UNIVERSITY OF
WEST LONDON
The Career University

Deep Learning Framework for cyber attack rate prediction

Dissertation Proposal

Research Methods

Date:

Supervisor:

Name:

Student ID:

Abstract

Nowadays “Cyberattack” is a very common word in our society. Most of the people talk about cyberattacks because of their popularity, the attention they gained and the damages they caused because of their high severities. According to the statistics 62% of the companies experienced phishing and social engineering attacks only in 2018(Milkovich, 2019). Not only that but also cyber-attacks increased by 37% over January to April this year with COVID19 pandemic (Muncaster, 2020). In order to minimize the cyberattacks, most of the company networks have set up honeypots and network telescopes to gather information about attacks. In this research proposal, I am proposing to implement a cyber attack rate prediction model based on deep neural networks using these valuable datasets. Due to the high accuracy prediction rates of this model, the network defenders can plan and allocate their resources to mitigate and minimise the attacks efficiently and protect the companies from threats.

This document consists of 5 sections. The first section is about the literature review conducted. It includes details about the past contribution to this research domain and explains the gap analysis. The second section describes the statement of the problem which is going to be addressed in this research. Then the next section provides more details about the aims and objectives of this research. The Methodology section outlines the planned steps to archive all the objectives mentioned in this proposal. The report concludes with a proposed schedule with a Gantt chart.

Table of Contents

Introduction

Information technology revolution opens thousands of new opportunities for the community to enhance their business and day to day life. Currently, researchers and manufacturers introduce new ways to enhance human life and to ease our manual tasks. For example, thousands of years ago humans used fire, pigeons as communication methods. But now, the world is developed to use communication through the internet in a single second. Everything is adapted to use the internet and service providers do a major role by providing internet services. Communication through the internet also consists of multiple layers including an end to end nodes such as computers, tablets or even IoT devices, the ports in those devices, wired connections, wireless connections, topologies in the network, network devices and many more. Most of them are full of known and unknown vulnerabilities. These vulnerabilities open up new opportunities for cyber hackers to attack the systems and networks to gain accessibility or damage the systems. These damages make reputational and monetary losses for these companies.

Hence, it is a major commitment of the company to protect the networks and resources from intruders and hackers. There are lots of services like intrusion prevention and detection management systems, anti-virus software, firewalls installed for the network and for its security. Apart from the cybersecurity team of the company, the attackers also learn and improve their attacking skills, identify new vulnerabilities day by day and use different patterns for their attacks. Therefore, it is a mandatory task to identify and predict the attack before it damages the system and allocates the necessary resources to secure the networks. Moreover, these resources should be allocated according to the priority of the attacks in order to safeguard the systems.

Honeypots and Network Telescopes are popular cyber defence instruments installed in networks to observe the internet traffic of a network. Honeypots are used to attract attackers by acting as real production servers and collect information on the attacks and attack types by interacting with them(Huang et al., 2019). There are 3 types of Honeypots available. They are,

- **Low - interaction honeypots**

These types of honeypots collect low-level information about attackers, IP addresses, attack frequencies etc. These honeypots will not provide any physical environment for the attackers.

- **Medium - interaction honeypots**

These types of honeypots are more interactive with the attacker to collect more information. But they also do not provide any physical environment for the attackers.

- **High - interaction honeypots**

These types of honeypots are more interactive than the other two types. It collects more information about attacks including attack types, originated countries, attacked ports etc.

Large networks consist of multiple honeypots and they are known as honeynets. AWS like services have honeynets to secure their networks all around the world. Network telescopes or darknets are also popular cyber defence instruments. They monitor all traffic directed to unused subnets within a locally allocated address space. Through honeypot data and telescope data, the companies can identify denial of service attacks, malware, botnets, worms, internet messaging threats and targeted attacks. Instead of analysing these honeypot and telescope datasets only by the intrusion detection systems, nowadays researchers have put more effort into identifying attack features, attackers behaviours, extract threat actor tactics, techniques and procedures accurately in order to predict attacks ahead of a few hours earlier(Huang et al., 2019). This provides the ability to allocate the resources accordingly to secure the whole network properly.

