Prediction of Cyber Attacks using Data Science Technique

Deepika A, Fiza K , Harani K , Pooja B

Mr.Manimaran (Assistant Professor/CSE)

8th semester,Department of Computer Science and Engineering

Dhaanish Ahmed College of Engineering,Chennai

Abstract-Cyber-attack, via cyberspace, targeting an enterprise's use of cyberspace for the purpose of disrupting, disabling, destroying, or maliciously controlling a computing environment/infrastructure; or destroying the integrity of the data or stealing controlled information. The state of cyberspace portends uncertainty for the future Internet and its accelerated number of users. New paradigms add more concerns with big data collected through device sensors divulging large amounts of information, which can be used for targeted attacks. Though a plethora of extant approaches, models and algorithms have provided the basis for cyber-attack predictions, there is the need to consider new models and algorithms, which are based on data representations other than task-specific techniques. However, its nonlinear information processing architecture can be adapted towards learning the different data representations of network traffic to classify type of network attack. In this paper, we model cyber-attack prediction as a classification problem, Networking sectors have to predict the type of Network attack from a given dataset using machine learning techniques. The analysis of dataset by supervised machine learning technique(SMLT) to capture several information's like, variable identification, uni-variate analysis, bi-variate and multivariate analysis, missing value treatments etc. A comparative study between machine learning algorithms had been carried out in order to determine which algorithm is the most accurate in predicting the type of cyber Attacks. We classify four types of attacks are DOS Attack, R2L Attack, U2R Attack, Probe attack. The results show that the effectiveness of the proposed machine learning algorithm technique can be compared with best accuracy with entropy calculation, precision, Recall, F1 Score, Sensitivity, Specificity and Entropy.

1.INTRODUCTION

Supervised Machine Learning is the majority of practical machine learning using supervised learning. Supervised learning is where you have input variables (X) and an output variable (y) and use an algorithm to learn the mapping function from the input to the output is y = f(X). The goal is to approximate the mapping function so well that when you have new input data (X) that you can predict the output variables (y) for that data. Techniques of Supervised Machine Learning algorithms include logistic regression, multi-class classification, Decision Trees and support vector machines etc. Supervised learning requires that the data used to train the algorithm is already labeled with correct answers. Supervised learning problems can be further grouped into Classification problems. This problem has as goal the construction of a succinct model that can predict the value of the dependent attribute from the attribute variables. The difference between the two tasks is the fact that the dependent attribute is numerical or categorical for classification. A classification model attempts to draw some conclusion from observed values. Given one or more inputs a classification model will try to predict the value of one or more outcomes. A classification problem is when the output variable is a category, such as "red" or "blue".

Title : A Prediction Model of DoS Attacks Distribution Discrete Probability

Authors: Wentao Zhao, Jianping Yin and Jun Long

The process of prediction analysis is a process of using some method or technology to explore or

stimulate some unknown, undiscovered or complicated intermediate processes based on previous and present states and then speculate the results. In an early warning system, accurate prediction of DoS attacks is the prime aim in the network offence and defense task. Detection based on abnormality is effective to detect DoS attacks. Various studies focused on DoS attacks from different respects. However, these methods required a priori knowledge being a necessity and were difficult to discriminate between normal burst traffics and flux of DoS attacks. Moreover, they also required a large number of history records and cannot make the prediction for such attacks efficiently. Based on data from flux inspecting and intrusion detection, it proposed a prediction model of DOS attack's distribution discrete probability based on clustering method of genetic algorithm and Bayesian method and the clustering problem first, and then utilizes the genetic algorithm to implement the optimization of clustering methods. Based on the optimized clustering on the sample data, we get various categories of the relation between traffic and attack amounts, and then build up several prediction sub-models about DoS attack. Furthermore, according to the Bayesian method, deduce discrete probability calculation about each sub-model and then get the distribution discrete probability prediction model for DoS attack. This paper begins with the relation between network traffic data and the amount of DoS attack, and then proposes a clustering method based on the genetic optimization algorithm to implement the classification of DoS attack data. This method first gets the proper partition of the relation between the network traffic and the amount of DoS attack based on the optimized clustering and builds the prediction sub-models of DoS attack. Meanwhile, with the Bayesian method, the calculation of the output probability corresponding to each sub-model is deduced and then the distribution of the amount of DoS attack in some range in future is obtained.

