Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 9, August2021 :7099 - 7110

Research Article

Traffic Analysis Based Intrusion Dection System For Wireless Systems Using Gated RNN

Mr.K.Naresh Kumar Thapa[a], Sadula Akshitha[b], Subha Reddy P[c], Poluru Likitha[d],
Dr.T.Suresh[e], Karthik[f]

[a] Assistant Professor, Department of ECE, R.M.K Engineering College, R.S.M. Nagar, Kavaraipettai, Tamil Nadu, India
[b,c,d,e,f] UG Student, Department of ECE, R.M.K Engineering College, R.S.M. Nagar, Kavaraipettai, Tamil Nadu, India
Email: [1]knt.ece@rmkec.ac.in, [2]sadu17337.ec@rmkec.ac.in, [3]subh17418.ec@rmkec.ac.in, [4]polu17313.ec@rmkec.ac.in

Abstract

Traffic analysis plays most important role in validating the performance and protection of the whole network environment. As the congestion of network traffic increasing day by day, network traffic analysis need to be practiced periodically for ensuring and enhancing security. Various kinds of analysis methods are proposed. In recent machine learning techniques are employed for detecting intrusions, analysing the behaviours of malwares and discriminating the network traffic. Detecting anomalies in certain period of time is considered as drawback of ML bases analysis techniques and it increased complexity also. In this proposed work, Gated Recurrent Neural Network is incorporated in order to analyse the traffic of the network effectively with aim of zero vulnerability. It overcomes the issues of earlier existing methods and provides better result while comparing with them. In addition, this proposed approach provides better accuracy in term of classifying vulnerabilities in network traffic.

Keywords: *Traffic Analysis, Wireless Systems, RNN, Dection System*

Introduction

Malware traffic may be of any kind where the system functionality changes completely [1] [6]. Traffic is a very sensitive data that deals with a Quality of services like gaming, surfing and social media and other packet-based data [12].Malware is a malicious software, which infects the computer via network [18]. Modern malwares are propagating via networks are very stronger and not captured by present antivirus or anti-malware systems [16][20]. Hence analyzing the network traffic and system traffic is much important and needed as per the present security compliance [22]. Ransomware is variety of malware is designed to threaten money from victims, who are protect and blocked her data from anonymous. The most commonly known types of ransomwares are screen lockers and encryptors. Screen lockers, as the name implies block access to the system data with lock screen. Encryptors are type of ransomware which can encrypt the data and replicating themselves. Encyptors are used to threaten victim's data. The content cannot view without decryption key. Some of the ransomeware were created major impacts in the IT sectors and corporates. They are reveton ransomeware in 2012,cryptolocker in 2013, torrent locker in 2014, cryptowall in 2014 and wannacry attack in 2017 which is one of epic attacks in the history to stolen the user to encrypted one. For decryption it ask ransome to retrieve the user data it may not in recover form.

Present day research mainly focuses on Artificial Intelligence (AI) development is more popular [9]. In Artificial Intelligence, Poet-based and Deep Packet inspection (DPI)-based are rule-based approaches, which is mainly used to perform the traffic classification by matching pre-defined hard-coded rules. But in case, Statistical-based and behavioral-based methods are classic machine approaches, which is used to classify a traffic by extracting patterns from the original data set using a set of features. A classical Machine Learning approaches solves many major issues such as encrypted traffic classification and high computational cost and its faces many crucial challenges of designing features creating a problem in recent studies [7] [24].

Contributions in this paper

- We proposed a novel malicious traffic classification system using GRU model. The highlight of using GRU is its nature of stacking the LSTM models and concurrently executing the same.
- We also profile the traffic and characterize the features available into the useful form. A generalized packet processing adapter is also designed and developed for live testing
- We also emphasis to attempt the auto-tuning of the features while processing it in GRU. However, the system cannot auto-tune the entire features but few highly correlated features are auto-tuned which are achieved in the very first time in the traffic classification.
- We also validated the proposed system with the MTA-KDD'19 dataset to understand the efficacy of the system. From the experimental evaluation, it is also observed that the proposed system can work as same in both the collected data and the live data, which indeed produced an average accuracy of 98% for the both.
- Finally, the proposed system would pave the pathway for the security researchers who are focusing more towards the online DL system add automation.

