Forecasting Cyber-Attacks to Destination Ports Using Machine Learning

Kostas Loumponias[a], Sotiris Raptis[b], Eleni Darra[c], Theodora Tsikrika[d],
Stefanos Vrochidis[e] and Ioannis Kompatsiaris[f]
Information Technologies Institute, Centre for Research and Technology Hellas-CERTH, GR-54124, Thessaloniki, Greece

Keywords: Cyber-Attack, Network Traffic, Forecasting, Destination Port.

Abstract: To anticipate and counter cyber-attacks that may threaten the stability of the economy, society, and governments around the world, significant efforts have made particularly towards the detection of cyber-attacks, while fewer studies have focused on their forecasting. This paper proposes a framework that provides forecasts of upcoming (within the next minute) cyber-attacks, as well as their type, to a specific destination port. To this end, several machine learning-based methods are applied on measurements (observations) obtained from the network traffic flow. The proposed method is supported by two major pillars: first, the selection of appropriate features generated by the network traffic and, second, in addition to the selected features, the detection of the type of cyber-attacks that occurred in the past. The proposed framework is evaluated on the CIC-IDS2017 synthetic dataset and provides a robust performance in forecasting the type of upcoming cyber-attack in terms of Accuracy, Precision, Recall, F1-score and confusion matrix.

1 INTRODUCTION

Cyber-attacks have been continuously increasing in recent years both in number and in type. Their impact can be significant as cyber-attacks can lead to power outages, equipment failure in several domains, leaks of sensitive information, e.g. related to national security, as well as to the theft of valuable and private information, such as medical records. Ultimately, they can result in the shutdown of critical systems, bringing down computer networks, and preventing access to data. It is not a stretch to argue that cyber-attacks have the potential to disrupt everyday life as we know it. Therefore, a variety of approaches have been proposed to stop or mitigate the devastating effects of cyber-attacks. In particular, cyber-attack countermeasure approaches can be categorized into (1) cyber-attack detection, (2) attack projection frameworks, (3) graphical systems, and (4) forecasting methods.

Typical *cyber-attack detection systems* use a misuse-based approach, in which events that are being watched are compared to the signatures of incidents that have already been seen. The difficulty of such systems to identify novel events whose fingerprints are unknown to the detection systems is one of their limitations (Dutta et al., 2022). Anomaly detection is another technique that aims at detecting deviations from normal behaviour and labeling them as malicious (Blowers and Williams, 2014), (Liakos et al., 2020). These methods frequently produce many false alarms because they may classify as anomalous actions that are in fact normal, but not previously seen.

Attack projection frameworks (Yang et al., 2014) simulate the development of an assault over time and can be described as a network attack modeling approach for threat projection. This modeling goes beyond the conventional definition of intrusion detection systems, which uses observables of active assaults to forecast future hostile behavior based on system flaws and attacker behavior. The focus of attack projection is on the traces left by multistage cyber-attacks. While these methods have found success, they require access to the victim's network and can only detect attacks that are in progress. The study and comprehension of complicated patterns can sometimes be challenging for viewers to visualize.

Graphical systems and well-designed diagrams can help in understanding the nature of cyber-attack and offer insights on how to best deal with them. Attack graphs and attack trees may represent cyber-

[a] https://orcid.org/0000-0002-6268-3893
[b] https://orcid.org/0000-0001-7040-966X
[c] https://orcid.org/0000-0002-8496-9999
[d] https://orcid.org/0000-0003-4148-9028
[e] https://orcid.org/0000-0002-2505-9178
[f] https://orcid.org/0000-0001-6447-9020

Loumponias, K., Raptis, S., Darra, E., Tsikrika, T., Vrochidis, S. and Kompatsiaris, I.
Forecasting Cyber-Attacks to Destination Ports Using Machine Learning.
DOI: 10.5220/0011891000003405
In *Proceedings of the 9th International Conference on Information Systems Security and Privacy (ICISSP 2023)*, pages 757-764
ISBN: 978-989-758-624-8; ISSN: 2184-4356

attacks in terms of their visual syntax to correlate, hypothesize, and predict intrusion alerts (Lallie et al., 2020). The main goal of an attack graph is to provide an effective representation and algorithmic tools to identify the scenarios in which system vulnerabilities in a network might be exploited, but this strategy heavily depends on having a thorough understanding of the firewall rules and system vulnerabilities in the network (Polatidis et al., 2020), (Polatidis and Georgiadis, 2016). However, this strategy offers little insight into the precise nature of potential assaults.