The researchers have introduced different techniques to predict cyber attacks and attack rates using honey pot and telescope data. Most of them are based on time series analysis using algorithms like ARIMA, FARIMA and Extreme value theory. These different research methodologies address different concepts in cyber attack rate predictions and overall all the methodologies try to provide high accuracy prediction results. The research proposed by this report is to introduce a framework for predicting cyberattack rates using deep neural networks. The literature review section explains more about the current research status in predicting cyberattack rates using honeypot and network telescope data.

Literature Review

The researchers have conducted different approaches to predict cyberattack types and attack rates based on different algorithms and theories using honeypot and telescope data. According to the literature review, there is very limited research conducted on this topic for honeypot and network telescope data. The following paragraphs explain more about the interesting and distinct approaches available for this research domain.

A novel concept of a mathematical approach to naturally predict cyber attacks called stochastic cyber attack process was introduced in a research led by Zhenxin in 2013. Furthermore, he had identified the occurrence of Long-Range Dependence(LRD) feature in honeypot data for the first time. Before this finding, the previous studies were conducted only for Short Range Dependencies(SRD) in honeypot data. This research proved that around 80% of network-level attacks, 44.5% port level attacks and 70% of victim level attacks are exhibits of LRD features in attack processes. Hence the introduction of the Gray box method to accommodate this LRD phenomenon in honeypot time series was a major achievement in this research. To facilitate this phenomenon, Gray box algorithm consists of three statistical models for LRD less, LRD aware attacks and SRD using LRD-Less ARMA(Auto Regressive Moving Average model) and LRD aware FARIMA(Fractional Autoregressive Integrated moving average) models respectively. The Gray box selects the best algorithm at the prediction step to increase the accuracy. Percent Mean Absolute Deviation(PMAD) is used for the accuracy calculations throughout the research. The whole research was conducted by assuming the honeypot dataset as stationary data. It is one of the main limitations of this research. Furthermore, this research was conducted only for 5 time periods, therefore the accuracy and the challenges with more data were not tested properly(Zhan et al., 2013).

Based on the previous research, Zhenxin and his team had improved the cyber attack predictions using the Gray box model which accommodates LRD features in the cyberattack rate time series. In this research, they had identified the drawbacks of using the Gray box method in extreme attack rates. Extreme attack rates are the time series which received frequent attacks for a given time unit. To calculate that, researchers set up a threshold value for the time series data. For example, in this dataset, the threshold was set to 90%. They introduced FARIMA(Fractional Autoregressive Integrated moving average) and GARCH(Generalized AutoRegressive Conditional Heteroskedasticity) methods to accommodate LRD features and extreme cyber-attack rate in order to predict attacks with high accuracy. They tested 3 honeypot datasets with Gray box approach and FARIMA+GARCH method with extreme value theory. They found an improved accuracy in the new model. Moreover, they tested the same dataset with Hidden Markov models and Symbolic Dynamic models as well. But the approach with FARIMA+ GARCH models had given the highest accuracy rates. Apart from increasing accuracy levels, this model can predict the cyberattack rates in the short term and long term datasets. The major gap in this research is that the researchers assumed that these honeypot datasets are stationary time series without any seasonality and trends. Hence they used classical point over threshold(POT) methods to accommodate extreme value theory in cyber attack rate time series data(Zhan et al., 2015).

In 2016, Cheng Peng conducted research to model and predict cyber-attack rates using extreme value property which is a statistical property that exists in the cyberattack rates time series data. This value calculates the number of attacks for some targets per given time unit. In his research, he found that the cyber attack rate time series does not follow a Poisson process. In the simplest words, he found that the occurrence of extreme values called inter exceedance time are independent of each other in this time series and follows an exponential distribution. He found that the classical point over threshold(POT) methods accommodate only the time series with Poisson distributions. In order to incorporate inter exceedance time to POT method, he introduced a marked point process which allows both arrivals of extreme attack rates and calculates the magnitude of the exceedance. By using the marked point process over other time series analysis methods mentioned above, this research had performed more accuracy in prediction, high performance in the process of sample fitting and out of space prediction performances. Furthermore, this research used value at Risk(VaR) as the natural measure of intense attack. A log autoregressive conditional duration approach was studied to describe the arrival of extreme cyber-attack rates. The main achievement of this research was the high accuracy of predicting cyberattack rates than the models described earlier(Peng et al., 2016). The drawbacks of this research were,

- This research was conducted only considering 1h as the optimal time resolution for an accurate cyber attack prediction.
- The accuracy of the prediction can be increased by understanding the correlation between the number of attackers, cyberattack victim time series with cyberattack rates(Peng et al., 2016).

