IJIST
P ISSN : 2618-1630
E ISSN : 2909-6130

Enterprise Network Infrastructure Malicious Activity Analysis

Original Article

Muhammad Shujat Ali[1], Ahsan Abbas[4] , Abdullah Faisal[2], Anza Riaz[3], Imran Siddiq[*2]
[1] Orange Networks, Lahore, Pakistan
[2] Afro-Asian Institute Affiliated with Government College University Faisalabad, Pakistan
[3] Government College University Faisalabad Layyah Campus, Pakistan
[4] Telenor bank, Lahore, Pakisatn
* **Correspondence**: Imran Siddiq[2],imrancpe4u@gmail.com

Citation | Shujat.A.M, Abbas, Abdullah Faisal.A, Riaz.A, Siddiq.I "Enterprise Network Infrastructure Malicious Activity Analysis". International Journal of Innovations in Science & Technology, Vol 04 Issue 04: pp 982-997, 2022.
DOI | https://doi.org/ 10.33411/IJIST/2022040404
Received |Sep 22, 2022; **Revised** | Oct 15, 2022; **Accepted** | Oct 21, 2022; **Published** | Oct 26, 2022.

Inter and intra-network connectivity for different organizations have become a useful resource for accessibility and flexibility for data connectivity many apps, and online services are increasing day by day, and because everything is available online, it generates a huge amount of data, which is why cyber security revolves around it. Because of an increase in internet use and cyber-attacks, companies such as business continuity failure, credibility, and other aspects are frequent and affect them. In the continued development of the threat environment, cyber security teams must now deal with numerous threats. As multiple attacks on computer networks and systems are becoming more harmful, current security tools are often inadequate to resolve issues relating to unauthorized users, reliability, and reliable network security. To maintain a safe environment and Intrusion-Detection Mechanisms (IDS) that control device functions and detect intrusions should typically be used to supplement other protection strategies; conventional security methods are not adequate. Actual users expect their requested information to be processed in real-time, while malicious traffic needs to be mitigated just as quickly as possible. As traffic increases, this problem becomes more complex. This paper contributes a detailed analysis of network packets to find anomaly detection based on the UNSW NB 15 dataset and investigate the the difference between IP packet behavior for both malicious and legitimate packets. Besides we acquaint with new methodologies to illuminate and appraise the network attack in a very proficient way using different machine learning algorithms which will accomplish locating the malicious traffic in the least execution time with precision.

Keywords: Cyber Security, Network Attack, Intrusion-Detection Mechanisms (IDS), Multi-Layer Perceptron (MLP), Distributed Denial-of-Service (DDoS).

Acknowledgment.
This declaration clarifies that all Authors have seen and approved the paper that has been submitted. We assure you that the article is the original work of the Authors. We ensure that the article has not been previously published and is not being considered for publication Anywhere.
Project details. NIL
Author's Contribution.
All authors contributed significantly to the study.

Conflict of interest
The authors declare no conflict of interest in publishing this manuscript in IJIST.

Introduction

Currently, communication networks are described by the assortment of recently deployed services and network features. Network administrators should accordingly have the option to continuously detect, analyze and troubleshoot connectivity issues before they cause a loss of connectivity or other distraction in network services [1]. Initially, anomaly detection has henceforth become an acute problem for network administrators and network service providers in their obligation to keep a contracted[2] Service Level Agreement (SLA). Network administrators are typically liable for handling and managing the connectivity from various network devices like bridges, switches hubs, and routers for keeping up at acceptable levels[3]. Identifying, analyzing and troubleshooting network faults are major areas in network[4] management[5]. Evaluation of network flow from various network devices is therefore mandatory to maintain the connectivity and to assure of guaranteed flow of traffic[6]. A fault management system is shown in Figure 1.

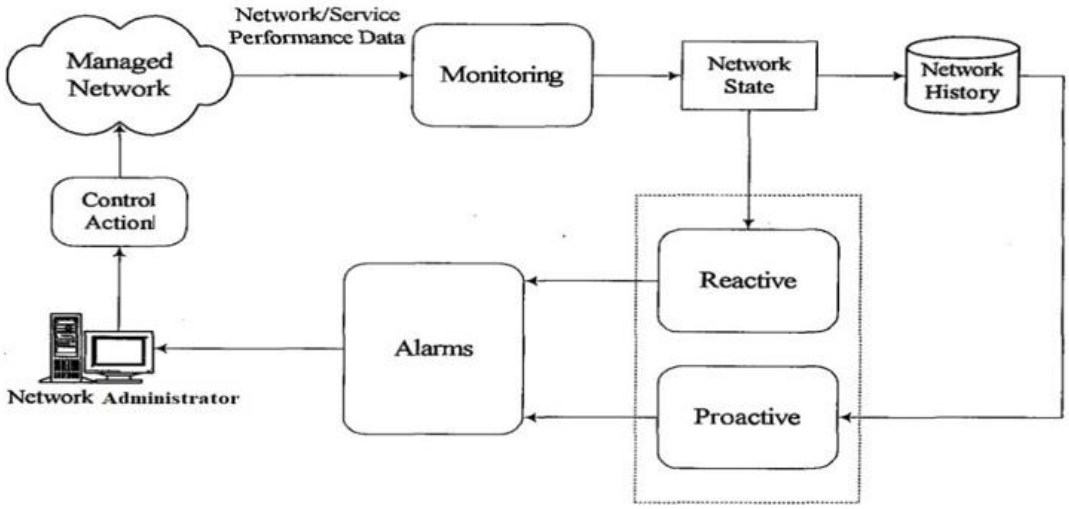


Figure 1: Fault Management System

About 145.5 million United States customers and some 44 million British citizens and a major part of Canadian citizens were affected by the Equifax cybersecurity identity theft event[7]. In initial exchange rates, the next day of the cyber-attack, Equifax shares dropped 13 percent and several proceedings against Equifax emerged as an outcome of the violation[8]. Equifax was suffering from brand damage[9]. In July 2019, Equifax accepted a $300 million deal with the FTC that includes the Victim Compensation Fund, the Convention's State and Territories of $175 million, and the penalties of $100 million[10]. Kaspersky attack analysis is shown in Figure 2.