Effective defenses could be designed to stop the devastating effects of cyber-attacks if they are predicted before they happen. *Forecasting methods* have been used to this end, with particular focus on predicting the number of attacks for the next time step using the historical cases of the previous period (Bakdash et al., 2018), (Kwon et al., 2017) and their impact (Ji et al., 2022). These methods are a relatively new area of research and while several machine learning (ML) techniques have already been used to predict cyber-attacks, Deep Learning methods have recently been found to be particularly effective in time series forecasting (Barreto and Koutsoukos, 2019). In particular, analysis of time series data with uncertainties and/or certain unobservable elements has been developed using Gaussian mixture models, hidden Markov models, and state-space models (Brockwell and Davis, 2016).

This paper proposes a framework that provides next-minute forecasts of cyber-attacks, where these forecasts also include the type of cyber-attack (DoS, DDoS, etc.), by considering network traffic data provided by the Intrusion Detection Evaluation dataset (CIC-IDS2017) (Sharafaldin et al., 2018). More specifically, the proposed framework is destination port (DP) oriented, since it considers only measurements of the target DP to provide forecasts of upcoming cyber-attacks (forecasts for the next minute). At this point, it is worth mentioning that the term *forecast* should not be confused with the term *prediction*, as cyber-attack prediction refers (in the literature) to the detection of cyber-attacks and not to the forecasting of new upcoming cyber-attacks, which is the case study in this paper. To the best of our knowledge, network traffic data has been extensively used to detect network anomalies and intrusions (Thapa et al., 2020), (Khan et al., 2021) or to forecast the frequency and the risk level of cyber-attacks (Yang et al., 2021), (Ji et al., 2022), but not to forecast the type of upcoming cyber-attacks.

The selection of appropriate features generated by analyzing the network traffic has substantial impact on providing robust forecasts (Ji et al., 2022). In this study, only features that do not strongly correlate with one another are selected since prediction models do not benefit from extra information from features with roughly comparable patterns. The detected type of cyber-attacks that occurred in the past is also utilized in order to forecast the cyber-attacks that will occur in the next minute, in addition to the chosen features. The type of previous cyber-attacks is detected by the Random Forest (RF) algorithm (Svetnik et al., 2003). Next, well established ML methods, such as Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997), Multilayer Perceptron (MLP), and Logistic Regression (LR) (Dreiseitl and Ohno-Machado, 2002) models are applied to the extracted features to provide cyber-attacks forecasts. The experimental results in synthetic datasets show that the proposed framework provides high accuracy for all ML methods in forecasting next minute cyber-attacks, as well as their type.

Overall, the main contribution of this work is a framework that provides forecasts of upcoming cyber-attacks as well as their type to a specific destination port, through the selection and subsequent use of specific features of network traffic in a novel way, along with the use of the prediction (i.e., detection) of the type of cyber-attacks that have occurred in the past.

The rest of the paper is organized as follows: In Section 2, related works are reported. In Section 3, the proposed framework is provided. In Section 4, experimental results are presented using synthetic datasets to demonstrate the effectiveness of the proposed framework. Finally, in Section 5, conclusions and future work are discussed.

2 RELATED WORK

In (Okutan et al., 2019), an automated system, named CAPTURE, is presented. This system uses a range of unconventional signals to forecast the occurrences of endpoint-malware and malicious email for a target organization. Novel methods have been developed such as Entropy-based Lagged Feature Selection (ELFS) that selects the significant signals with specific lags, and Concept Drift based Training Window (CDTW) that dynamically finds the non-stationary relationships between the unconventional signals and the attack occurrences. Integrating both of these methods, along with other components, CAPTURE is developed. CAPTURE selects the relevant signals with the right lags and the corresponding training set to produce better forecasts. A detailed examination of the individual forecast confidences shows that CAPTURE offers better differentiation between the days cyber-attacks occur from those without. Furthermore, CAPTURE is able to allow the analysts to evaluate the relevant lagged signals and how they collectively lead to the forecasts.

In (Ahmet Okutan and McConky, 2018) the occurrence of a cyber-attack towards an entity is forecast by using unconventional signals from various data sources that may or may not be related to that target entity. They make use of Twitter and the open source GDELT project (Leetaru and Schrodt, 2013) for unconventional signals. The signals are not directly linked to specific vulnerabilities. Additionally, a methodology based on Bayesian networks is presented, which can treat a variety of unconventional signals to forecast events that do not necessarily have balanced positive and negative ground truth instances.