With the development of machine learning the researchers tend to use other new technologies like deep learning to these prediction approaches. One of the main studies conducted using deep neural networks is the introduction of BRNN-LSTM by Fang and his team(Fang et al., 2019). This study used a novel bi-directional recurrent neural network with long short term memory framework(BRNN-LSTM). The bi-directional recurrent neural network(RNN) is a feed-forward network which helps the network to train itself and increase the prediction accuracy with the frequency of usage. LSTM is in-memory states which are used to store memory status at different nodes and they help to increase the performance. The training process of RNN can cause gradient vanishing problems and LSTM are used to fix it. The whole framework uses statistical properties of cyberattack rates time-series data(Fang et al., 2019).

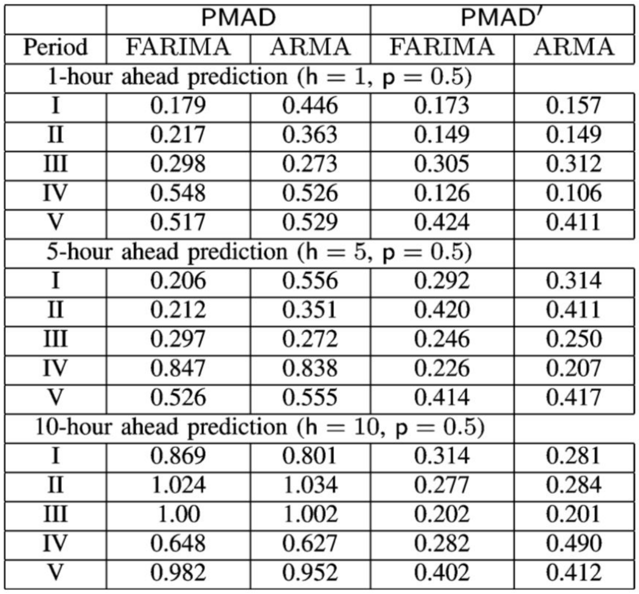
The accuracy measurement like present mean absolute deviation and mean absolute percentage error values achieved remarkably high prediction accuracy rate for BRNN-LSTM than other models. The data preprocessing step is avoided in the

BRNN-LSTM framework and the selection of fitted values is calculated using an algorithm. Furthermore, the researchers conducted a comparison against other analytical approaches with this deep learning approach and found that the deep learning approach is more accurate and reduces error rates than other models(NAMINAMIN, 2018).

However, the authors found that this BRNN-LSMT framework had missed the observed values on some occasions and these occasions are not predictable. The authors assumed the prediction accuracy as sufficient throughout the paper but this can vary from different situations. So there is more to improve in this concept to maximize accuracy and performance.

The following table illustrates the accuracy comparison for different models discussed above.

Accuracy comparison

<table>
<tr><th>Process</th><th>Accuracy</th></tr>
<tr><td>Gray Model(Zhan et al., 2013)</td><td>