Organization of the paper is as follows: Section II covers the state of the art literature. Section III explains the proposed malicious traffic classification system using GRU. Here, in this chapter it discusses about the GRU in the deeper aspect followed by how GRU systems are applied for malicious traffic classification. Section IV discusses about the results obtained during experimentation. This section also highlights the experimental setup used for experimentation. This section also emphasis on the performance of the proposed malicious traffic classification system with the other state of the art systems. Finally, the paper is concluded.

Literature survey

Rajendra Prasad et al. (2017) presented a review on MANET intrusion detection system with multitier energy system. The multitier energy system solves the intrusion detection issues like authentication, data integrity, and confidentiality. They discusses about the various wireless intrusion detection system and analyse the performance in terms of detection rate. It results in the designing challenges of the MANET IDS. Vasumathi et al. (2017) presented review on the anomaly-based intrusion detection system. They use different data set and algorithms to analyze the performance and security aspects of the anomaly-based traffic classification system. They use KDD data set to analyze different proportion of the system. It results the drawbacks of the previous IDS system and they describe the solution to overcome the problem. Masdariet al. (2020) proposed the optimization techniques and neural network-based traffic classification using the fuzzy min-max logic. Ranjan et al. (2015) They use KDD’99 data set, the system provides the adaption facility on online system and minimize the learning timing of the system. It results the improvement of system classification error and accuracy. They also improve the performance of the system.

Zhan Xin et al. 2019 proposed the IDS framework to identify the threads in WLAN networks. They acquainted the framework engineering with embrace the program/worker mode, and the framework shows the outcome to customer, it comprises of the customer and worker communications through the internet browsers. The general framework engineering is center around information stockpiling layer, information securing layer, result investigation layer, and recognition and examination layer. They likewise presented the square chain interruption discovery for safer and solid administrations. And furthermore, conquer the security challenges by utilizing the square chain interruption identification strategy, it results the more proficiently shield the framework from malevolent strings. Yi Aung et al. 2017 introduced the community interruption discovery dependent on k-implies procedure to improve the location exactness in interruption recognition framework, they utilize the information mining in half breed technique and single strategy. It results relatively diminished the time complicity of the framework between the single strategy and information mining in crossover technique. The creators likewise portray a technique called idea of projective versatile reverberation, which is utilized to especially decrease the framework model preparing time and keeping up the discovery exactness. This outcome the information mining calculation job in IDS entombs of time complexity.

Yihan Xiao *et al.* 2019 proposed the CNN based intrusion detection system. The authors use the KDD-CUP99 dataset to compare the performance of the IDS by using the CNN. It results the CNN based IDS model provide the higher detection rate and reduce the false negative rate of the system. Kumar & Sharam *et al*. 2018 proposed the IDS for signature and anomaly-based methods. The authors describe the intrusion detection in the cloud computing environment and hybrid IDS algorithm for improves the detection rate in the private cloud environment. Mohammed hasan ali *et al.* 2018 proposed the model for apple-based intrusion detection and validation. They use NSL-KDD data set. And compare this model with the Extreme Learning Machine (ELM) approach for IDS with the hybrid Particle Swarm Optimization (PSO) technique. It results the model PSO-ELM shows improve the accuracy in intrusion detection system. W. Hu *et al.* 2008 Proposed an ID algorithm using AdaBoost technique was used in decision stumps as weak classifiers. Their system performed better than other published results with a lower false alarm rate, a higher detection rate, and a computationally faster algorithm. However, the drawback is that it failed to adopt the incremental learning approach.

Md zahangir *et al.* 2017 Describe the cyber security serve issues in the cyber spaces. Author includes the neuromorphic cognitive computing method for IDS network in cyber security using the deep learning. They use the NSL-KDD data set with vector factorization approach. It result the increase accuracy in the classification in rage of 81.31% to 90.12%. Deep learning method reduces the human effect in the task, and improves the performance of the system. M. Safwan Mawlood Hussein *et al*. 2012 proposed the effectiveness of the hybrid ids with snort with native bayes network to improve the performance of the hybrid ids system. They used KDD cup 99 dataset for her research of intrusion detection. It results the average false alarm rate is improved and bayes networks gives the j48 graft response. Souparnika jayaprakash *et al*. 2018 proposed a system for data base intrusion detection using the octraplet and machine learning based on anomaly detection system. They create the architecture to implement the role based access control and implement new data structure is called octraplet, which is used to store the sql queries. This method is improving the performance of the system and detection rate.