In (Goyal et al., 2018) two concepts of ML, LBFGS method (Seabold and Perktold, 2010) and Adaptive Moment Estimation (Kingma and Ba, 2014a), for forecasting cyber-attacks are used. The two methods take as input historical cyber-attacks to train ML models that provide forecasts about the frequency of malware attacks. These models capture patterns present in historical data that enhance the forecasting accuracy. Authors propose that they can increase the forecasting accuracy of these models by leveraging signals from external Web data sources. From these data sources, a variety of time series is extracted, each representing the number of daily occurrences of cyber security-related terms. The time series are used as external signals in the forecasting task. The ground truth data about cyber-attacks is used to train the forecasting models, and to evaluate their predictions.

In (Qasaimeh et al., 2022) a network-based cyber-attacks forecasting model is designed to protect the entire bank or financial institution from unknown suspicious activities by anticipating the emergence of new cyber-attacks with novel patterns that made of combination of existing attack. The proposed model was able to forecast new types of network-based cyber behaviors, which were generated from the patterns, features, and activities of well-known attacks, with 99.67% accuracy. The accuracy of the model is 90.36% when it is evaluated and verified in a real life banking test environment that is controlled by specific proactive controls.

The work presented in (Ivanyo et al., 2018) focused on the interval forecasting results of cyber-attacks based on intelligent modeling. A probabilistic neural network (NN) with a dynamic updating value of the smoothing parameter is used. This approach allows carrying out the cyber-attack interval forecasting with a pre-set intensity level of cyber-attacks. The approach demonstrates the high accuracy of cyber-attack interval forecasts for selected data. At the same time, the necessary practical recommendations on an application of the interval forecasting results to the protection against the cyber-attacks in industrial control systems were formulated.

In (Tavabi et al., 2020) some of the challenges associated with forecasting cyber-attacks are identified. The small number of attacks that do penetrate the target's defenses follow a different generative process compared to the whole data which is much harder to learn for predictive models. The loss of predictability is quantified by using real-world data from two organizations. The proposed work identifies the limits to forecasting cyber-attacks from highly filtered data.

In (Yang et al., 2021) a forecasting method based on simulated annealing algorithm, ARIMA and NN techniques is presented in order to forecast the the network traffic (in bytes). The proposed method extracts features from traffic data, combing a linear (ARIMA) and and non-linear (NN models) method in order to forecast the network traffic with high accuracy. In (Ji et al., 2022) the authors forecast the risk level (low, medium and high) and the frequency of upcoming cyber-attacks using network traffic data. Initially, wavelet transform are used to extract features form the network traffic data. Then the vector auto regression with eXogenous variables (VARX), is utilized to forecast future network traffic events (frequencies). Finally, cyber-attack risks for network events are estimated with an adaptive threshold method and assessed by using the support vector machine (SVM) and LR model.

3 PROPOSED FRAMEWORK

In this section, the proposed method to forecast the type of cyber-attacks (if any) to a specific destination port (DP) in the next minute is described in detail. The proposed framework includes the following steps-stages:

Stage 1: Selection of Features: Network traffic measurements include many different features, $\{\mathbf{f}_1, \mathbf{f}_2, ..., \mathbf{f}_n\}$, such as *total packets in the forward direction*, *total packets in the backward direction*, and others. However, some of these features $\mathbf{f}_i$ may present a high degree of correlation (linear or non-linear) among them. Thus, we can take advantage of this observation and reduce the number of features. For instance, in the case where $\mathbf{f}_i$ and $\mathbf{f}_j$ have a high correlation value (e.g. greater than 95%), one of them can be removed, since it does

