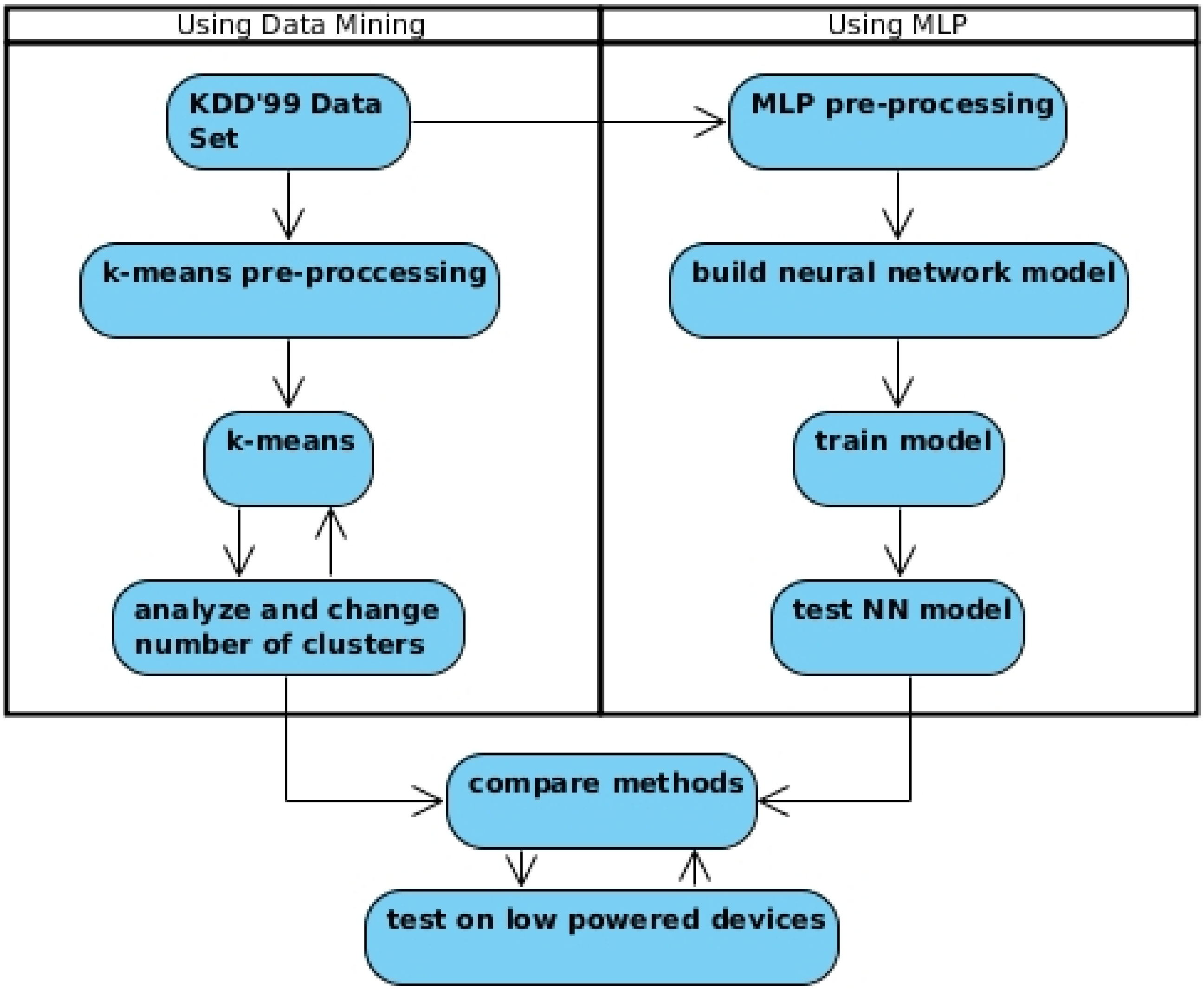
This work will benefit cyber security companies or agencies to understand how different types of IDS perform on similar environment and may use this research to influence their decision in choosing one. This research will help individuals working in IoT or cyber security domain in making the right choice for their specific case scenario. This work may also help future research in developing an all round hybrid IDS capable of performing equally well in both, high end and low powered devices.

5 Literature Survey

[1] proposes the use of k-means algorithm for IDS. The algorithm is tested with various number of clusters and the best result was shown to be when the number of clusters equalled 22 for the static database the algorithm was tested with. In a dynamic network, however, the number of clusters will greatly affect the results for different data sets and thus, identifying the number of clusters becomes an important issue.

[2] proposes the use of multi-layer perceptrons for building an IDS. The model performs better than other methods, however, is computationally expensive. The authors lowered the computational load by binarizing the weights from the neural network. This allowed the use of this algorithm in low powered devices, but also resulted in a drop in performance by 4 times.

6 Methodology



7 Work done so far

- 30 Jan 2018: Acquire and understand KDD'99 data-set
- 12 Feb 2018: Implement [1] on KDD'99
- 28 Feb 2018: Implement [2] on KDD'99
- 07 Mar 2018: Analyze and compare the two on regular devices
- 20 Mar 2018: Test on low powered devices
- 15 Apr 2018: Scale down algorithms for low powered devices
- 01 May 2018: Re-run tests and understand pros and cons of scaling down

8 Remaining work

References

[1] S. Duque and D. M. N. bin Omar. (2015) Using data mining algorithms for developing a model for intrusion detection system (ids). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915029750

[2] P. V. S. Alpano, J. R. I. Pedrasa, and R. Atienza. (2017) Multilayer perceptron with binary weights and activations for intrusion detection of cyber-physical systems. [Online]. Available: http://ieeexplore.ieee.org/document/8228342/