Dimitar Nikolov *et al*. 2018 proposed the recurrent neural network classifier for network intrusion detection based on short- or long-term memory units. This approach is mainly focus on HTTP server-based intrusion detection. Marin E Pamukov *et al*. 2017 proposed the IoT IDS to improve the detection rate and performance of the detection system in IoT devices. The authors use multiple negative selection algorithms to reduce the errors in intrusion detection and it can runs without input operators. It results the device detect the intrusion with 90 % succession rate. Chung-ming *et al*. 2019 proposed a host-based IDS by using the machine learning, which is inspired by adoptable agent based artificial intelligence. This approach is detects the malicious attacks from the system call and protect the system from host based intrusion attacks. It also shows the exchange of packets between the computers by detection signals. Ved prakash Mishra *et al*. 2017 proposed the simulator system for IDS to detect the DDoS attacks and alarm about the attack to the administrator. It approaches uses the core of IDS and IPS to simulate the software in self-execution mode. And protect the system from traffic information. It results the system with increase the performance of the detection system with factors of accuracy, security. This approach used for education purpose because of some implementation difficulty in the real time network devices.

Methodology

Gated Recurrent Units (GRU)

Gated Recurrent Units (GRU) provides solution to the vanishing gradient problem and short-term memory problem. GRU is similar to LSTM with less parameters, so GRU is faster to train than LSTM. Using the internal gates GRU regulates the flow of information. GRU uses hidden states to transfer information instead of cell state. It contains only reset gate and update gate. These two gates can retain information for a long time.

Update Gate

They are similar to the forget gate and input gate of LSTM. Deletion or adding of an incoming information is decided by the update gate of GRU. The update gate helps in retaining of past information. If the model retains all the past information, the vanishing gradient problem will be eliminated.

The update gate for time t is computed as follows:

$$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1})$$

Table 1. Represents notation and its abbreviation

Notation	Abbreviation
z_t	Update gate vector
x_t	Input vector
$W^{(z)}$	Weight of the input vector
h_{t-1}	Output vector of previous state.
$U^{(z)}$	Weight of past output vector
σ	Sigmoid activation function
LSTM	Long Short Term Memory
GRU	Gated Recurrent Unit
RNN	Recurrent Neural Network
CNN	Convolutional Neural Network
MTA	Malicious Traffic Analysis

The result of update gate varies from 0 to 1.

Reset gate

Reset gate decide on the how much past information to be forgotten. Following is the formula to calculate the value of reset gate.

$$r_t = \sigma(W^{(r)}x_t + U^{(r)}h_{t-1})$$

wherer_t is reset gate vector, $W^{(r)}$ and $U^{(z)}$ is weight of input and past output vector.

Current memory content

A new memory content which will use the reset gate to store the past relevant information. It is calculated as follows

$$h_t = tanh(Wx_t + r_t \odot Uh_{t-1})$$

The Hadamard (element-wise) product is calculated between the reset gate r_t and Uh_{t-1}, which decides the information removal of previous state.

Final memory at current time step

Finally, theh_t vector is calculated, the information of current unit is passed down for update gate. The current memory content h'_tgives information to the output vector. The h_t is calculates as follows:

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h'_t$$

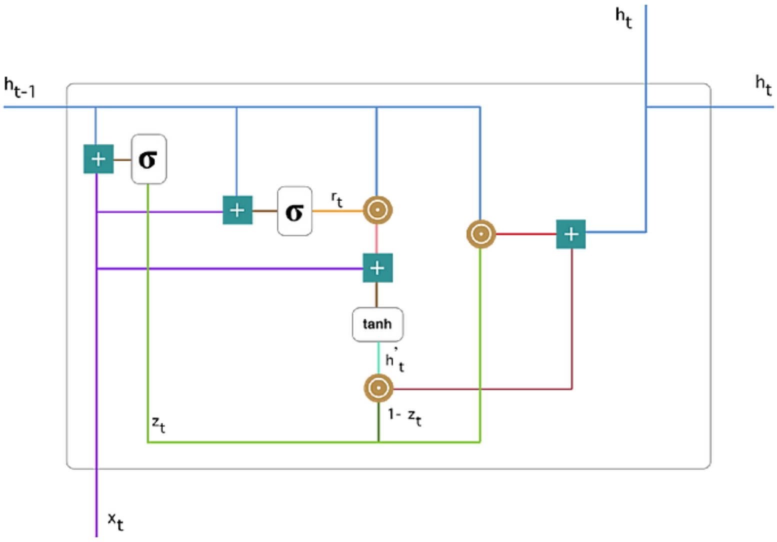

Figure 1: Gated Recurrent Unit

GRU can be used to efficiently predict network traffic. Prior prediction of huge volume of network traffic might help in a DDoS attack detection, Some of the influx traffic of HTTP,TCP and ICMP initiated by DDoS attack can be disguised as normal traffic. GRU works best in such scenarios.

Experimental Setup

Figure 2 shows the experimental setup used for the dataset collection. A single standalone computer with three independent VM is used to generate the traffic. Normal transactionsare performed for 24 hours to collect the normal data whereas for abnormal or malicious traffic generation the attacks as listed in the Table 1 are launched and all the traffic were tapped for a duration of 10 hours.

Legitimate traffic is captured and saved as different pcap files. Each pcap file size is around 700MB. The split up among pcap file is to avoid the bigger sized file, which may leads to loading and processing problem. Each pcap file is extracted for useful information with the self-written python script. The extracted information are processed and stored in the separate csv file with label for each record.

Legitimate traffic is captured and saved as different pcap files. Each pcap file size is around 400MB. The extracted information are processed and stored in the separate csv file with label for each record. The test bed hosting machine is protected with anti-virus and intrusion detection system (Suricata[25]). AV and IDS are deployed to flag the attack traffic as malicious. The present pcap file totally sized around 4.79 GB.

The conventional approach to process the data is to build a detection model which

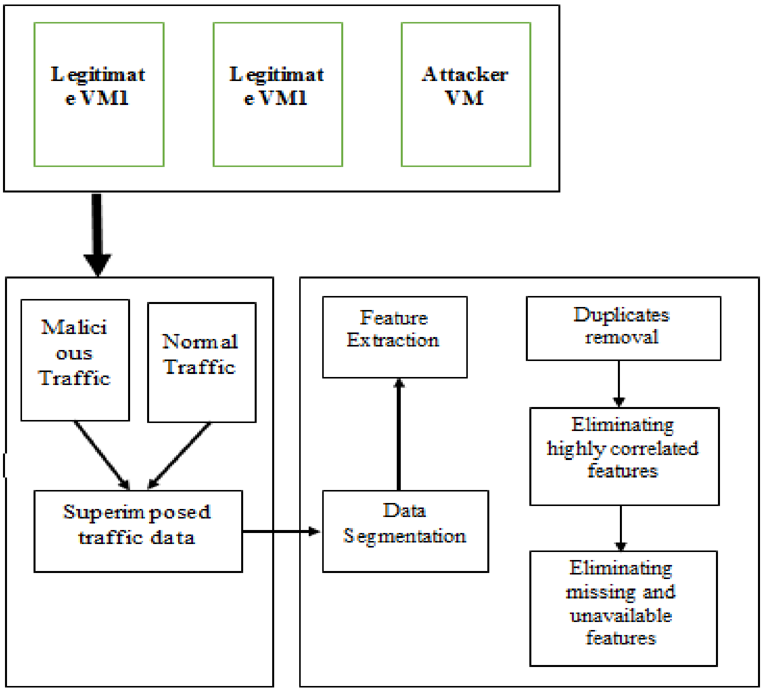

Figure2. Experimental setup used for dataset collection.

Dataset Description

The dataset used for experimentation includes the self-developed dataset (Refer Table 2) and MTA-KDD’19 (Refer Table 3).

Table2. Description of the dataset processed and validated in real time.