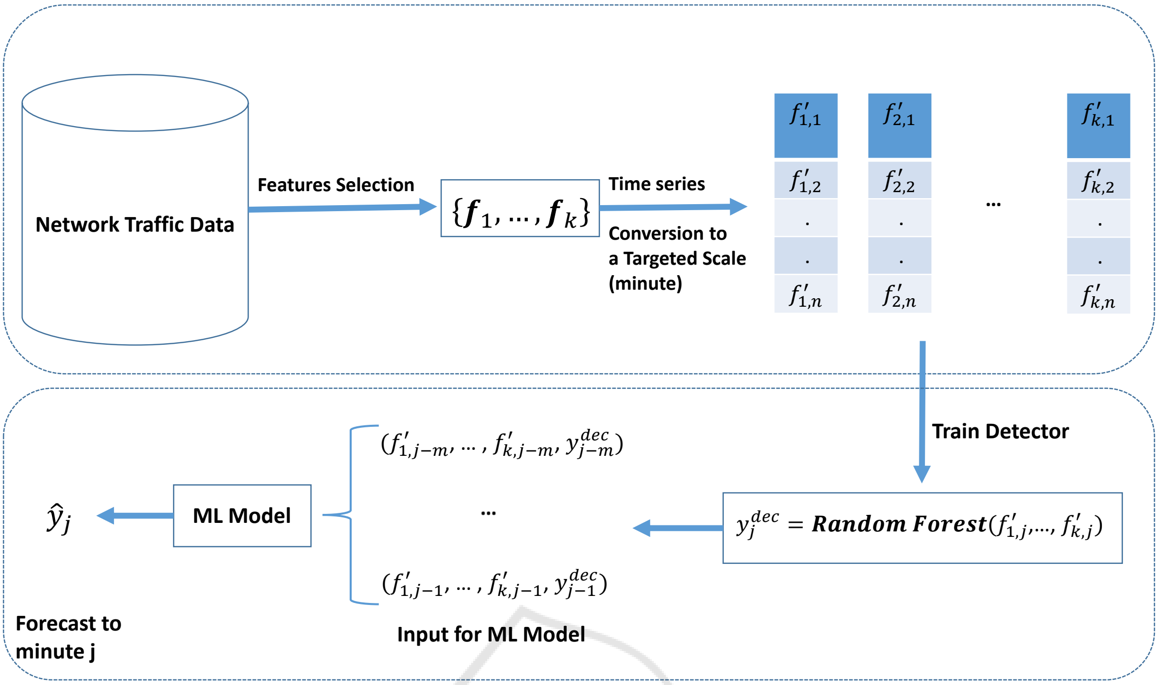
Network Traffic Data
Features Selection
$\{f_1, \ldots, f_k\}$
Time series
Conversion to
a Targeted Scale
(minute)
$f'_{1,1}$ $f'_{2,1}$ $f'_{k,1}$
$f'_{1,2}$ $f'_{2,2}$ $f'_{k,2}$
...
$f'_{1,n}$ $f'_{2,n}$ $f'_{k,n}$
Train Detector
$y_j^{dec} = \textbf{Random Forest}(f'_{1,j}, \ldots, f'_{k,j})$
$(f'_{1,j-m}, \cdots, f'_{k,j-m}, y_{j-m}^{dec})$
...
$(f'_{1,j-1}, \cdots, f'_{k,j-1}, y_{j-1}^{dec})$
Input for ML Model
ML Model
$\hat{y}_j$
Forecast to
minute j

Figure 1: Framework of the proposed approach.

not provide any additional information. In the proposed method, features with Spearman correlation coefficient (De Winter et al., 2016) greater than 95% are considered to have the same impact on predictive models and one of them is removed. From now on, $\{\mathbf{f}_1, \mathbf{f}_2, \ldots., \mathbf{f}_k\}$ denotes the reduced set of features.

Stage 2: Definition of Time-Frame: Without loss of generality, $f_{i,j}$ stands for feature i at time j. In network traffic records, the timestamp typically contains the hour, the minute, and the second, however the latter is not always provided. Furthermore, the number of measurements for a feature may vary significantly at different minutes, e.g., only a couple of measurements may be provided for feature i (i.e., $\mathbf{f}_i$) at one minute, while hundreds of measurements may be provided for the same feature at a different minute. Therefore, to provide a forecast for the *next time-step*, we should define what the *time-step* is based on the available data. In order to tackle the cases where the seconds are not provided, i.e., the timestamp only includes the hour and the minute, the maximum value of $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$ for a period of one minute is considered, i.e.:

$$f'_{i,j} = max\{f_{i,\nu_j:\nu_j+t}\}, \tag{1}$$

where t is the total number of measurements within the minute j and ν_j is the starting point of minute j. Thus, the derived features, $\mathbf{f}'_i$, include one measurement per minute.

Stage 3: Detection of Cyber-Attacks: Detecting the type of cyber-attack at each minute (if any) can be used to forecast future threats. Therefore, the RF algorithm is used to predict (i.e., detect) the type of cyber-attack at every minute given the features $\left(\mathbf{f}'_1, \mathbf{f}'_2, \ldots, \mathbf{f}'_k\right)$. However, one limitation of RF process is the unbalanced data. In order to tackle this limitation, the SMOTE (Chawla et al., 2002) oversampling method is applied. Thus, given the $\left(f'_{1,j}, f'_{2,j}, \ldots, f'_{k,j}\right)$ at minute j, the type of attack y_j^{dec} is predicted (detected), i.e.:

$$y_j^{dec} = RF\left(f'_{1,j}, f'_{2,j}, \ldots, f'_{k,j}\right). \tag{2}$$

Stage 4: Forecasting Models: In the proposed framework, three methods are applied to forecast whether there will be a cyber-attack in the next step (minute), as well as the type of attack. In the first approach, the LSTM method is applied to forecast the type of cyber-attack given the features in the m previous time steps, as well as the corresponding detected cyber-attacks, i.e.,