Title : Apriori Viterbi Model for Prior Detection of Socio-Technical Attacks in a Social Network

Authors : Preetish Ranjan, Abhishek Vaish

Socio-technical attack is an organized approach which is defined by the interaction among people through maltreatment of technology with some malicious intent to attack the social structure based on trust and faith. Awful advertisements over the internet and mobile phones may defame a person, organization, group and brand value in society which may prove to be fatal. People are always very sensitive towards their religion therefore mass spread of manipulated information against their religious belief may create pandemonium in the society and can be one of the reasons for social riots, political imbalance etc. Cyber-attack on water, electricity, finance, healthcare, food and transportation system are may create chaos in society within few minutes and may prove even more destructive than that of a bomb as it does not attack physically but it attacks on the faith and trust which is the basic pillar of our social structure. Trust is a belief that the person who is being trusted will do what is being expected for and it starts from the family which grows to build a society. Trust for information may be established if it either comes from a genuine source or information is validated by an authentic body so that there is always a feeling of security and optimism. In the huge and complex social network formed using cyberspace or telecommunication technology, the identification or prediction of any kind of socio-technical attack is always difficult. This challenge creates an opportunity to explore different methodologies, concepts and algorithms used to identify these kinds of communities on the basis of certain patterns, properties, structure and trend in their linkage. It tries to find the hidden information in huge social network by compressing it in small networks through apriori algorithm and then diagnosed using viterbi algorithm to predict the most probable pattern of conversation to be followed in the network and if this pattern matches with the existing pattern of criminals, terrorists and hijackers then it may be helpful to generate some kind of alert before crime.

Due to the emergence of the internet on mobile phones, the different social networks such as on social networking sites, blogs, opinion, ratings, review, serial bookmarking, social news, media sharing, Wikipedia led the people to disperse any kind of information very easily.

Rigorous analysis of these patterns can reveal some very undisclosed and important information explicitly whether that person is conducting malignant or harmless communications with a particular user and may be a reason for any kind of socio-technical attacks. From the above simulation done on CDR, it may be concluded that if this kind of simulation applied on networks based on the internet and if we are in the position to get the data which could be transformed in transition and emission matrix then several kind of prediction may be drawn which will be helpful to take our decisions.

Title : New Attack Scenario Prediction Methodology
Author : Seraj Fayyad, Cristoph Meinel

Intrusion detection systems (IDS) are used to detect the occurrence of malicious activities against IT systems. Through monitoring and analyzing IT system activities the malicious activities will be detected. In the ideal case IDS generates alert(s) for each detected malicious activity and stores it in the IDS database. Some of the stored alerts in the IDS database are related. Alerts relations are differentiated from duplication relation to same attack scenario relation. Duplication relation means that the two alerts generated as a result of the same malicious activity. Where the same attack scenario relation means that the two related alerts are generated as a result of related malicious activities. Attack scenario or multi-step attack is a set of related malicious activities run by the same attacker to reach a specific goal. Normal relation between malicious activities belonging to the same attack scenario is causal relation. Causal relation means that current malicious activity output is pre-condition to run the next malicious activity. Possible multi-step attacks against a network start with information gathering about the network and the information gathering is done through network Reconnaissance and fingerprinting. Through reconnaissance network configuration and running services are identified. Through the fingerprint process Operating system type and version are identified. propose a real time prediction methodology for predicting most possible attack steps and attack scenarios. Proposed methodology benefits from attack history against the network and from attack graph source data. it comes without considerable computation overload such as checking of attack plans library. It provides parallel prediction for parallel attack scenarios. Possible third attack step is to identify the attack plan based on the modeled attack graph in the past step. The attack plan usually will include the exploitation of a sequence of found vulnerabilities. Mostly this sequence is distributed over a set of network nodes. This sequence of node vulnerabilities is related through causal relation and connectivity. Lastly Attacker start orderly exploits the attack scenario sequences till reaching his/her goal. Attack plan consisting of many correlated malicious activities ends up with an attacking goal.