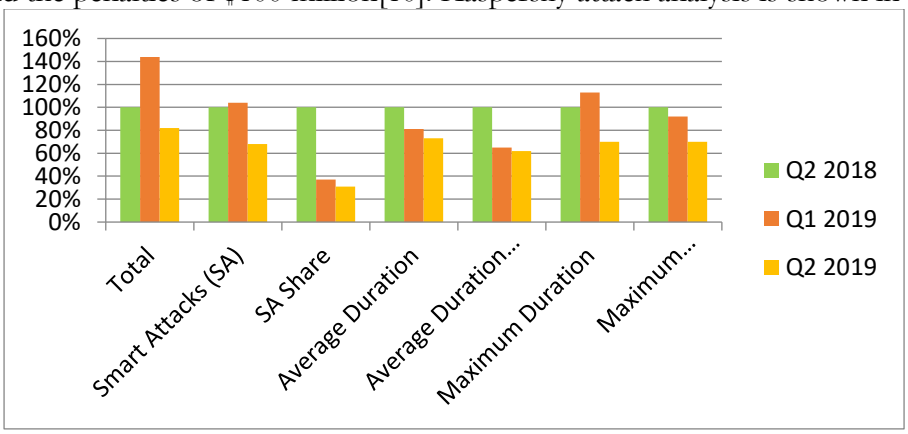


Figure 2: Kaspersky's DDOS Attacks Analysis

According to Kaspersky's research, the overall number of DDoS attacks increased by 18% as compared to the same span in 2018 in the second quarter of 2019[11]. The more difficult to

coordinate and defend from application-layer attacks saw robust growth of 32 percent compared to Q2 in 2018. It now accounts for almost (46%) of all Kaspersky DDoS security prevention assaults. Botnets are ready to launch DDoS attacks that can disrupt critical services or serve as a resource diversion screen for another operation on the network and both. DDoS attacks can trigger significant service disruption. Social manipulation and spear phishing are also used to carry a bomber into an opponent's system. Georgian Leader Mikheil Saakashvili's website was struck by a DDoS attack in July 2008

Threats, Attacks, and Vulnerabilities

Threats to the protection of information may include attacks by malware, theft of intellectual property, identity theft, robbery of equipment or information, sabotage, and extortion of information. A danger can be something that can circumvent an infringement of security and adversely impact, delete, and destroy artifacts or items of interest. ttack software includes Virus attacks, Worm attacks, Trojan horses, etc. Many users think malware, viruses, worms, and bots are all similar things. But they're not similar, just that they all have different habits of malicious software. In corporate ventures, a business network plays a critical role to perform and move forward, the upgrading of network devices with the latest technologies and security techniques is necessary. Trustworthiness and security are essential concerns for a business to expand and develop. A centralized security strategy no longer suffices for the evolving and indefinable nature of security today, as people are dependent on mobile devices, home computers, and laptops to link to company and business networks. Endpoint Protection combines centralized security controls with extra protection for certain threats at the access level for sensitive data. By having endpoints on devices to fulfill security criteria before granting network access, businesses may maintain greater control over the ever-increasing number and effectively block access threats and efforts before entering. While managing access, the endpoint often has features like monitoring for dangerous or abnormal behaviors and blocking them. Endpoint working topology is shown in Figure 3.

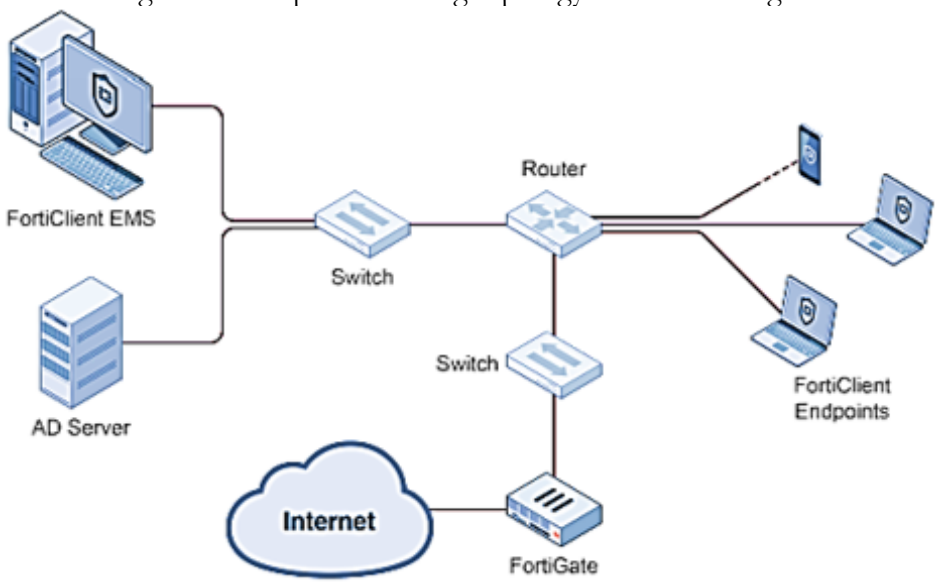


Figure 3: Endpoint Working Topology (Fortinet)

The parallel identical mechanism correlates the payload of a file with the DB. If any packet meets Snort policies, the ASIC hardware is in an idle state and sends a signal to the module for confirmation that combines signals to take decisions and decides if there is an irregular payload Here in the policy engine, the authors built a half mesh architecture that permits the network traffic to be equated with a different set of rules. Figure 4 shows the Parallel path Processing Workflow Topology.