$$\hat{y}_j = LSTM\left(\mathbf{c}_{j-1}, \ldots, \mathbf{c}_{j-m}\right), \tag{3}$$

where $\mathbf{c}_j = \left(f'_{1,j}, \ldots, f'_{k,j}, y_j^{dec}\right)$ and $\hat{y}_j$ denotes the forecast label (type) of cyber-attack at minute j. Furthermore, the LR model for the multiple classification case and MLP are applied using only the last measurement, $\mathbf{c}_{j-1}$, i.e.,

$$\hat{y}_j = LR(\mathbf{c}_{j-1}), \tag{4}$$

$$\hat{y}_j = MLP(\mathbf{c}_{j-1}). \tag{5}$$

In Figure 1 a framework of the proposed approach for forecasting the type of cyber-attack (if any) in the next minute is illustrated.

4 EXPERIMENTAL EVALUATION

In this section, we conduct experimental evaluation on the framework presented in Section 3 using the three predictive models, LSTM, LR and MLP. Initially a detailed description of the used dataset, CIC-IDS2017, is provided, before the presentation of the experimental results.

4.1 Data Description

CIC-IDS2017 is a widely used publicly available dataset provided and generated by the Canadian Institute for Cybersecurity (CIC). It includes common up-to-date types of cyber-attacks such as DoS, DDoS, Web-Attack (WA), Infiltration, Port Scan (PS), Botnet etc., that meet real worlds criteria. There are few other more recent network traffic datasets, such as CSE-CIC-IDS2018 (Sharafaldin et al., 2018) and CIC-DDOS2019 (Sharafaldin et al., 2019), however, they do not include as many types of cyber-attacks as CIC-IDS2017.

The CIC-IDS2017 dataset capture took place over the course of 5 days, from Monday, July 3, 2017, at 9:00 a.m. through Friday, July 7, 2017, at 17:00. More specifically, the Victim-Network and the Attack-Network, which are two entirely different networks, were created in order to generate a realistic background traffic. Then, CIC employed the CICFlowMeter, a flow-based feature extractor that can extract 80 characteristics (or features) of network traffic from a pcap file. More details, as well as a statistic analysis for the CIC-IDS2017 dataset are provided in (Panigrahi and Borah, 2018).

4.2 Experimental Setup

The scope of this paper is to provide forecasts for the cyber-attacks to a specific DP; to that end, only the features relevant to this DP will be considered. In CIC-IDS2017, only the DP 80 and the DP 22 include different types of cyber-attacks. More precisely, DP 80 includes DDoS, DoS, WA, and PS attacks, while DP 22 includes SSH Patror (SSH-P) and PS attacks.

Next, the same features of network traffic are selected for both DPs, using the Spearman correlation values as described in stage 1 (Section 3). The resulting features (nine in total) are: *Flow Duration*, *Total Fwd Packets*, *Fwd Packet Length Max*, *Fwd Packet Length Min*, *Bwd Packet Length Max*, *Bwd Packet Length Min*, *Fwd Packets/s*, *Fwd Packets/s*, *Bwd Packets/s* and *Min Packet Length*.

In all records of the CIC-IDS2017 dataset, there is only one type of cyber-attack that occurs per minute. Therefore, the new features (derived during stage 2 in Section 3) are labeled with the cyber-attack indicator if there is at least one cyber-attack within the minute under consideration. Moreover, the task of forecasting the type of cyber-attacks can be approached as a multi-classification problem, since in every case (minute) there is one class (type of cyber-attack) occurring among the different types of cyber-attacks.

Next, the RF algorithm and the SMOTE oversampling method are applied to the derived features and labels, in order to predict (detect) the type of attack given the features (stage 3 in Section 3). The proposed cyber-attack detector for both DPs provides a very high accuracy in predicting the type of cyber-attack. The accuracy in DP 80 is 97.55%, while in the DP 22 is 98.2% (see Table 1) for balanced data.

Table 1: Detection of cyber-attack type.

	DP 80	**DP 22**
RF + SMOTE	97.55%	98.2%

Finally, each record of CIC-IDS2017 contains only one type of cyber-attack, however in order for the training dataset to contain cases from all types of cyber-attacks, we proceed as follow: (1) each record is split into training and testing dataset, such as both datasets to contain benign and non-benign labels, (2) all derived training datasets are concatenated (joined) into a single training dataset, and (3) we proceed in the same way for the testing datasets.