Title : Cyber Attacks Prediction Model Based on Bayesian Network
Author: Jinyu W1, Lihua Yin and Yunchuan Guo

The prediction results reflect the security situation of the target network in the future, and security administrators can take corresponding measures to enhance network security according to the results. To quantitatively predict the possible attack of the network in the future, attack probability plays a significant role. It can be used to indicate the possibility of invasion by intruders. As an important kind of network security quantitative evaluation measure, attack probability and its computing methods have been studied for a long time. Many models have been proposed for performing evaluation of network security. Graphical models such as attack graphs have become the main-stream approach. Attack graphs which capture the relationships among vulnerabilities and exploits show us all the possible attack paths that an attacker can take to intrude all the targets in the network. The traffic to different hosts or servers may differ from each other. The hosts or servers with big traffic may be more risky since they are often important hosts or servers, and intruders may have more contacts and understanding with them. In our cyber-attacks prediction model, they used attack graphs to capture the vulnerabilities in the network. In addition we consider 3 environment factors that are the major impact factors of the cyber-attacks in the future. They are the value of assets in the network, the

usage condition of the network and the attack history of the network. Cyber-attacks prediction is an important part of risk management. Existing cyber-attacks prediction methods did not fully consider the specific environment factors of the target network, which may make the results deviate from the true situation. In this paper, we propose a cyber-attacks prediction model based on the Bayesian network. We use attack graphs to represent all the vulnerabilities and possible attack paths. Then we capture the environmental factors using the Bayesian network model. Cyber-attacks predictions are performed on the constructed Bayesian network.

Title : A Prediction Model of DoS Attacks Distribution Discrete Probability
Author: Wentao Zhao, Jianping Yin

This paper begins with the relation between network traffic data and the amount of DoS attack, and then proposes a clustering method based on the genetic optimization algorithm to implement the classification of DoS attack data. This method first gets the proper partition of the relation between the network traffic and the amount of DoS attack based on the optimized clustering and builds the prediction sub-models of DoS attack. Meanwhile, with the Bayesian method, the calculation of the output probability corresponding to each sub-model is deduced and then the distribution of the amount of DoS attack in some range in future is obtained. This paper describes the clustering problem first, and then utilizes the genetic algorithm to implement the optimization of clustering methods. Based on the optimized clustering on the sample data, we get various categories of the relation between traffic and attack amounts, and then build up several prediction sub-models about DoS attack. Furthermore, according to the Bayesian method, we deduce discrete probability calculation about each sub-model and then get the distribution discrete probability prediction model for DoS attack.

Title : Adversarial Examples: Attacks and Defenses for Deep Learning
Author: Xiaoyong Yuan , Pan He, Qile Zhu, and Xiaolin Li

It reviewed the recent findings of adversarial examples in DNNs. We investigated the existing methods for generating adversarial examples. A taxonomy of adversarial examples was proposed. We also explored the applications and countermeasures for adversarial examples. This paper attempted to cover the state-of-the-art studies for adversarial examples in the DL domain. Compared with recent work on adversarial examples, we analyzed and discussed the current challenges and potential solutions in adversarial examples. However, deep neural networks (DNNs) have been recently found vulnerable to well-designed input samples called adversarial examples. Adversarial perturbations are imperceptible to humans but can easily fool DNNs in the testing/deploying stage. The vulnerability to adversarial examples becomes one of the major risks for applying DNNs in safety-critical environments. Therefore, attacks and defenses on adversarial examples draw great attention. In this paper, we review recent findings on adversarial examples for DNNs, summarize the methods for generating adversarial examples, and propose taxonomy of these methods. Under the taxonomy, applications for adversarial examples are investigated. We further elaborate on countermeasures for adversarial examples. In addition, three major challenges in adversarial examples and the potential solutions are discussed.