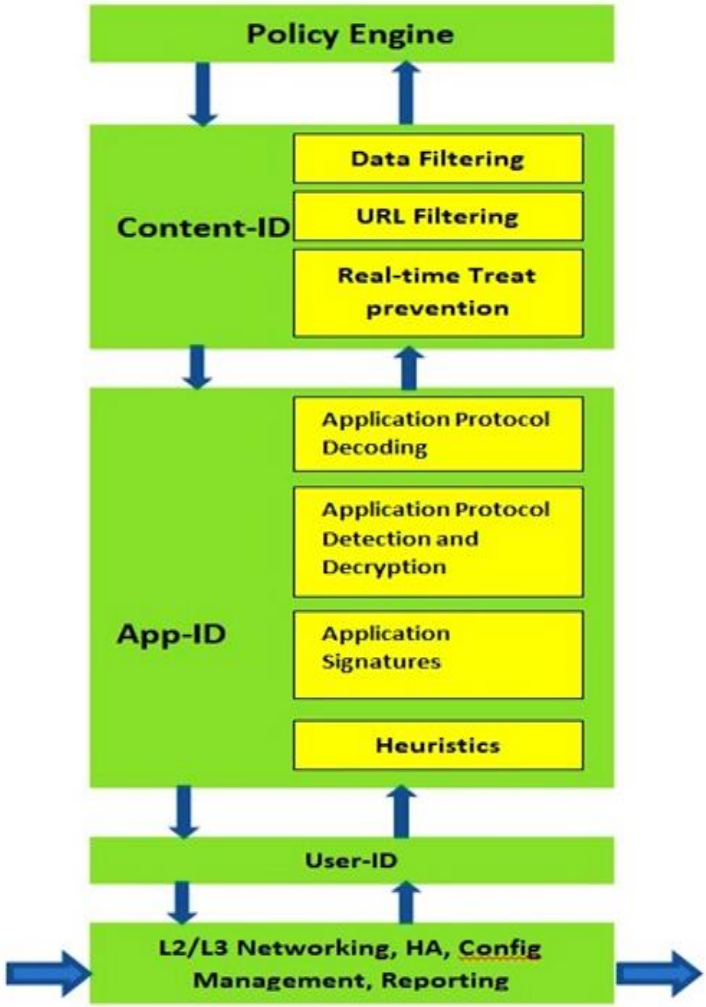


Figure 4: Parallel path Processing Workflow Topology

LITERATURE REVIEW

Recent events in Central America, in which botnets have repeatedly made down the national infrastructure of one country, underline the severity of this hazard. A one million bots botnet can handle 128 gigabits of traffic at a moderate 128kBps broadband upload per compromised machine. A study implemented two algorithms: the decision trees and multilayer neural network for intrusion strategies grounded on anomalies. NN displayed a capacity for the precision of generalization but was weak in sensing new threats. Decision trees have hewed good effectiveness in the generalization and detection of new threats. A dataset that is used is KDD, derived from the DARPA dataset. This dataset consists of 24 types of attacks which will be assembled into four kinds, name DOS, R2L, U2R, and Probe. Additionally, they apply their methodology from their laboratory to a campus network, which includes unknown for DARPA. Although the outcomes obtained are compared to previous studies, the identification performance of a variety of types of attacks remains low. In the study, introduced a methodology that generates a collection of rules for classifying network traffic as natural or invasive and specifies the type of breach to be defined. A technique for dimension reduction, PCA, is recognized as Karhunen Loève transforms, and it was used to describe a subgroup of attributes that preserve important details. The KDD dataset was used for changing rules. PCA cut the 41 attributes in the dataset to just 3 initially. Then it generates a set of rules using a genetic algorithm to identify actions as normal or abnormal. Every intrusion detection rule is a variation of the binary from (0, 1). The provisional portion of the instruction consists of the attribute related to the AND function. The significance

of rues is a confirmation of a classification of the intrusion. This methodology can provide higher detection before the execution of the real-time network and low false-positive. Nonetheless, they found three forms of breaches that are not tested by this method.

Ambusaidi generates an IDS using the SVM algorithm combined with the function selection algorithm: Versatile Shared Knowledge-Based attribute chosen. This study aims to remove unnecessary and irrelevant features from data since these attributes decelerate the classification procedure and also disturb the accuracy of cataloging. The study presented that the attribute chosen technique contributes to attributes that are important to achieving better results for LSSVM-IDS and reduces computational costs. Najafabadi tried ML techniques to create forecasting from malicious networks to differentiate against regular traffic. The proposed architecture requires a method of data reduction (i.e. collection of features) to eliminate unnecessary and redundant features that contribute to a reduction in process time. Along with four separate feature selection techniques, they assessed three categories. These models were trained and tested by using Dataset Kyoto 2006. The outcomes of the differentiation showed that while the collection of attributes decreases the number of attributes, it preserves the previous value or does not reduce results. They accomplish that while choosing attributes in the anomaly detection areas is a significant preprocessing stage and it will not be entertained casually.

Xiang [12] introduced an IDS hybrid multilevel classifier. This method merged an unmonitored apprentice with managed learning. The dataset used for this analysis is KDD. This model is categorized into three instances such as DoS, R2L, and Other categories at the first classification level. At this point, all other protocols are categorized as others such as R2L and U2R. In the second phase, they use Bayesian clustering to organize U2R and R2L as attacks and start occurrences by using 4 attributes out of 41. Such features were chosen by measuring different parameters. Using 14 functions only, this methodology differentiates others in stage 3. This move is performed much easier because at stage 2 the Benign instances were separated. Finally, this model is used additionally categorize all attack classes into their disparities. The outcomes after this experiment presented an efficient discovery frequency with a low false-negative, compared with other common methods, and retained a fair false alarm rate level at the same time. Unknown attacks however are absent from the training.

In his study [13] they suggested a solution containing three classifiers (Decision Trees, SVM, and SVM mixture method. The anomaly detection methodology is developed by using KDD for the training. For this experiment, for categorization, every trained model received corresponding details about the experiential behaviors [14]. The Ensemble classifier's final result is determined by the base model's highest score. - Model's score is calculated through assigned weights that reflect the output of the individual prediction on the training dataset [15]. Therefore, the ensemble classifier uses a model score for a confident instance that will be classified, and if all representations give different views. The highest score model is pondered by the champion and it is responsible for the result. This approach increases outcomes by using gaps in misclassification [16] .

METHODOLOGY

Besides, computer-based crime has evolved and is becoming increasingly dangerous in many aspects of life as technology has advanced over the last few decades. Crime behavior, in particular, has become a major issue because of financial benefits. Such an operation exploits the computer system of a victim to obtain sensitive personal data to be used for business fraud or identity theft or reputation loss for a well-known brand or any Government level institution. Also, these offenders use computer tools to initiate DDoS attacks. The main objective of cybercriminals is to produce a large-scale attack with little effort on standardized software platforms. The growth of cybercrime technologies has intensified these side effects. A botnet is a series of computers that have been infected. The botnet has different features than previous malware. Malicious activity produces a momentous threat to the existing interconnected networks. Anomaly detection will be

analyzed by using the TTL values of TCP, UDP, and OSPF features of network traffic flow, and machine learning algorithms are a common approach to identifying malicious events.
The main objective of this research is to perform an analysis of network traffic flow, to identify the legitimate and malicious traffic that will cause spamming or disruption of legitimate services in real-time enterprise networking environments such as universities, healthcare infrastructures, Financial Institutions, and other organizations, etc.

Research Contribution

On the basis of analysis of normal & abnormal network traffic flow (explained under Analyze of Normal & Abnormal Network Traffic Flow Using the deployed protocols in Data set) using the deployed protocols data set employing our proposed algorithm that is based on Decision Tree, quite suitable for the underlying algorithm, we are confident that it will enhance the accuracy and reducing the overall processing time of the whole process. In view of this fact, the following objectives in respect of our research are being presented but not limited:

I. to enhance the success ratio and decrease the overall execution time
II. to compare the working of each algo with the proposed algorithm
III. to evaluate the result of the proposed algorithm with the results of other researchers for their employed algorithms using the same data set

IV.Algorithms used:

a. Random Forest Classifier (Information Gain)
b. Decision Tree (Proposed algorithm)
c. Gaussian Naive Bayes
d. Gradient Boosting Classifier
e. Multi-Layer Perceptron (MLP) Neural Network Classifier

V.Compile results to check

a. Accuracy of each algorithm
b. Comparison of each algorithm

VI.Data Set

a. The data set is collected from UNBC Australia
b. It includes different types of traffic captures, for example, protocol,ots spots, pets, et state, details, sjit, it, etc)
c. This dataset consists of 45 attributes (features), which includes (Id, protocol, dst, src, src byte, DST byte, etc.). The output variable is ttl.
d. Is consists of 82,333 records.
e. TTL value at source & destination is less than 64