4.3 Experimental Results in DP 80

The performances of the three ML methods (LSTM, LR, MLP) in forecasting next minute cyber-attacks are evaluated by calculating the *weighted* Accuracy (Acc.), Precision (Prec.), Recall (Rec.), F1-score, since the dataset is unbalanced, as well as the confusion matrix (Hossin and Sulaiman, 2015). The parameters used in LSTM model are one LSTM layer with 50 units and one hidden layer with 5 units (as the number of classes), while for the MLP one hidden layer with 150 units is used. The Adam optimization process (Kingma and Ba, 2014b) is used for 20 epochs for both methods. The LR method uses the default parameters, which are defined in the Python library *sklearn* (Pedregosa et al., 2011).

Table 2 illustrates the results of the three methods for DP 80, where LSTM$_m$ stands for the LSTM method given the m previous features (see stage 4 in Section 3). From now on, only LSTM$_m$ methods, which provide the best performance are reported. It can be seen that all methods provide a good performance regarding all the four metrics. LSTM$_1$ provides the best performance over Prec. (94.23%) and F1-score (93.32%) metrics, while LR over the Acc. (93.49%) and the Rec. (93.49%) metrics.

Table 2: Performance of ML methods for DP 80.

	Acc.	**Prec.**	**Rec.**	**F1-score**
LSTM$_1$	93.33%	**94.23%**	93.33%	**93.32%**
LSTM$_2$	92.50%	93.80%	92.50%	93.06%
LR	**93.49%**	93.33%	**93.49%**	93.13%
MLP	93.17%	93.54%	93.17%	93.07%

However, as it can be seen in the confusion matrix $Conf^{80}_{LR}$, the LR method cannot forecast the PS attacks, since it misclassifies them as WA, while LSTM$_2$ ($Conf^{80}_{LSTM_2}$) provides more robust forecasting regarding the PS attacks. Finally, LSTM$_1$ ($Conf^{80}_{LSTM_1}$) and MLP ($Conf^{80}_{MLP}$) provides almost similar performances.

$$Conf^{80}_{LSTM_1} = \begin{array}{r|ccccc} & Benign & DoS & WA & PS & DDoS \\ \hline Benign & 540 & 6 & 13 & 1 & 1 \\ DoS & 6 & 6 & 0 & 0 & 0 \\ WA & 2 & 1 & 15 & 0 & 0 \\ PS & 5 & 0 & 5 & 2 & 0 \\ DDoS & 1 & 0 & 0 & 0 & 11 \end{array}$$

$$Conf^{80}_{LSTM_2} = \begin{array}{r|ccccc} & Benign & DoS & WA & PS & DDoS \\ \hline Benign & 534 & 14 & 2 & 10 & 0 \\ DoS & 6 & 6 & 0 & 0 & 0 \\ WA & 5 & 0 & 12 & 1 & 0 \\ PS & 5 & 0 & 1 & 6 & 0 \\ DDoS & 1 & 0 & 1 & 0 & 10 \end{array}$$

$$Conf^{80}_{LR} = \begin{array}{r|ccccc} & Benign & DoS & WA & PS & DDoS \\ \hline Benign & 544 & 2 & 13 & 1 & 1 \\ DoS & 6 & 6 & 0 & 0 & 0 \\ WA & 2 & 0 & 16 & 0 & 0 \\ PS & 5 & 0 & 7 & 0 & 0 \\ DDoS & 1 & 0 & 2 & 0 & 9 \end{array}$$

$$Conf^{80}_{MLP} = \begin{array}{r|ccccc} & Benign & DoS & WA & PS & DDoS \\ \hline Benign & 541 & 5 & 13 & 1 & 1 \\ DoS & 6 & 6 & 0 & 0 & 0 \\ WA & 2 & 1 & 14 & 1 & 0 \\ PS & 5 & 0 & 6 & 1 & 0 \\ DDoS & 1 & 0 & 0 & 0 & 11 \end{array}$$

4.4 Experimental Results in DP 22

The same metrics and methods' parameters as in the case of DP 80 are used in the case of DP 22. The only difference is the hidden layer in LSTM method, where the number of units are 3, since the number of classes is 3 (Benign, SSH-P, and PS).

The performances of the three ML methods are provided in Table 3, where all methods provide a good performance. However, LSTM$_1$ provides the best performance in all metrics, while LR the worst. Moreover, as it can be seen in the confusion matrices ($Conf^{22}_{LSTM_1}$) and ($Conf^{22}_{MLP}$), MLP has only one more false positive forecasting compare to LSMT$_1$.