Title : Distributed Secure Cooperative Control Under Denial-of-Service Attacks From Multiple Adversaries
Author: Wenying Xu, Guoqiang Hu

This paper has investigated the distributed secure control of multiagent systems under DoS attacks. We focus on the investigation of a jointly adverse impact of distributed DoS attacks from multiple adversaries. In this scenario, two kinds of communication schemes, that is, sample-data and event-triggered communication schemes, have been discussed and, then, a fully distributed control protocol has been developed to guarantee satisfactory asymptotic consensus. Note that this protocol has strong robustness and high scalability. Its design does not involve any global information, and its efficiency has

been proved. For the event-triggered case, two effective dynamical event conditions have been designed and implemented in a fully distributed way, and both of them have excluded Zeno behavior. Finally, a simulation example has been provided to verify the effectiveness of theoretical analysis. Our future research topics focus on fully distributed event/self-triggered control for linear/nonlinear multiagent systems to gain a better understanding of fully distributed control.

4.EXISTING-SYSTEM

They proposed first to create a contrastive self-supervised learning to the anomaly detection problem of attributed networks. CoLa mainly consists of three components: contrastive instance pair sampling, GNN-based contrastive learning model, and multi round sampling-based anomaly score computation. Their model captures the relationship between each node and its neighbouring structure and uses an anomaly-related objective to train the contrastive learning model. We believe that the proposed framework opens a new opportunity to expand self-supervised learning and contrastive learning to increasingly graph anomaly detection applications. The multi round predicted scores by the contrastive learning model are further used to evaluate the abnormality of each node with statistical estimation. The training phase and the inference phase. In the training phase, the contrastive learning model is trained with sampled instance pairs in an unsupervised fashion. After that the anomaly score for each node is obtained in the inference phase.

Limitations

- The performance is not good and it gets complicated for other networks.
- The performance metrics like recall F1 score and comparison of machine learning algorithms is not done.

5.PROPOSED SYSTEM

The proposed model is to build a machine learning model for anomaly detection. Anomaly detection is an important technique for recognizing fraud activities, suspicious activities, network intrusion, and other abnormal events that may have great significance but are difficult to detect. The machine learning model is built by applying proper data science techniques like variable identification that is the dependent and independent variables. Then the visualisation of the data is done to insights of the data .The model is built based on the previous dataset where the algorithm learns data and gets trained different algorithms are used for better comparisons. The performance metrics are calculated and compared

Logistic Regression :

In other words, the logistic regression model predicts P(Y=1) as a function of X. Logistic regression Assumptions:

- Binary logistic regression requires the dependent variable to be binary.
- For a binary regression, the factor level 1 of the dependent variable should represent the desired outcome.
- Only the meaningful variables should be included.
- The independent variables should be independent of each other. That is, the model should have little.
- The independent variables are linearly related to the log odds.
- Logistic regression requires quite large sample sizes.

Decision Tree:

It is one of the most powerful and popular algorithms. Decision-tree algorithms fall under the category of supervised learning algorithms. It works for both

continuous as well as categorical output variables. Assumptions of Decision tree:

- At the beginning, we consider the whole training set as the root.
- Attributes are assumed to be categorical for information gain, attributes are assumed to be continuous.
- On the basis of attribute values records are distributed recursively.
- We use statistical methods for ordering attributes as root or internal nodes.

Decision trees build classification or regression models in the form of a tree structure. It breaks down a data set into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. A decision node has two or more branches and a leaf node represents a classification or decision. The topmost decision node in a tree which corresponds to the best predictor called root node. Decision trees can handle both categorical and numerical data. Decision trees build classification or regression models in the form of a tree structure. It utilizes an if-then rule set which is mutually exclusive and exhaustive for classification. The rules are learned sequentially using the training data one at a time. Each time a rule is learned, the tuples covered by the rules are removed.

This process is continued on the training set until meeting a termination condition. It is constructed in a top-down recursive divide-and-conquer manner. All the attributes should be categorical. Otherwise, they should be discretized in advance. Attributes in the top of the tree have more impact in the classification and they are identified using the information gain concept. A decision tree can be easily over-fitted generating too many branches and may reflect anomalies due to noise or outliers.