Analyze Normal & Abnormal Network Traffic Flow using the deployed protocols in Data Set Based on the decision tree algorithm network traffic flow analyzed the value of source & destination TTL of TCP, UDP, and OSPF protocol.
Normal & Abnormal Packet statistics:

a- Protocol: UDP
Service: DNS
State: CONN

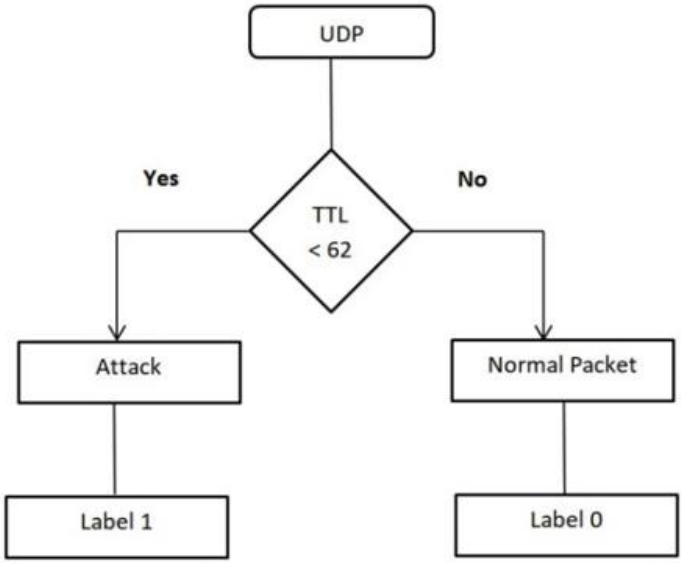


Figure 5: The process of detecting legitimate and malicious traffic

Normal State:

i. If the TTL value at the source & destination is less than 64 then the packet would be considered normally shown in Table 1

ii. The minimum time duration of a normal packet resides in the cache

iii. Ttl value range is between the same sequence flow

Table 1: TTL Values

Sr	Id	Purdue	proto	service	state	sttl	data	attack_cat	label
1	23379	0.00111	UDP UDP	DNS DNS	CON	31	29	Normal	0
2	23413	0.001017	UDP	DNS DNS	CON	31	29	Normal	0
3	23668	0.000987	UDP UDP	DNS	CON	31	29	Normal	0
4	23852	0.001051	UDP	DNS	CON	31	29	Normal	0
5	23872	0.00113	UDP	DNS	CON	31	29	Normal	0

Abnormal State:

i. If the TTL value at the source & destination is more than 64 then the packet would be considered abnormal (Malicious Packet) shown in Table 2

ii. The maximum time duration of abnormal packets reside in the cache

iii. Ttl value range is between the same sequence flow

Table 2: Malicious Packet

Sr	Id	Purdue	proto	service	state	sttl	data	attack_cat	label
1	1041	0.008666	UDP	Dns	CON	62	252	DoS	1
2	6518	0.04725	UDP	Dns	CON	254	60	DoS	1
3	6938	0.217753	UDP	Dns	CON	62	252	DoS	1
4	10087	0.074168	UDP	Dns	CON	254	60	DoS	1
5	13686	0.009394	UDP	DNS	CON	62	252	DoS	1

b- Protocol: TCP

Service: HTTP

State: CONN

Given Below Flow chart highlights the process of detecting legitimate and malicious traffic of TCP protocol whereas Table 3 & Table 4 shows features respectively

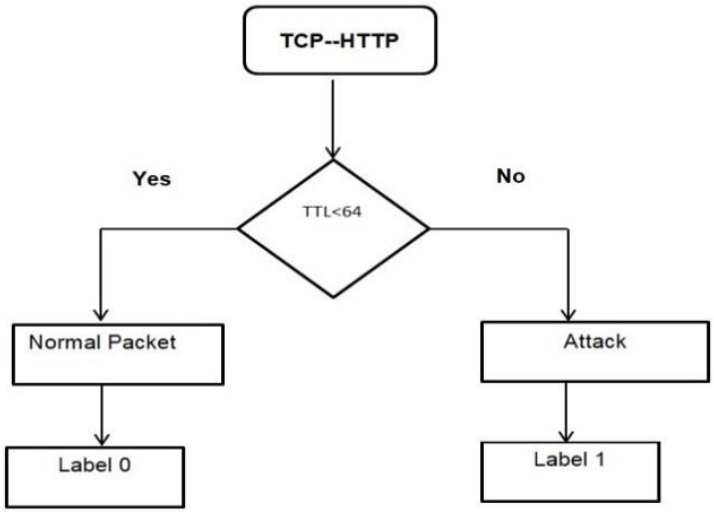


Figure 6: Detecting legitimate and malicious traffic

Normal State:

i. If the TTL value at the source & Destination is less than 64 then the packet would be considered normally shown in Table 3

ii. The maximum time duration of an abnormal packet sresiding in the cache

iii. Ttl value range is between the same sequence flow

Table 3: TTL value at source & Destination is less than 64

Sr	Id	Due	proto	Service	state	sttl	Titel	attack_cat	Label
1	25570	1.226977	TCP	HTTP	CON	32	30	Normal	0
2	31341	2.166709	TCP	HTTP	CON	31	30	Normal	0
3	31478	1.034461	TCP	HTTP	CON	31	29	Normal	0
4	31485	0.005322	TCP	HTTP	CON	31	29	Normal	0
5	32347	0.004343	TCP	HTTP	CON	31	29	Normal	0

Abnormal State:

i. If the TTL value at the source & destination is more than 64 then the packet would be considered abnormal (Malicious Packet) shown in Table 4

ii. The maximum time duration of an abnormal packet sresiding in the cache

iii. Ttl value range is between the same sequence flow

Table 4: TTL value at source & destination is more than 64

Sr	Id	Dur	proto	Service	state	sttl	dttl	attack_cat	Label
1	10155	59.96682	tcp	http	CON	62	252	DoS	1
2	59530	0.726271	tcp	http	CON	62	252	DoS	1

c- Protocol: OSPF

Service:

State: REQ

Given Below Flow chart highlight the process of detecting legitimate and malicious traffic of OSPF whereas Table 5 & Table 6 shows features respectively

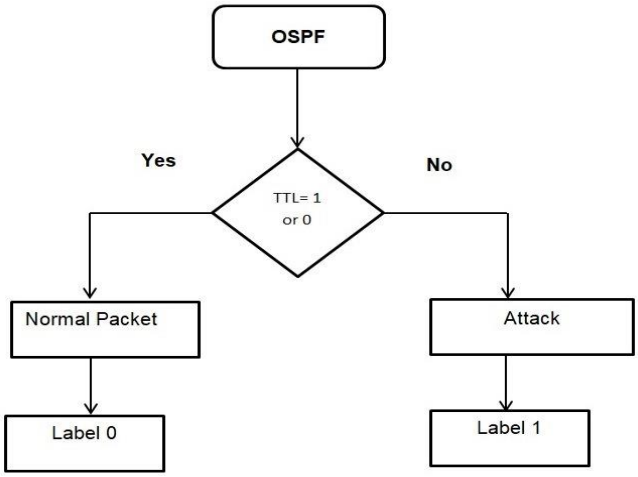


Figure 7: detecting legitimate and malicious traffic of OSPF

Normal State:

i. If the TTL value of the source & destination is between 1 & 0 respectively then the packet would be considered normal
ii. The minimum time duration of a normal packet resides in the the cache
iii. Ttl value range is between 0 and 1