It is clear that forecasting SSH-P attacks is easier than forecasting PS attacks. All methods correctly predict 22 SSH-P attacks, while they misclassify only one. Meanwhile, from the 5 in total PS attacks, LSTM$_1$ and MLP forecast correctly 3, while LR forecasts 2. Therefore, it can be concluded that for both DPs that the PS attacks are the hardest to forecast compare to other types of cyber-attacks.

Table 3: Performance of ML methods for DP 22.

	Acc.	**Prec.**	**Rec.**	**F1-score**
LSTM$_1$	**95.04%**	**95.43%**	**95.04%**	**95.21%**
LR	92.56%	93.63%	92.56%	93.07%
MLP	94.21%	94.60%	94.21%	94.38%

$$Conf^{22}_{LSTM_1} = \begin{array}{r|ccc} & Benign & SSH-P & PS \\ \hline Benign & 90 & 1 & 2 \\ SSH-P & 0 & 22 & 1 \\ PS & 2 & 0 & 3 \end{array}$$

$$Conf^{22}_{LR} = \begin{array}{r|ccc} & Benign & SSH-P & PS \\ \hline Benign & 88 & 1 & 4 \\ SSH-P & 0 & 22 & 1 \\ PS & 3 & 0 & 2 \end{array}$$

$$Conf^{22}_{MLP} = \begin{array}{r|ccc} & Benign & SSH-P & PS \\ \hline Benign & 89 & 1 & 3 \\ SSH-P & 0 & 22 & 1 \\ PS & 2 & 0 & 3 \end{array}$$

5 CONCLUSIONS AND FUTURE WORK

The aim of this paper was to provide next-minute forecasts about the type of cyber-attack to a specific DP. To that end, the network traffic flow was considered and, more precisely, the features of CIC-IDS2017 were taken into account. The CIC-IDS2017 dataset is mainly used in the literature for detecting the type of cyber-attacks. Therefore, there were many limitations that needed to be addressed to forecast cyber-attacks, such as the large number of different features, the non constant frequency of measurements within a predetermined time period, and the distribution of the cyber-attacks over time.

To tackle these limitations, only features with Spearman coefficient value of less than 95% were initially considered. Next, we set minute as the time-step and the new features were generated taking into account only the maximum value of the features within one minute. Then, in addition to the new features, the detected type of cyber-attacks were also used to forecast next minute's cyber-attacks. The RF algorithm was used to detect the type of cyber-attacks.

Finally, three ML methods (LSTM, MLP and LR) were utilised to provide forecasts for next minute's cyber-attacks. All methods performed well at both DPs that were considered (i.e., 80 and 22), regarding the four metrics, Acc., Prec., Rec. and F1-score. However, the LSTM method had the most robust performance being able to forecast all types of cyber-attacks.

As a step further, the proposed framework will be extended to forecast cyber-attacks in the next several minutes. Moreover, it would be interesting to extend the proposed framework in real-life datasets with cyber-attacks incidents, which would include more types of cyber-attacks occurring at the same time.

ACKNOWLEDGEMENTS

This research is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under AIDA (Grant Agreement No 883596).

REFERENCES

Ahmet Okutan, Gordon Werner, S. J. Y. and McConky, K. (2018). Forecasting cyber attacks with imbalanced data sets and different time granularities. *Cybersecurity*, 1:1–15.

Bakdash, J. Z., Hutchinson, S., Zaroukian, E. G., Marusich, L. R., Thirumuruganathan, S., Sample, C., Hoffman, B., and Das, G. (2018). Malware in the future? forecasting of analyst detection of cyber events. *Journal of Cybersecurity*, 4(1):tyy007.

Barreto, C. and Koutsoukos, X. (2019). Design of load forecast systems resilient against cyber-attacks. In *International Conference on Decision and Game Theory for Security*, pages 1–20. Springer.

Blowers, M. and Williams, J. (2014). Machine l earning applied to cyber operations. In *Network science and cybersecurity*, pages 155–175. Springer.

Brockwell, P. J. and Davis, R. A. (2016). Nonstationary and seasonal time series models. In *Introduction to Time Series and Forecasting*, pages 157–193. Springer.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357.

De Winter, J. C., Gosling, S. D., and Potter, J. (2016). Comparing the pearson and spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological methods*, 21(3):273.

Dreiseitl, S. and Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, 35(5-6):352–359.

Dutta, N., Jadav, N., Tanwar, S., Sarma, H. K. D., and Pricop, E. (2022). Intrusion detection systems fundamentals. In *Cyber Security: Issues and Current Trends*, pages 101–127. Springer.