<table>
<tr><th></th><th>PMAD</th><th></th><th>PMAD′</th><th></th></tr>
<tr><th>Period</th><th>FARIMA</th><th>ARMA</th><th>FARIMA</th><th>ARMA</th></tr>
<tr><td colspan="4">1-hour ahead prediction (h = 1, p = 0.5)</td><td></td></tr>
<tr><td>I</td><td>0.179</td><td>0.446</td><td>0.173</td><td>0.157</td></tr>
<tr><td>II</td><td>0.217</td><td>0.363</td><td>0.149</td><td>0.149</td></tr>
<tr><td>III</td><td>0.298</td><td>0.273</td><td>0.305</td><td>0.312</td></tr>
<tr><td>IV</td><td>0.548</td><td>0.526</td><td>0.126</td><td>0.106</td></tr>
<tr><td>V</td><td>0.517</td><td>0.529</td><td>0.424</td><td>0.411</td></tr>
<tr><td colspan="4">5-hour ahead prediction (h = 5, p = 0.5)</td><td></td></tr>
<tr><td>I</td><td>0.206</td><td>0.556</td><td>0.292</td><td>0.314</td></tr>
<tr><td>II</td><td>0.212</td><td>0.351</td><td>0.420</td><td>0.411</td></tr>
<tr><td>III</td><td>0.297</td><td>0.272</td><td>0.246</td><td>0.250</td></tr>
<tr><td>IV</td><td>0.847</td><td>0.838</td><td>0.226</td><td>0.207</td></tr>
<tr><td>V</td><td>0.526</td><td>0.555</td><td>0.414</td><td>0.417</td></tr>
<tr><td colspan="4">10-hour ahead prediction (h = 10, p = 0.5)</td><td></td></tr>
<tr><td>I</td><td>0.869</td><td>0.801</td><td>0.314</td><td>0.281</td></tr>
<tr><td>II</td><td>1.024</td><td>1.034</td><td>0.277</td><td>0.284</td></tr>
<tr><td>III</td><td>1.00</td><td>1.002</td><td>0.202</td><td>0.201</td></tr>
<tr><td>IV</td><td>0.648</td><td>0.627</td><td>0.282</td><td>0.490</td></tr>
<tr><td>V</td><td>0.982</td><td>0.952</td><td>0.402</td><td>0.412</td></tr>
</table>
Table 1: Accuracy calculations in Gray Model
This figure illustrates the calculated PMAD and PMAD` values for 5 time periods of honeypot data using the Gray model. From these accuracy calculations, the researchers expressed that this model is able to predict cyberattack rates in 10h ahead with high accuracy.
</td></tr>
<tr><td>FARIMA + GARCH models with extreme values (Zhan et al., 2015)</td><td>
Prediction accuracy rates for 2 honeypot data is calculated for the Gray Box model and FARIMA+ GARCH models.
<table>
<tr><th>Model</th><th>DataSet 1</th><th>DataSet 2</th><th>DataSet 3</th></tr>
<tr><td>Gray Box</td><td>86.2%</td><td>87.9%</td><td>86.0%</td></tr>
<tr><td>FARIMA+ GARCH</td><td>90.3%</td><td>97.5%</td><td>101.8%</td></tr>
</table>
Table 2: Accuracy calculations in FARIMA+GARCH Model
These values were compared with hidden Markov models and Symbolic dynamics
</td></tr>
</table>

<table>
<tr><td></td><td>models as well and found that the FARIMA+GARCH models predict the cyberattack rates in higher accuracy than the other models(Zhan et al., 2015).</td></tr>
<tr><td>Marked point processes(Peng et al., 2016)</td><td>Performance of the prediction is calculated for different hours ahead(1h,2h,10h) using observed and expected violations measured using VaR values of LR_{uc}, LR_{ind} and LR_{cc}. According to the values, this model is capable of predicting the cyberattack rates ahead of 10h correctly. That is a high achievement compared to other models discussed in this literature review(Peng et al., 2016).</td></tr>
<tr><td>BRNN-LSTM(Fang et al., 2019)</td><td>
<table>
<tr><td>Dataset</td><td>MSE</td><td>MAD</td><td>PMAD</td><td>MAPE</td></tr>
<tr><td>1</td><td>3,628,266</td><td>463.2715</td><td>.0124374</td><td>.0138781</td></tr>
<tr><td>2</td><td>16,497,941</td><td>1036.604</td><td>.0401286</td><td>.0481919</td></tr>
<tr><td>3</td><td>30,637,599</td><td>675.7551</td><td>.0429913</td><td>.0230468</td></tr>
<tr><td>4</td><td>2,165,707</td><td>508.3557</td><td>.1658243</td><td>.2656372</td></tr>
<tr><td>4*</td><td>1,085,361</td><td>297.3440</td><td>.1034426</td><td>.1338577</td></tr>
<tr><td>5</td><td>20,415,119</td><td>1396.763</td><td>.0356409</td><td>.0478739</td></tr>
</table>
Table 3: Accuracy calculations in BRNN-LSTM Model
According to the accuracy calculations, BRNN-LSTM had given high accuracy values for dataset 1,2,3 and 5 by having less than 5% of error rates. But for Dataset 4 it received around 17% error rate, hence the model is recalculated using a rolling approach(Dataset 4*). After the recalculation, the dataset 4 also provided the same high accuracy with less error rate. Furthermore, these datasets are calculated against ARIMA, FARIMA_GRACH models. It proved</td></tr>
</table>

	that this neural network approach provided higher accuracy than other models. The accuracy results can be found in the appendix section (Table 6).