Table 5: TTL value of source & destination between 1 & 0

Sr	Id	Dur	proto	service	state	sttl	dttl	attack_cat	label
1	23458	50.00436	ospf	-	REQ	1	0	Normal	0
2	27927	50.00435	ospf	-	REQ	1	0	Normal	0
3	27934	50.00435	ospf	-	REQ	1	0	Normal	0
4	27949	50.00435	ospf	-	REQ	1	0	Normal	0

Abnormal State:

i. If the TTL value of source & destination is more than 1 & 0 respectively then the packet would be consider as abnormal (Malicious Packet). Shown in Figure 8
ii. Maximum time duration of abnormal packet reside in cache
iii. Ttl value is also much higher than normal packet range

Table 6: TTL value of source & destination is more than 1 & 0

Sr	Id	Dur	proto	Service	state	sttl	dttl	attack_cat	label
1	9706	59.40567	ospf	-	REQ	254	0	DoS	1
2	9842	59.40567	ospf	-	REQ	254	0	DoS	1
3	9916	58.42214	ospf	-	REQ	254	0	DoS	1
4	9967	59.74488	ospf	-	REQ	254	0	DoS	1

Proposed Model

We proposed a decision tree algorithm to overcome the accuracy and execution issues.

a- What are Decision Trees?

Decision trees are highly versatile predictive models that allow items to be rapidly classified, grouped, or valued against a range of parameters. They are also a great way to simply visualize a decision involving multiple factors.

b- How does decision Tree work?

It's a collection of 'if/then' rules which will be applied to the branches of etrees. Based on data input 'if/then' rules move downward to end up at a leaf which is an outcome or reasonable prediction of a target variable.

c- Functioning of a Decision Tree:

A decision Tree is a decision-making tool that uses a flowchart-like tree structure or is a model of decisions and all of their possible results, including outcomes, input costs, and utility. This algorithm falls under the category of supervised learning algorithms. It works for both continuous as well as categorical output variables.

The branches/edges represent the result of the node and the nodes have either:

- Conditions [Decision Nodes]
- Result [End Nodes]

The branches/edges represent the truth/falsity of the statement and take makes a decision based on that shown in the figure 8 which shows a decision tree that evaluates the mallets of three numbers:

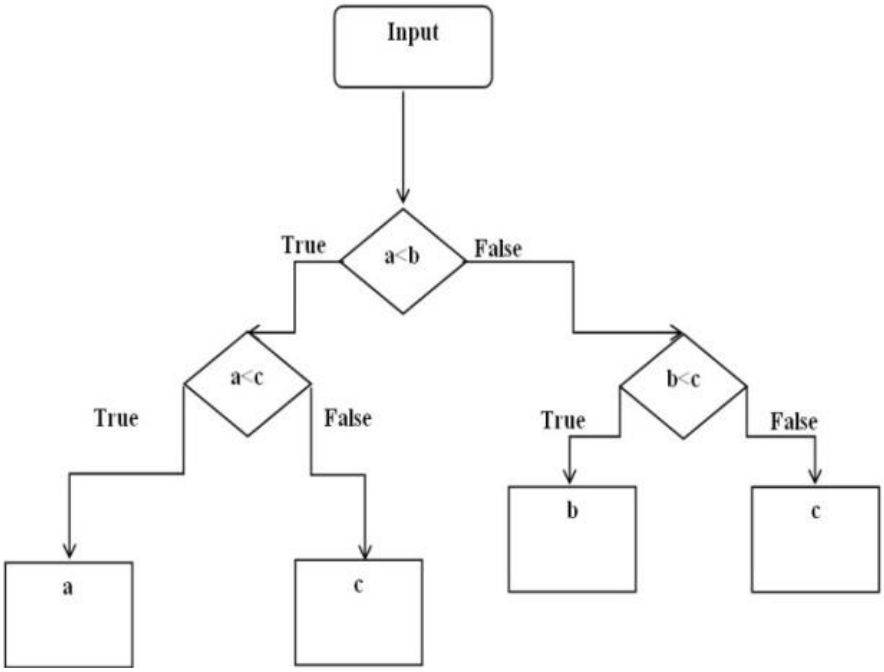


Figure 8: Decision Trees

Decision trees are made up of three elements: nodes, branches, and leaves.

- **Root Nodes:**
  Contain the overarching question or the decision your tree is seeking to answer.
- **Branches:**
  Indicated by an arrow line, represent options, criteria, or courses of action
- **Leaf Nodes:**
  It appears at the end of branches to represent either a further question to be asked (decision node), outcome (end node), or probability/uncertainty of an outcome (chance/change node)

Factors for Classification of Normal and Abnormal Traffic

In this study, we examined different factors that affect network speed and cause congestion during peak time and disrupt some critical services. In real environments at large-scale networks such as enterprises, universities, health care systems, or corporate levels this disruption will cause intrusion in legitimate services like web Services or other services.

Below are some factors which should be considered while the classification of legitimate and non-legitimate traffic.

- Source & Destination Byte
- Flag Type and Service
- Destination Byte
- Source & Destination of TTL

- Source & Destination Port
- Protocol (OSI Layers)