Goyal, P., Hossain, K., Deb, A., Tavabi, N., Bartley, N., Abeliuk, A., Ferrara, E., and Lerman, K. (2018). Discovering signals from web sources to predict cyber attacks. *arXiv preprint*, 1:1–11.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Hossin, M. and Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2):1.

Ivanyo, Y. M., Krakovsky, Y. M., and Luzgin, A. N. (2018). Interval forecasting of cyber-attacks on industrial control systems. *IOP Conference Series: Materials Science and Engineering*, 327:1–6.

Ji, S.-Y., Jeong, B. K., Kamhoua, C., Leslie, N., and Jeong, D. H. (2022). Forecasting network events to estimate attack risk: Integration of wavelet transform and vector auto regression with exogenous variables. *Journal of Network and Computer Applications*, 203:103392.

Khan, A. S., Ahmad, Z., Abdullah, J., and Ahmad, F. (2021). A spectrogram image-based network anomaly detection system using deep convolutional neural network. *IEEE Access*, 9:87079–87093.

Kingma, D. and Ba, J. (2014a). Adam: A method for stochastic optimization. *arXiv preprint*, 1.

Kingma, D. P. and Ba, J. (2014b). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kwon, D., Kim, H., An, D., and Ju, H. (2017). Ddos attack volume forecasting using a statistical approach. In *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, pages 1083–1086. IEEE.

Lallie, H. S., Debattista, K., and Bal, J. (2020). A review of attack graph and attack tree visual syntax in cyber security. *Computer Science Review*, 35:100219.

Leetaru, K. and Schrodt, P. A. (2013). Gdelt: Global data on events, location, and tone, 1979–2012. In *ISA annual convention*, volume 2, pages 1–49. Citeseer.

Liakos, K. G., Georgakilas, G. K., Moustakidis, S., Sklavos, N., and Plessas, F. C. (2020). Conventional and machine learning approaches as countermeasures against hardware trojan attacks. *Microprocessors and Microsystems*, 79:103295.

Okutan, A., Jay Yang, S., McConky, K., and Werner, G. (2019). Capture: Cyberattack forecasting using non-stationary features with time lags. In *IEEE Conference on Communications and Network Security (CNS)*, pages 205–213. IEEE.

Panigrahi, R. and Borah, S. (2018). A detailed analysis of cicids2017 dataset for designing intrusion detection systems. *International Journal of Engineering & Technology*, 7(3.24):479–482.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Polatidis, N. and Georgiadis, C. K. (2016). A multi-level collaborative filtering method that improves recommendations. *Expert Systems with Applications*, 48:100–110.

Polatidis, N., Pimenidis, E., Pavlidis, M., Papastergiou, S., and Mouratidis, H. (2020). From product recommendation to cyber-attack prediction: Generating attack graphs and predicting future attacks. *Evolving Systems*, 11(3):479–490.

Qasaimeh, M., Abu Hammour, R., Yassein, M. B., Al-Qassas, R. S., Lara Torralbo, J. A., and Lizcano, D. (2022). Advanced security testing using a cyber-attack forecasting model: A case study of financial institutions. *Software: Evolution and Process*, 1:1–22.

Seabold, S. and Perktold, J. (2010). Econometric and statistical modeling with python. *in Proceedings of the 9th Python in Science Conference*, 57:1–16.

Sharafaldin, I., Lashkari, A. H., and Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSP*, 1:108–116.

Sharafaldin, I., Lashkari, A. H., Hakak, S., and Ghorbani, A. A. (2019). Developing realistic distributed denial of service (ddos) attack dataset and taxonomy. In *2019 International Carnahan Conference on Security Technology (ICCST)*, pages 1–8. IEEE.

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R. P., and Feuston, B. P. (2003). Random forest: a classification and regression tool for compound classification and qsar modeling. *Journal of chemical information and computer sciences*, 43(6):1947–1958.

Tavabi, N., Abeliuk, A., Mokhberian, N., Abramson, J., and Lerman, K. (2020). Challenges in forecasting malicious events from incomplete data. *Companion Proceedings of the Web Conference 2020*, 1:603–610.

Thapa, N., Liu, Z., Kc, D. B., Gokaraju, B., and Roy, K. (2020). Comparison of machine learning and deep learning models for network intrusion detection systems. *Future Internet*, 12(10):167.

Yang, H., Li, X., Qiang, W., Zhao, Y., Zhang, W., and Tang, C. (2021). A network traffic forecasting method based on sa optimized arima–bp neural network. *Computer Networks*, 193:108102.

Yang, S. J., Du, H., Holsopple, J., and Sudit, M. (2014). Attack projection. *Cyber Defense and Situational Awareness*, pages 239–261.