Table 4: Accuracy calculations

Statement of Problem

According to the gap analysis conducted, different researchers contributed different ways to predict cyberattack rates in high accuracy. Most of the latest researches are based on the previous ones and have fixed most of the drawbacks and gaps available on those research models. Not only that by taking all the challenges and the gap analysis conducted in the literature review it shows that current prediction methodologies need to be updated with high accuracy prediction models to warn the network defenders about cyber attacks in a few hours ahead. It will allow the defenders to allocate adequate defence resources purposefully to manage the attacks and protect the whole network(Peng et al., 2016).

The main problems identified in this research domain are,

1. Most of the currently available models use traditional statistical approaches like ARIMA, FARIMA models to predict the attack rates. These models are outdated these days. Therefore the models need to be upgraded.
2. These models need to update to handle new attack patterns as the attacks are varying massively and getting complex day by day
3. Most of the current models cannot predict extreme cyberattack rates
4. Most of the currently available models are not able to learn by themself and upgrade the model
5. Most of the currently available models cannot predict the attacks more than 5-10h ahead.
6. These models cannot predict the cyberattack types because they use low interaction honeypot data

7. Most of the available models can predict cyberattacks based on honeypot or telescope data only. They cannot predict cyber attack data based on both types of datasets.

Aims and Objectives

This proposed research is planned to implement on top of the research conducted by Fang(Fang et al., 2019) to build a deep neural network approach to predict the cyberattack rates. This proposed solution will contribute to the following gaps in the current domain.

- Increase the accuracy level of the deep neural network by increasing the performance of the network to identify long-range and short-range of dependence and high non-linearity of the data.
- Introduce a marked point process or better solution to the deep neural network model to handle extreme attack rates.
- Provide more data about the predicted attacks based on the honeypot and telescope data. For example the origin, type of attack etc.
- Provide predicted attack data more than 10h ahead of time with high accuracy.
- Increase the responsiveness of the attack patterns by training the network to learn by itself.
- Decrease the computational power for the prediction and increase the performance.
- Design the model to predict real-time data.
- Provide more details about the predicted attack in order to allocate sufficient defender resources by the defenders.

Methodologies

This section describes the methodology of the proposed research in detail.

1. Conduct a comprehensive literature review -

   In this step, I am planning to conduct a comprehensive study in order to fulfil the aims and the objectives of this research. Hence this step will include a literature review from understanding the honeypot and telescope data, until troubleshooting the neural network in order to increase the performance of the prediction.

2. Collect honeypot and telescope datasets including high interaction honeypots datasets

   This step is to focus on collecting meaningful and publicly available honeypot and network telescope data sets.

   Example:

   Network Telescope data: The CAIDA UCSD Network Telescope Educational Dataset: Analysis of Unidirectional IP Traffic to Darkspace

   Honeypot Data:

   https://www.kaggle.com/jonathanbouchet/aws-honeypot#detection-by-host

3. Select the best technologies to use

   This step is mainly focused on identifying the best technologies to implement prediction models using deep neural networks, distributed cluster computing frameworks, technologies to handle real-time processing, technologies to visualize data.

4. Analyse the datasets and get insights

   In this step, the dataset will be analyzed to identify the data available. In addition to that, the honeypot data will be mapped with telescope data and identify important features in the datasets.

5. Preprocess the datasets

   The dataset will be preprocessed according to the procedures introduced in the following papers and the procedures collected from the literature review.

   Honeypot data: Characterization of Attackers' Activities in Honeypot Traffic Using Principal Component Analysis (Almotairi et al., 2008)

   Network Telescope data: A parameterizable methodology for Internet traffic flow profiling(Claffy et al., 1995)

6. Model creation

   This model will be implemented to archive all the aims and objectives mentioned in the earlier section. The model creation can be subdivided into the following phases.