Proposed Model

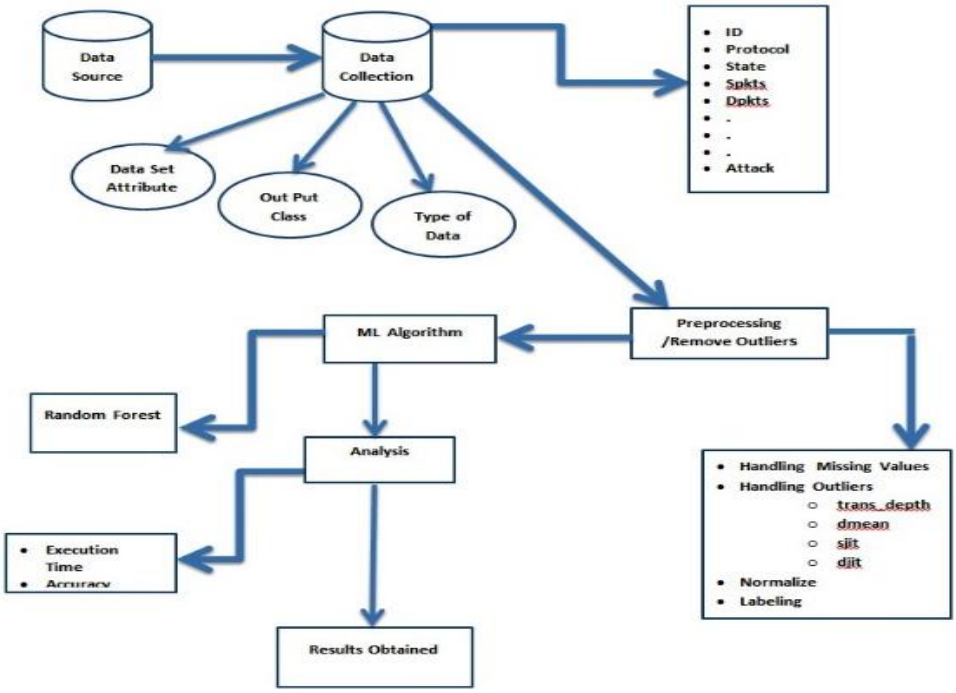


Figure 9: Malicious Traffic Detection Model

Proposed Model Explanation

The research methodology includes the review of existing research in adopting cloud-based solutions in national as well as international universities, analyzing current models and methods, technical and high-tech solutions and proposing an optimum solution. Microsoft Solutions Framework will be used as research methodology throughout the entire research process to design and propose a clouds-based hybrid model.

The basic objective of this proposed model is to make sure the availability and enhance the services required by the users from end to end within a university network. Microsoft Solutions Framework (MSF) is a set of principles, disciplines, concepts, and guidelines for delivering information technology solutions from Microsoft. No doubt, it doesn't provide the solution only to the problem significantly it ensures that if followed in the prescribed manner, the required goal can be achieved as the model being proposed in this proposal. Each phase of the research methodology is associated with the other phase.

Experiments and Results

The simulation tool MATLAB is used for the evaluation of the proposed model. Dataset is used to predict the results of using different machine-learning approaches. We have applied the following approach and got the results follows

Random Forest Classifier (Information Gain)

In this approach, the aforementioned machine learning algorithm shows the accuracy, accuracy of CV, and execution time in Table 7 whereas the Figure 10 shows the graphical representation.

Table 7: Accuracy of CV and Execution Time (Random Forest Classifier)

Accuracy	Accuracy of CV	Execution Time
99.74	92.6	460.2896

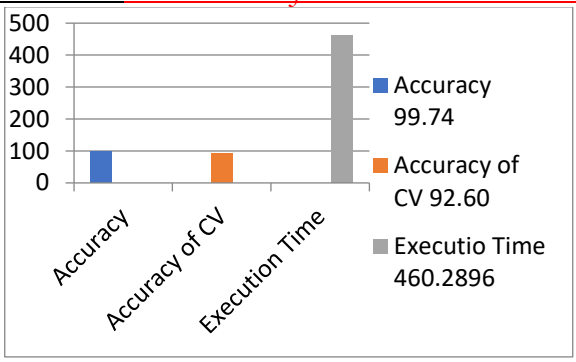


Figure 10: Random Forest Algorithms

Decision Tree Classifier

In this approach, the a for mentioned machine learning algorithm shows the accuracy, accuracy of CV and execution time in the Table 8 whereas the Figure 11 shows the graphical representation

Table 8: Accuracy of CV and Execution Time (Decision Tree Classifier)

Accuracy	**Accuracy of CV**	**Execution Time**
99.74	91.44	70.1151

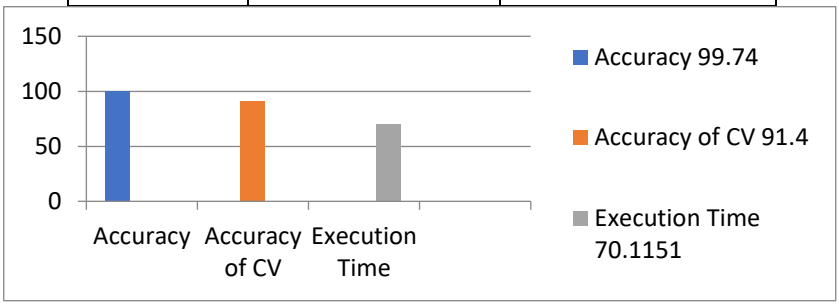


Figure 11: Decision Tree Algorithms

Gaussian Naive Bayes

In this approach, the a for ementioned machine learning algorithm shows the accuracy, accuracy o,f CV, and execution time in the Table 9 whethe reas the Figure 12 shows the graphical representation

Table 9: Accuracy of CV and Execution Time (Gaussian Naive Bayes)

Accuracy	Accuracy of CV	Execution Time
50.50	50.46	17.1392

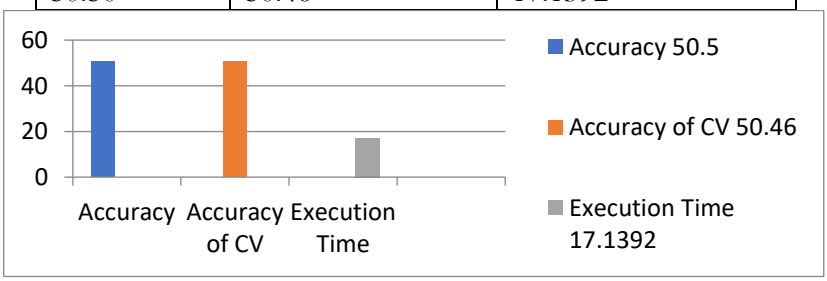


Figure 12: Gaussian Naïve Bayes Algorithm

Gradient Boosting Classifier

In this approach, a for mentioned machine learning algorithm shows the accuracy, accuracy of CV and execution time in the Table 10 whereas the Figure 13 shows the graphical representation

Table 10: Accuracy of CV and Execution Time (Gradient Boosting Classifier)

Accuracy	Accuracy of CV	Execution Time
93.38	92.09	1432.8694

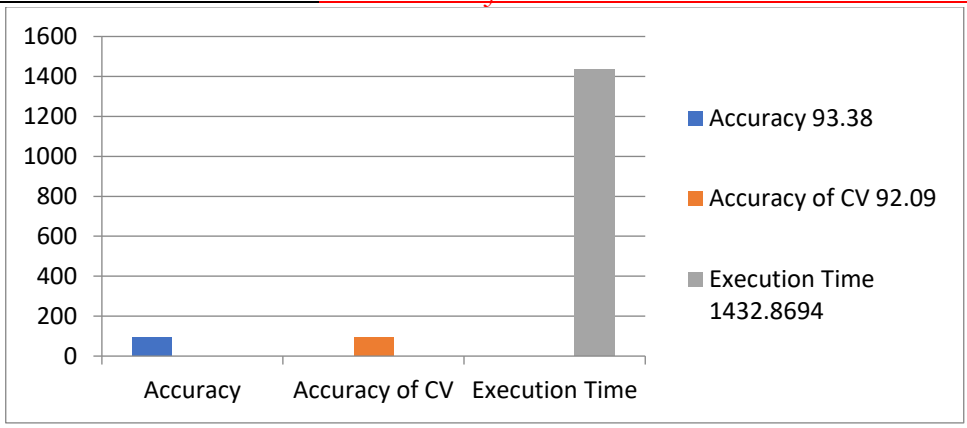


Figure 13: Gradient Boosting Algorithm

Multi-Layer Perceptron (MLP) Neural Network Classifier

In this approach, a for mentioned machine learning algorithm shows the accuracy, accuracy of CV and execution time in the Table 11 whereas the Figure 14 shows the graphical representation