Random Forest:

The following are the basic steps involved in performing the random forest algorithm:

- Pick N random records from the dataset.
- Build a decision tree based on these N records.
- Choose the number of trees you want in your algorithm and repeat steps 1 and 2.
- In case of a regression problem, for a new record, each tree in the forest predicts a value for Y (output). The final value can be calculated by taking the average of all the values predicted by all the trees in the forest. Or, in case of a classification problem, each tree in the forest predicts the category to which the new record belongs. Finally, the new record is assigned to the category that wins the majority vote.

Support Vector Machines:

A classifier that categorizes the data set by setting an optimal hyperplane between data. I chose this classifier as it is incredibly versatile in the number of different kernelling functions that can be applied and this model can yield a high predictability rate. Support Vector Machines are perhaps one of the most popular and talked about machine learning algorithms. They were extremely popular around the time they were developed in the 1990s and continue to be the go-to method for a high-performing algorithm with little tuning.

- How to disentangle the many names used to refer to support vector machines.
- The representation used by SVM when the model is actually stored on disk.
- How a learned SVM model representation can be used to make predictions for new data.
- How to learn an SVM model from training data.
- How to best prepare your data for the SVM algorithm.
- Where you might look to get more information on SVM.

7.SYSTEM SPECIFICATION

HARDWARE REQUIREMENTS

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. The minimal hardware requirements are as follows:

1. Processor : Pentium IV/III
2. Hard disk : Minimum 80 GB

SOFTWARE REQUIREMENTS

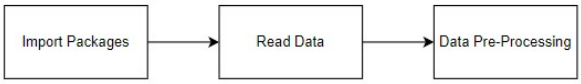
Software requirements deal with defining resource requirements and prerequisites that need to be installed on a computer to provide functioning of an application. The minimal software requirements are as follows,

1. Operating System : Windows
2. Tool : Anaconda with Jupyter Notebook

8.MODULE DESCRIPTION

Data Pre-processing:

Pre-processing refers to the transformations applied to our data before feeding it to the algorithm. Data Preprocessing is a technique that is used to convert the raw data into a clean data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not feasible for the analysis. To achieve better results from the applied model in Machine Learning the method of the data has to be in a proper manner. Some specified Machine Learning model needs information in a specified format; for example, Random Forest algorithm does not support null values. Therefore, to execute random forest algorithms null values have to be managed from the original raw data set. And another aspect is that data sets should be formatted in such a way that more than one Machine Learning and Deep Learning algorithms are executed in a given dataset.

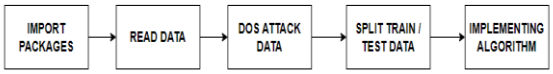


Module-02:

In computing, a denial-of-service attack (DoS attack) is a cyber-attack in which the perpetrator seeks to make a machine or network resource unavailable to its intended users by temporarily or indefinitely disrupting services of a host connected to the Internet. Denial of service is typically accomplished by flooding the targeted machine or resource with superfluous requests in an attempt to overload systems and prevent some or all legitimate requests from being fulfilled. In a distributed denial-of-service attack (DDoS attack), the incoming traffic flooding the victim originates from many different sources. This effectively makes it impossible to stop the attack simply by blocking a single source. A DoS or DDoS attack is analogous to a group of people crowding the entry door of a shop, making it hard for legitimate customers to enter, disrupting trade.

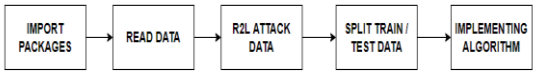
A distributed denial-of-service (DDoS) is a large-scale DoS attack where the perpetrator uses more than one unique IP address, often thousands of them.[10] A distributed denial of service attack typically involves more than around 3–5 nodes on different networks; fewer nodes may qualify as a DoS attack but is not a DDoS attack.[11][12] Since the incoming traffic flooding the victim originates from different sources, it may be impossible to stop the attack simply by using ingress filtering. It also makes it difficult to distinguish legitimate user traffic from attack traffic when spread across multiple points of origin. As an alternative or augmentation of a DDoS, attacks may involve forging of IP sender addresses (IP address spoofing) further complicating identifying and defeating the attack. An application layer DDoS attack (sometimes referred to as layer 7 DDoS attack) is a form of DDoS attack where attackers target application-layer processes. The attack over-exercises specific functions or features of a website with the intention to disable those functions or features. This application-layer

attack is different from an entire network attack, and is often used against financial institutions to distract IT and security personnel from security breaches.