   I. Neural Network implementation by identifying the best number of layers, models and neural network types to use in order to achieve maximum prediction power, high performance and low computational power.
   II. Implement the network as described as in the Fang's research paper(Fang et al., 2019) and find solutions for the drawbacks and gaps mentioned in his paper.
   III. Convert this model to handle the real-time streaming data processing in order to make it highly responsive to the time series.
   IV. Introduce a marked point process to handle extreme attack rates exhibited in honeypot and telescope data.
   V. Introduce a model to get more data about the attack including the possible time of the attack, possible types of attack. For example attack to a port etc.

7. Test and verifying the results

   Training and testing the model with real data. Then compare with the actual data and calculate error values. Based on the results fine-tune the neural network to provide maximum accuracy.

8. Visualise/Provide meaningful and human-understandable data based on the prediction

   Provide the prediction about 10h ahead of the attack and provide more data related to the attack in a human-understandable way.

9. Compare the results with other models

   Compare the prediction results with other models discussed in the literature review.

10. Complete the dissertation report
11. Publish the work

Schedule

The following Gantt chart exemplifies the planned schedule for the proposed research.

Task	Starting Date	End Date	Duration(days)
Literature Review	05-25-2020	06-07-2020	14
Collect Datasets	05-25-2020	06-07-2020	14
Selection of Technologies	05-25-2020	05-31-2020	7
Analyse the datasets	06-01-2020	06-07-2020	7
Preprocess the datasets	06-08-2020	06-21-2020	14
Model creation			
1. Initializing the neural network implementation	06-22-2020	07-05-2020	14
2. Update the neural network according to Fang's research	06-29-2020	07-05-2020	7
3. Handle real-time data	07-06-2020	07-08-2020	3
4. Introduce marked point process	07-09-2020	07-12-2020	4
5. Create a model to get more data	07-13-2020	07-26-2020	14
Test and verifying the results	07-15-2020	07-29-2020	14
Visualise data	07-27-2020	08-02-2020	7
Analysis and evaluation	08-03-2020	08-16-2020	14
Complete the dissertation report	08-17-2020	09-11-2020	26