Table 11: Accuracy of CV and Execution Time
(Multi-Layer Perceptron (MLP) Neural Network Classifier

Accuracy	Accuracy of CV	Execution Time
93.67	91.13	524.8119

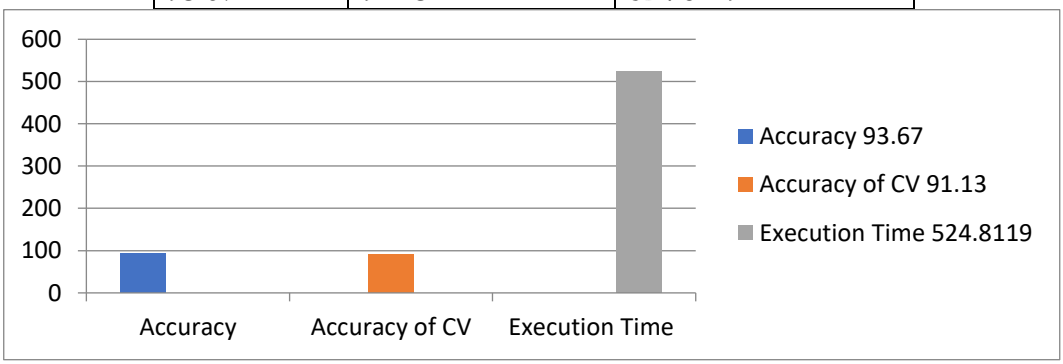
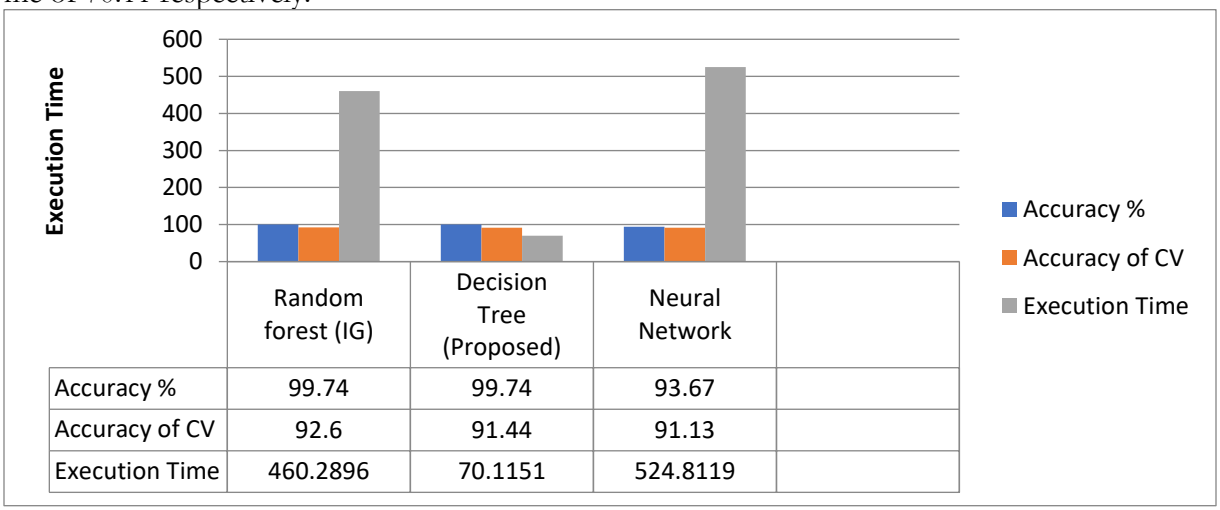


Figure 14: Neural Network Classifier

Comparison of Algorithm Accuracy

The experimental result has clearly shown that the accuracy of the Random Forest and D ecision Tree approach of machine learning attained the same accuracy. But the decision tree algor ithm shows the best performance of attaining good accuracy of 99.74% in minimum execution ti me of 70.11 respectively.



	Random forest (IG)	Decision Tree (Proposed)	Neural Network	
Accuracy %	99.74	99.74	93.67	
Accuracy of CV	92.6	91.44	91.13	
Execution Time	460.2896	70.1151	524.8119	

Figure 15: Execution Time

Table 12: Accuracy & Execution Time Comparison

Comparison Between Algorithms				
Sr.	Algorithms	Accuracy (%)	Execution Time	Missing Rate
1	Random Forest IG	99.74	460.28	0.26
2	Decision Tree	99.74	70.11	0.26
3	Gaussian Naive Bayes GNB	50.50	17.13	49.50
4	Gradient Boosting Classifier GBC	93.38	1432.86	6.62
5	Neural Network	93.67	524.811	6.33

Algorithms Accuracy Comparative Analysis

A comparative analysis is shown in Figure 16 between different algorithms and it has bee n clearly shown that the decision tree approach of machine learning is better than other algorithm s used for comparison respectively.

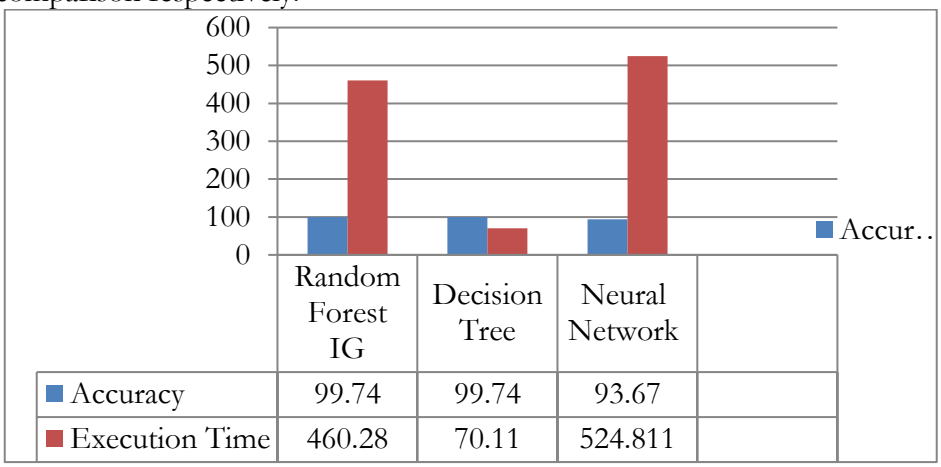


Figure 16: Algorithms Accuracy Comparative Analysis

Comparative Analysis of Accuracy with Other Researchers

DNN gains 75.75% accuracy while Nova shield Malware detection gains 75% accuracy through the decision tree best among them. The comparative analyses are shown in Figure 17.