Category	Description	Size in GB
Benign	Normal traffic which includes traffic of torrent client such as utorrent, Office 365, Zoom, Teams, Watchdog, WhatsApp web	45 GB
Malicious	Malicious traffic includes attack traffic of BlackNurse, a self-developed DDoS attack tool, Various malicious malwares downloaded from virustotal	50 GB

Table 3. Description of MTA-KDD’19

Category	Description	Size in GB
Benign	Normal traffic which includes traffic features such as Number of connections, SynAcksynRatio, TCP flags {Ack,Syn,Fin,Psh,Urg,Rst}, IP features {TCP,UDP,DNS}, Other statistical features such as MaxLen, MinLen, AvgLen, StdDevLen, MaxIAT, MinIAT, AvgIAT, AvgDeltaTime, MaxLenRx, MinLenRx, AvgLenRx, StdDevLenRx, MaxIATRx, MinIATRx, AvgIATRx,StartFlow, EndFlow, DeltaTime, FlowLen, FlowLenRx, packet features such as packet information, packet input/output ratio, packet length, Repeated packet length ratio,	7 GB

	small and large packet length ratio, DNS features such as DNSQDist,DNSADist,DNSRDist,DNSSDist, URL features such as AvgDomainChar, AvgDomainDot, AvgDomainHyph, AvgDomainDigit, ValidUrlRatio, average time-to-live (TTL), Number of destination address, Number of ports, Distinct User Agent, Average Distinct User Agent Length, HTTP packets etc.	
Malicious	Abnormal traffic which includes traffic features such as Number of connections, SynAcksynRatio, TCP flags {Ack,Syn,Fin,Psh,Urg,Rst}, IP features {TCP,UDP,DNS}, Other statistical features such as MaxLen, MinLen, AvgLen, StdDevLen, MaxIAT, MinIAT, AvgIAT, AvgDeltaTime, MaxLenRx, MinLenRx, AvgLenRx, StdDevLenRx, MaxIATRx, MinIATRx, AvgIATRx,StartFlow, EndFlow, DeltaTime, FlowLen, FlowLenRx, packet features such as packet information, packet input/output ratio, packet length, Repeated packet length ratio, small and large packet length ratio, DNS features such as DNSQDist,DNSADist,DNSRDist,DNSSDist, URL features such as AvgDomainChar, AvgDomainDot, AvgDomainHyph, AvgDomainDigit, ValidUrlRatio, average time-to-live (TTL), Number of destination address, Number of ports, Distinct User Agent, Average Distinct User Agent Length, HTTP packets etc.	4.8 GB

Results and Discussion

In this section we highlights the performance of the proposed GRU based malicious traffic classification system with the existing state of the art deep learning models.The use of random forest for best feature selection plays a major role. Further, the use of GRU on the MTA KDD'19 dataset allows for a much lesser human guesswork required than other approaches such as mRMR or MIFS. The dataset while being optimized for malware analysis still has 33 features, most of which do not have high correlations with each other, this fact alone causes severe problems during training using methodologies such as a standard dense network or even while using a CNN based neural network. The inherent nature of Random Forest of "bagging" different decision trees to find the best feature possible allows us to skip the step of optimizing the dataset further such as the methods used by (Letteri et al., 2020 [26]).

Table 4. Comparison with MI Algorithm Ranking

	Accuracy	Precision	Recall
mRMR	97.07	97.09	97.07
MIFS	97.06	97.08	97.06
CIFE	88.46	88.70	86.41
JMI	84.35	85.25	84.35

CMIM	88.19	88.52	88.21
DISR	81.27	81.56	81.26
Random Forest + GRU	99.98	99.98	99.96

From Table 4, we can clearly see that while MI ranking algorithms such as the mRMR or MIFS applied on to the dataset provide for very high scores, they still do not match up with the robust approach of the random forest whilst finding the most relevant feature in the dataset. Beyond that, (Letteri et al., 2020 [26]) also applied the ranking while reducing the overall number of features from the dataset, see Figure 3, to make the process of finding the most relevant feature from the 33 features, this process can also be ignored while using Random Forests since they adds additional randomness to the model, while growing the trees. Instead of searching for the most important feature while splitting a node, it searches for the best feature among a random subset of features. This results in a wide diversity that generally results in a better model while still being able to operate on all the 33 features.