Table 5: Planned schedule for this proposed solution

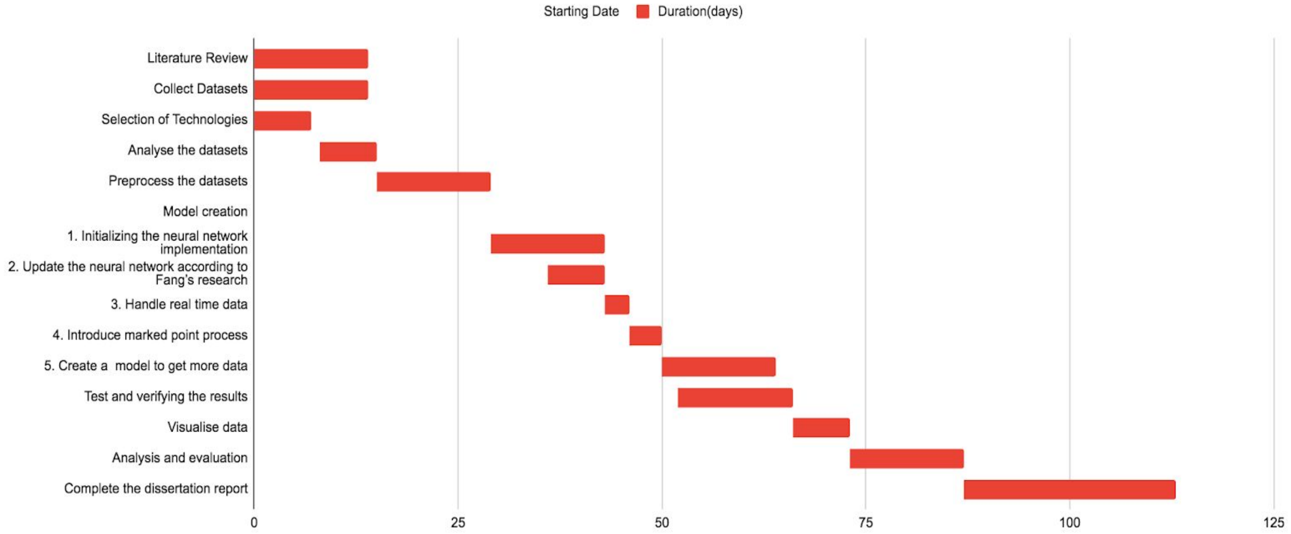


Figure 1: Gantt Chart for the planned schedule

References

Almotairi, S., Clark, A., Mohay, G. and Zimmermann, J. (2008) Characterization of Attackers' Activities in Honeypot Traffic Using Principal Component Analysis. *2008 IFIP International Conference on Network and Parallel Computing*. *IEEE* [Online]. Available at: doi:10.1109/npc.2008.82 [Accessed: 3 June 2020].

Claffy, K., Braun, H. and Polyzos, G. (1995) A parameterizable methodology for Internet traffic flow profiling. *IEEE Journal on Selected Areas in Communications*, 13 (8), p.1481-1494. *Institute of Electrical and Electronics Engineers (IEEE)* [Online]. Available at: doi:10.1109/49.464717 [Accessed: 1 June 2020].

Fang, X., Xu, M., Xu, S. and Zhao, P. (2019) A deep learning framework for predicting cyber attacks rates. *EURASIP Journal on Information Security*, 2019 (1). *Springer Science and Business Media LLC* [Online]. Available at: doi:10.1186/s13635-019-0090-6 [Accessed: 15 May 2020]

Huang, C., Han, J., Zhang, X. and Liu, J. (2019) Automatic Identification of Honeypot Server Using Machine Learning Techniques. *Security and Communication Networks*, 2019, p.1-8. *Hindawi Limited* [Online]. Available at: doi:10.1155/2019/2627608 [Accessed: 27 May 2020].

Milkovich, D. (2019) 15 Alarming Cyber Security Facts and Stats | Cybint. *Cybint*. [Online]. Available at: https://www.cybintsolutions.com/cyber-security-facts-stats/ [Accessed: 6 June 2020].

Muncaster, P. (2020) Cyber-Attacks Up 37% Over Past Month as #COVID19 Bites. *Infosecurity Magazine*. [Online]. Available at: https://www.infosecurity-magazine.com/news/cyberattacks-up-37-over-past-month/ [Accessed: 2 April 2020].

NAMIN, S. and NAMIN, A. (2018) *FORECASTING ECONOMIC AND FINANCIAL TIME SERIES: ARIMA VS. LSTM*. [Online]. Available at: doi:https://arxiv.org/pdf/1803.06386.pdf [Accessed: 21 May 2020].