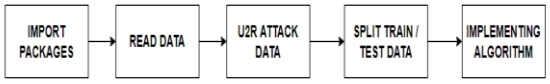
Module-03:

Now-a-days, it is very important to maintain a high level of security to ensure safe and trusted communication of information between various organizations. But secured data communication over the internet and any other network is always under threat of intrusions and misuses. To control these threats, recognition of attacks is a critical matter. Probing, Denial of Service (DoS), Remote To User (R2L) attacks are some of the attacks which affect a large number of computers in the world daily. Detection of these attacks and prevention of computers from it is a major research topic for researchers throughout the world.



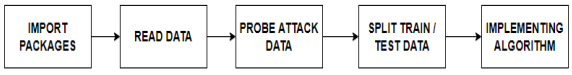
Module-04:

Remote to local attack (r2l) has been widely known to be launched by an attacker to gain unauthorized access to a victim machine in the entire network. Similarly user to root attack (u2r) is usually launched for illegally obtaining the root's privileges when legally accessing a local machine. Buffer overflow is the most common of U2R attacks. This class begins by gaining access to a normal user while sniffing around for passwords to gain access as a root user to a computer resource. Detection of these attacks and prevention of computers from it is a major research topic for researchers throughout the world.



Module-05:

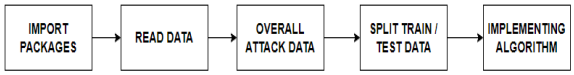
Probing attacks are an invasive method for bypassing security measures by observing the physical silicon implementation of a chip. As an invasive attack, one directly accesses the internal wires and connections of a targeted device and extracts sensitive information. A probe is an attack which is deliberately crafted so that its target detects and reports it with a recognizable “fingerprint” in the report. The attacker then uses the collaborative infrastructure to learn the detector's location and defensive capabilities from this report. This is an attack where the attacker attempts to gather information about the target machine or the network, to map out the network. Information about the target may reveal useful information such as open ports, its IP address, hostname, and operating system. Network Probe is the ultimate network monitor and protocol analyzer to monitor network traffic in real-time, and will help you find the sources of any network slow-downs in a matter of seconds.



Module-06:

Increasingly, attacks are executed in multiple steps, making them harder to detect. Such complex attacks require that defenders recognize the separate stages of an attack, possibly carried out over a longer period, as belonging to the same attack. Complex attacks can be divided into exploration and exploitation phases. Exploration involves identifying vulnerabilities and scanning and testing a system. It is how an attacker gathers information about the system. Exploitation involves gaining and maintaining access. At this stage, the attacker applies the know-how gathered during the exploration stage. An example of a complex attack that combines

exploration and exploitation is a sequence of a phishing attack, followed by an exfiltration attack. First, attackers will attempt to collect information on the organization they intend to attack, e.g., names of key employees. Then, they will craft a targeted phishing attack. The phishing attack allows the attackers to gain access to the user's system and install malware. The purpose of the malware could be to extract files from the user's machine or to use the user's machine as an attack vector to attack other machines in the organization's network. A phishing attack is usually carried out by sending an email purporting to come from a trusted source and tricking its receiver to click on a URL that results in installing malware on the user's system. This malware then creates a backdoor into the user's system for staging a more complex attack. Phishing attacks can be recognized both by the types of keywords used in the email (as with a spam email), as well as by the characteristics of URLs included in the message. Features that have been used successfully to detect phishing attacks include URLs that include IP addresses, the age of a linked-to domain, and a mismatch between anchor and text of a link.