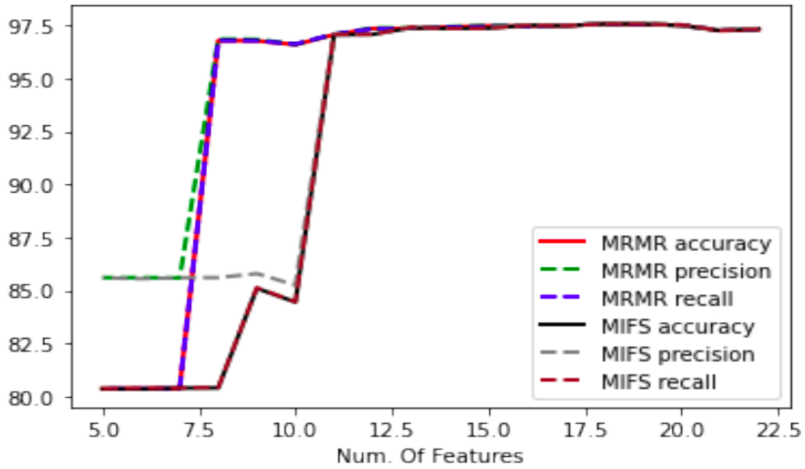

Figure 3. RF predictions while reducing the number of features using mRMR and MIFS

Looking at the ROC curve obtained, see Figure 4, by using Random Forests we can see that we achieve the best possible results between the true positive rate and false positive rate for a predictive model using different probability thresholds. As we can see clearly, the all the evaluation metrics have improved rather than fallen while using Random Forests even while applying no feature selection or dimension reduction algorithms. While optimization of the dataset allows for further advantages such as having a smaller model for online training, the time taken to train the Random Forest was negligible compared to strategies such as using Auto Encoders as experimented (Letteri et al., 2020 [26]). Thus the usage of our model should be applicable in most scenarios whilst still providing for a small and fast model that out performs current state of the art models.

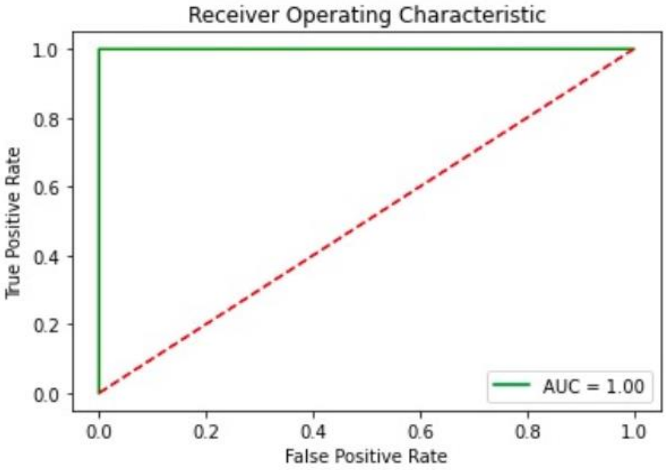

Figure 4. ROC curve plot for the proposed GRU based traffic classification system.

Conclusion

In this research, we presented a robust malicious traffic classification system. From the experimental results, it is clearly shown that the proposed method is effective in classifying the malicious traffic. Furthermore, the classification results exhibits that the proposed GRU base neural model can accurately classify the traffic with nearly 99% accuracy in overall with less than 1% False Positives and False negatives. The robustness of the proposed system is less packet flow inspection due to reduced and pre-processed dataset. The proposed system examines the flow by averaging the packets sum costing around 4 packets per flow and not more than 100 bytes from each packets. Hence, the classification time taken to detect malicious traffic is much reduced at the rate of 2.78% in increased efficiency. In future, the security researchers may focuses on the optimization methods used in the functional components of neural networks for building an effective online traffic classification system.

References

[1] P. P. Rajendra and Shivashankar, "Multitier energy system review on secure intrusion detection system in MANETs," 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2017, pp. 1722-1726, doi: 10.1109/RTEICT.2017.8256894.

[2] R. Samrin and D. Vasumathi, "Review on anomaly based network intrusion detection system," 2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2017, pp. 141-147, doi: 10.1109/ICEECCOT.2017.8284655.

[3] O. H. Ahmed, J. Lu, A. M. Ahmed, A. M. Rahmani, M. Hosseinzadeh and M. Masdari, "Scheduling of Scientific Workflows in Multi-Fog Environments Using Markov Models and a Hybrid Salp Swarm Algorithm," in IEEE Access, vol. 8, pp. 189404-189422, 2020, doi: 10.1109/ACCESS.2020.3031472.

