Cybersecurity Analytics: A New Approach to Network Segmentation

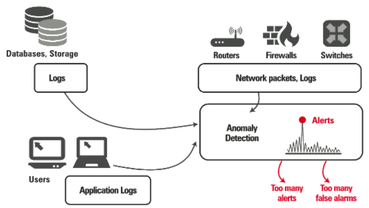
Lakshminarayana Mudradi, Geetika Mittal, Mani Prakash Rathna Kumar, Matthew A. Lanham
Purdue University, Krannert School of Management
lmudradi@purdue.edu; mittalg@purdue.edu; mrathnak@purdue.edu; lanhamm@purdue.edu

ABSTRACT

We have designed and built various network designs and ran simulations to understand the network vulnerability for a cybersecurity breach for each individual design. Using data mining and predictive modelling techniques on various network and port parameters, we attained an optimal network design that minimizes the risk of infiltration. Having a foolproof network design that minimizes the risk of infiltration becomes imperative.

INTRODUCTION

It is estimated that cybercrime will cost approximately $6 trillion per year on average through 2021 (Forbes, 2017). Also, Forbes estimates ~85% of the business assets are in digital form, making them vulnerable for cyber-attacks. With the impact of cyber security attacks, it is evident that all companies that have IT infrastructure, irrespective of their size need a robust fool-proof cyber networks. However, based on the company background, size and needs, the network infrastructure differs from one company to another. In this project, we discuss about the philosophy in choosing the optimal design with minimal risk of infiltration. The project discusses about the various factors that influences the strength of the network which can help us design and build multiple networks. Consequently, we also discuss about the data mining and predictive modelling techniques to arrive at the optimal network that minimizes the risk of infiltration.



Reference: Analytics Vidhya India

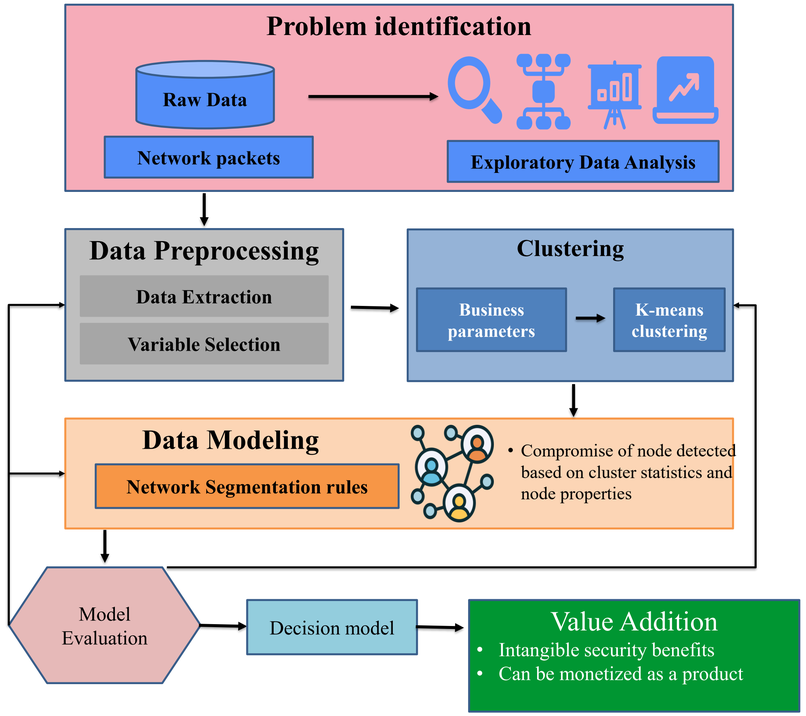
Research Questions:

- What is the optimal network design to minimize the risk of infiltration?
- What are the factors that affect the network strength?

LITERATURE REVIEW

Our work intends to suggest a network segmentation algorithm based on AI/ML techniques. Buczak et al. (2016) detail a focused literature survey of machine learning (ML) and data mining (DM) methods in support of intrusion detection.
Our study is based primarily on network packets that are transmitted over LAN. A specific application programming interface (API) called pcap is used to capture the network packets received and transmitted at the physical interface (e.g., Ethernet port) of the computer. Our work aims to perform judicious feature selection and data preprocessing with the goal of impactful network segmentation and anomaly detection. To that end, we devised an exhaustive set of features for creating network segmentation using K-means clustering algorithm. The anomaly detection was based on the cluster characteristics and historical data for the network.

METHODOLOGY



STATISTICAL RESULTS

The optimal set of features for model creation was established and clusters were created using K-means algorithm. The network was consequently divided into 7 segments which identifies various node groups in the network. Malicious traffic detection was based on historical data analysis and any deviation from the historical data was flagged as a violation or an anomaly.

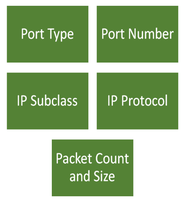


Fig 3. Model Parameters

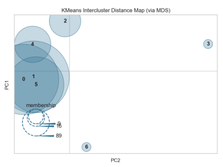


Fig 4. Network Segmentation Representation

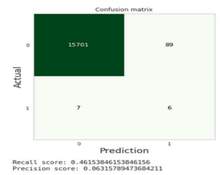


Fig 5. Confusion Matrix

BUSINESS IMPACT

INTAGIBLE SECURITY BENEFITS:
Network segmentation and monitoring helps companies avoid to pay ransom to recover from cyberattacks.

- Enhanced brand value makes customers and suppliers more loyal and thereby reduce the cost of doing business for the company.
- Protects the intellectual property and competitive advantage of a company from outside cyber threats.

PRODUCT MONETIZATION:

- Network segmentation and network security requirements are universal for all companies and our cybersecurity client has the potential to develop a product and service around network security for companies with small IT teams.
- Given first market entrant, assuming 20% market share the company has potential of $100 million increased revenues annually.

CONCLUSIONS

- Port parameters, network direction and the cluster to which the network node belongs significantly determines the type of traffic.
- Decision model effectively identifies the malicious traffic and blocks it thereby decreasing the risk of infiltration.
- With the Cost Benefit Analysis, Model helps to save the business more than $144K in value
- Network segmentation decision model is universal and can be replicated across different industries to protect businesses from cyber-attacks.

FUTURE SCOPE

- Our study incorporates TCP, UDP packets only for network segmentation quality and anomaly detection results. Further analysis based on other protocols would enhance segmentation quality.
- Present analysis is based on anomaly detection from a huge history data about the network. Future scope of work can be based on the deception triggered anomaly detection.
- Insights from our anomaly detection framework can be integrated with the enterprise site reliability engineering solutions to correlate other security event data with our anomaly alerts to generate a lot of insights about the adversary behavior in the network.

ACKNOWLEDGEMENTS

We would like to thank Professor Matthew Lanham and our corporate partners for their guidance and support on this project.