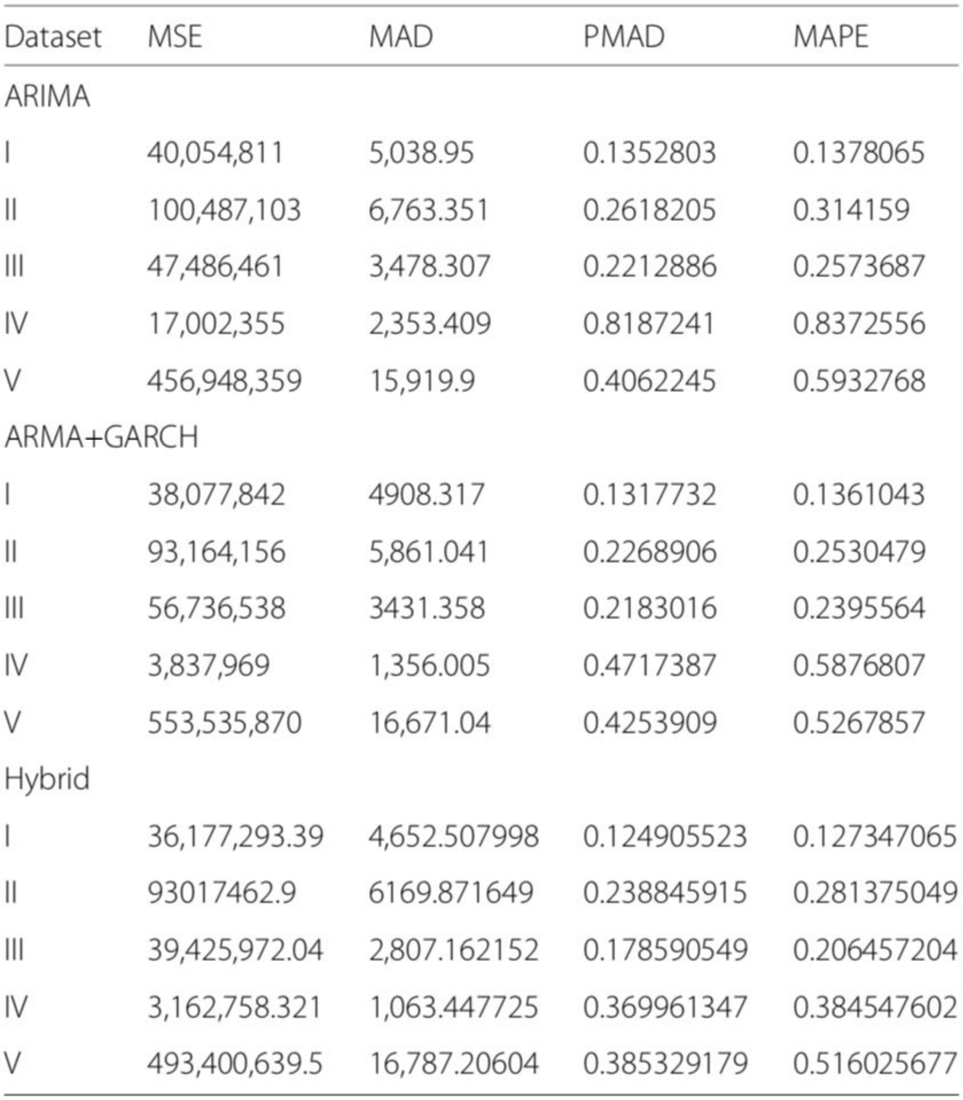
Peng, C., Xu, M., Xu, S. and Hu, T. (2016) Modeling and predicting extreme cyber attack rates via marked point processes. *Journal of Applied Statistics*, 44 (14), p.2534-2563. *Informa UK Limited* [Online]. Available at: doi:10.1080/02664763.2016.1257590 [Accessed: 31 May 2020].

Zhan, Z., Xu, M. and Xu, S. (2013) Characterizing Honeypot-Captured Cyber Attacks: Statistical Framework and Case Study. *IEEE Transactions on Information Forensics and Security*, 8 (11), p.1775-1789. *Institute of Electrical and Electronics Engineers (IEEE)* [Online]. Available at: doi:10.1109/tifs.2013.2279800.

Zhan, Z., Xu, M. and Xu, S. (2015) Predicting Cyber Attack Rates With Extreme Values. *IEEE Transactions on Information Forensics and Security*, 10 (8), p.1666-1677. *Institute of Electrical and Electronics Engineers (IEEE)* [Online]. Available at: doi:10.1109/tifs.2015.2422261.

.

Appendix



Dataset	MSE	MAD	PMAD	MAPE
ARIMA				
I	40,054,811	5,038.95	0.1352803	0.1378065
II	100,487,103	6,763.351	0.2618205	0.314159
III	47,486,461	3,478.307	0.2212886	0.2573687
IV	17,002,355	2,353.409	0.8187241	0.8372556
V	456,948,359	15,919.9	0.4062245	0.5932768
ARMA+GARCH				
I	38,077,842	4908.317	0.1317732	0.1361043
II	93,164,156	5,861.041	0.2268906	0.2530479
III	56,736,538	3431.358	0.2183016	0.2395564
IV	3,837,969	1,356.005	0.4717387	0.5876807
V	553,535,870	16,671.04	0.4253909	0.5267857
Hybrid				
I	36,177,293.39	4,652.507998	0.124905523	0.127347065
II	93017462.9	6169.871649	0.238845915	0.281375049
III	39,425,972.04	2,807.162152	0.178590549	0.206457204
IV	3,162,758.321	1,063.447725	0.369961347	0.384547602
V	493,400,639.5	16,787.20604	0.385329179	0.516025677

Table 6: Accuracy calculations with other models including ARIMA, ARIMA+GARCH and a hybrid solution(Fang et al., 2019)