[4] S. K. Sharma, D. Bhattacharyya, M. R. Patra and T. Kim, "A New Parallel Hybrid Model - Intrusion Prevention Systems," 2015 8th International Conference on Security Technology (SecTech), 2015, pp. 17-24, doi: 10.1109/SecTech.2015.17.

[5] X. Zhan, H. Yuan and X. Wang, "Research on Block Chain Network Intrusion Detection System," 2019 International Conference on Computer Network, Electronic and Automation (ICCNEA), 2019, pp. 191-196, doi: 10.1109/ICCNEA.2019.00045.

[6] Y. Y. Aung and M. M. Min, "A collaborative intrusion detection based on K-means and projective adaptive resonance theory," 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2017, pp. 1575-1579, doi: 10.1109/FSKD.2017.8393000.

[7] Y. Xiao, C. Xing, T. Zhang and Z. Zhao, "An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks," in IEEE Access, vol. 7, pp. 42210-42219, 2019, doi: 10.1109/ACCESS.2019.2904620.

[8] P. Kumar and V. Sharma, "Insecurity of a Secure Certificate-Based Signature Scheme," 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2018, pp. 371-373, doi: 10.1109/ICACCCN.2018.8748312.

[9] M. Z. Alom and T. M. Taha, "Network intrusion detection for cyber security on neuromorphic computing system," 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 3830-3837, doi: 10.1109/IJCNN.2017.7966339.

[10] M. H. Ali, M. Fadlizolkipi, A. Firdaus and N. Z. Khidzir, "A hybrid Particle swarm optimization -Extreme Learning Machine approach for Intrusion Detection System," 2018 IEEE Student Conference on Research and Development (SCOReD), 2018, pp. 1-4, doi: 10.1109/SCORED.2018.8711287.

[11] W. Hu, W. Hu and S. Maybank, "AdaBoost-Based Algorithm for Network Intrusion Detection," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 38, no. 2, pp. 577-583, April 2008, doi: 10.1109/TSMCB.2007.914695.

[12] S. Pan, T. Morris and U. Adhikari, "Developing a Hybrid Intrusion Detection System Using Data Mining for Power Systems," in IEEE Transactions on Smart Grid, vol. 6, no. 6, pp. 3104-3113, Nov. 2015, doi: 10.1109/TSG.2015.2409775.

[13] M. Z. Alom and T. M. Taha, "Network intrusion detection for cyber security on neuromorphic computing system," 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 3830-3837, doi: 10.1109/IJCNN.2017.7966339.

[14] S. M. Hussein, F. H. M. Ali and Z. Kasiran, "Evaluation effectiveness of hybrid IDS using Snort with Naïve Bayes to detect attacks," 2012 Second International Conference on Digital Information and Communication Technology and it's Applications (DICTAP), 2012, pp. 256-260, doi: 10.1109/DICTAP.2012.6215386.

[15] S. Jayaprakash and K. Kandasamy, "Database Intrusion Detection System Using Octraplet and Machine Learning," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018, pp. 1413-1416, doi: 10.1109/ICICCT.2018.8473029.

[16] D. Nikolov, I. Kordev and S. Stefanova, "Concept for network intrusion detection system based on recurrent neural network classifier," 2018 IEEE XXVII International Scientific Conference Electronics - ET, 2018, pp. 1-4, doi: 10.1109/ET.2018.8549584.

[17] M. E. Pamukov and V. K. Poulkov, "Multiple negative selection algorithm: Improving detection error rates in IoT intrusion detection systems," 2017 9th IEEE International

Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2017, pp. 543-547, doi: 10.1109/IDAACS.2017.8095140.

[18] V. P. Mishra and B. Shukla, "Development of simulator for intrusion detection system to detect and alarm the DDoS attacks," 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS), 2017, pp. 803-806, doi: 10.1109/ICTUS.2017.8286116.

[19] C. M. Ou, "Host-based Intrusion Detection Systems Inspired by Machine Learning of Agent-Based Artificial Immune Systems," 2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), 2019, pp. 1-5, doi: 10.1109/INISTA.2019.8778269.