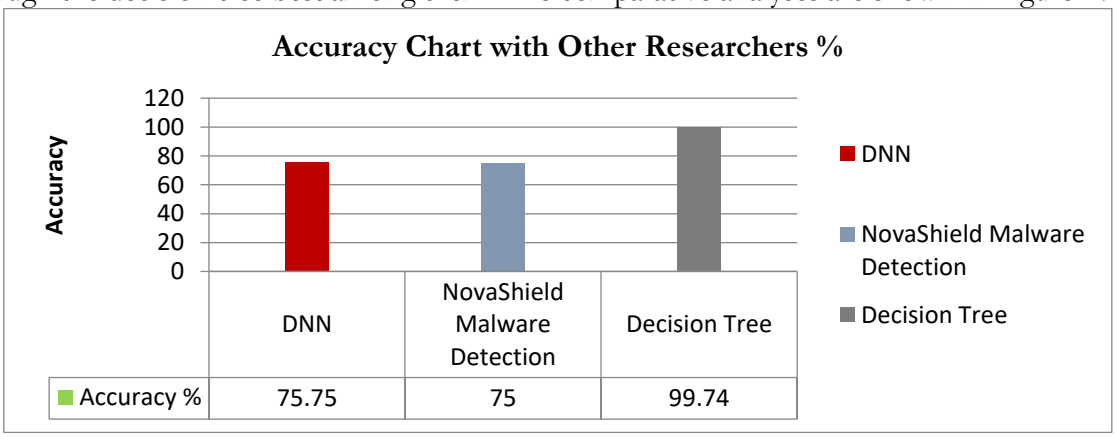


Figure 17: Comparative Analysis of Algorithms

CONCLUSIONS

The basic purpose of the proposed research how would a better accuracy rate achieve in network traffic by using various techniques of machine learning. The proposed model consisted of a machine-intelligent prediction system for the detection of malicious traffic in networks with ensemble different approaches adopted to improve efficiency and accuracy. The implementation of the proposed model can increase productivity and efficiency in the detection of malicious traffic in the network.

In this research methodology, different approaches of machine learning are used to find the malicious traffic route in the network. The result shows that the decision tree performs well, it gained an accuracy of 91.44 in 70.11 execution time which is the minimum time of execution. Whereas the random forest and gradient boosting algorithm achieved accuracy of 92.60 % and 92.09% which shows higher accuracy but on the other hand it takes too much time for execution 460.28 and 1432.86 respectively.

Future Work

The work stated in this research thesis gives a promising way for solving anomaly detection and prediction in communication networks; nonetheless, much work ruins to be done around there.

- The conventional distance metrics in the neighbor-based methods can't work well for the high-dimensional data in view of the equidistant attributes. The mixed-type features make irregularity discovery more troublesome. Presenting effective distance metrics for the high-dimensional and mixed-type data is essential.
- The neighbor-based malicious discovery algorithms are subtle to the nearest neighbors' selection for the models. Decisive the right number of neighbors is a difficult issue for the neighbor-based methods.

References

[1] Saravanan A, Bama SS. A review on cyber security and the fifth generation cyberattacks. Oriental Journal of Computer Science and Technology. 2019;12(2):50-6.

[2] Lobastova S. Geopolitics of Cyberspace and Virtual Power. Journal of Liberal Arts and Humanities. 2020(3):97-113.

[3] Celedonia KL, Valenti MW, Corrales Compagnucci M, Lowery Wilson M. Community-based health care providers as research participant recruitment gatekeepers: ethical and legal issues in a real-world case example. Research Ethics. 2021 Apr;17(2):242-50.

[4] Maurya RK, Bruce MA, Therthani S. Counselors' perceptions of distance counseling: A national survey. Journal of Asia Pacific Counseling. 2020 Aug 1;10(2):1-22. https://psycnet.apa.org/record/2021-31046-001

[5] Pattanaik SS. SAARC COVID-19 fund: calibrating a regional response to the pandemic. Strategic Analysis. 2020 May 3;44(3):241-52.

[6] Cisco U. Cisco annual internet report (2018–2023) white paper. Cisco: San Jose, CA, USA. 2020 Mar. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[7] Kaartinen H, Pieskä S, Vähäsöyrinki J. Digital manufacturing toolbox for supporting the manufacturing SMEs. In2016 7th IEEE International Conference on Cognitive Info communications (CogInfoCom) 2016 Oct 16 (pp. 000071-000076). IEEE.

[8] Malwarebytes 2020 State of Malware Report [Online]. https://www.malwarebytes.com/blog/news/2020/02/malwarebytes-labs-releases-2020-state-of-malware-report

[9] iDefense Security Intelligence Services. "Cyber-threats against 2018 PyeongChang Winter Olympics." February 7, 2019. IntelGraph reporting; iDefense Security Intelligence Services. "Secure Olympics Tokyo 2020: Is Japan Prepared for the Games?" April 29, 2019. IntelGraph reporting.

[10] iDefense Security Intelligence Services. "Technical Analysis of HWP-based Malware Targeting Current Events." June 21, 2018. IntelGraph reporting; iDefense Security Intelligence Services. "Hacktivist Activity for Sept. 1-8, 2016." September 9, 2016. IntelGraph reporting; iDefense Security Intelligence Services. "Phishing Attack Targeting Tibetan Organizations uses the 2014 G20 Summit to Deliver MNkit and Lurk Malware." November 13, 2014. IntelGraph reporting

[11] iDefense Security Intelligence Services. "SNAKEMACKEREL Campaign Likely Targeting NATO Members, Defense, and Military Outlets." December 21, 2018. IntelGraph reporting.
[12] iDefense Security Intelligence Services. "Overview of Recent Ransomware Activity." March 29, 2019. IntelGraph reporting. https://www.coursehero.com/file/pr2clq/98-IntelGraph-reporting-97-ibid-98-iDefense-Security-Intelligence-Services/
[13] Minnaar A. Cybercriminals, cyber-extortion, online blackmailers and the growth of ransomware. Acta Criminologica: African Journal of Criminology & Victimology. 2019 Aug 1;32(2):105.
[14] Badawi E, Jourdan GV. Cryptocurrencies emerging threats and defensive mechanisms: A systematic literature review. IEEE Access. 2020 Oct 29;8:200021-37.
[15] Price G. Cisco annual internet report (2018–2023) white paper. Cisco, San Jose, CA, USA, Tech. Rep. 2020.
[16] Grass E, Pagel C, Crowe S, Ghafur S. A Stochastic Optimisation Model to Support Cybersecurity within the UK National Health Service. Available at SSRN 4042065. 2022 Feb

Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.