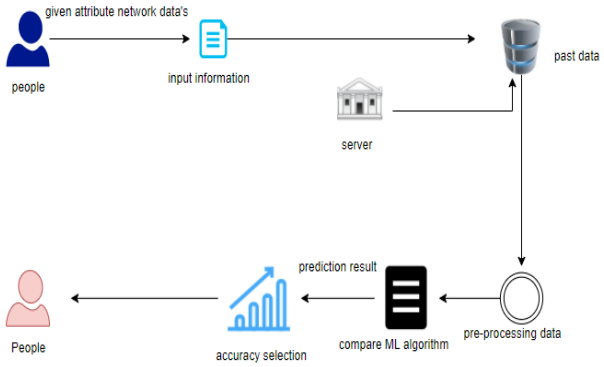


Module-07:

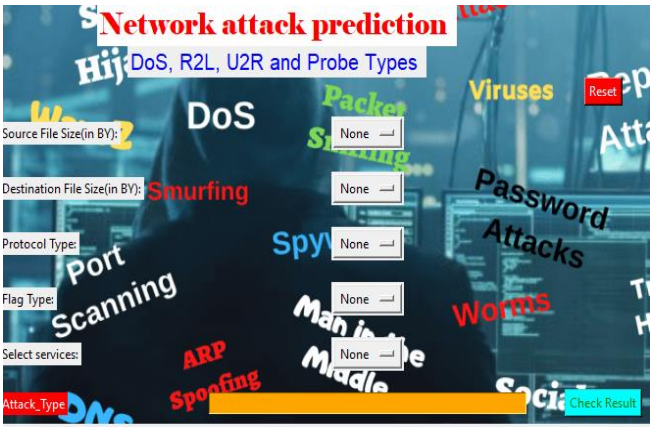
GUI means Graphical User Interface. It is the common user Interface that includes Graphical representation like buttons and icons, and communication can be performed by interacting with these icons rather than the usual text-based or command-based communication. A common example of a GUI is Microsoft operating systems.

The graphical user interface (GUI) is a form of user interface that allows users to interact with electronic devices through graphical icons and audio indicator such as primary notation, instead of text-based user interfaces, typed command labels or text navigation. GUIs were introduced in reaction to the perceived steep learning curve of command-line interfaces (CLIs) which require commands to be typed on a computer keyboard.

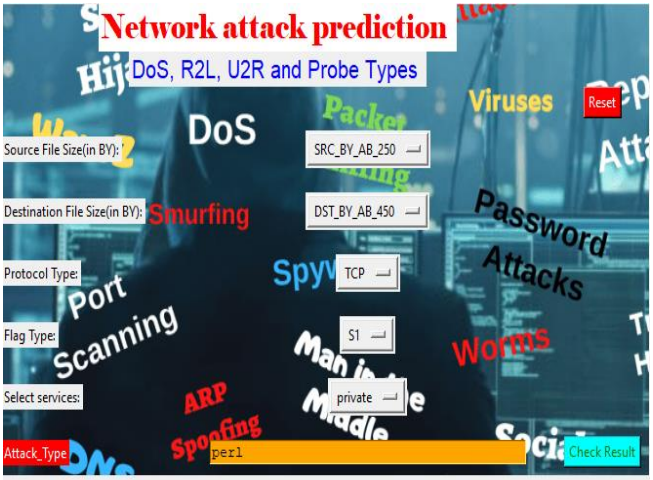
9. ARCHITECTURE DIAGRAM



10.SCREENSHOTS



RESULT PAGE :



11.CONCLUSION

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on a public test set is higher accuracy score will be found by comparing each algorithm with the type of all network attacks for future prediction results by finding best connections. This brings some of the following insights about diagnosing the network attack of each new connection. To present a prediction model with the aid of artificial intelligence to improve over human accuracy and provide with the scope of early detection. It can be inferred from this model that area analysis and use of machine learning techniques is useful in developing prediction models that can help network sectors reduce the long process of diagnosis and eradicate any human error.

12.REFERENCES

[1] Z. Liu, C. Chen, X. Yang, J. Zhou, X. Li, and L. Song, "Heterogeneous graph neural networks for malicious account detection," in Proc. 27th ACM Int. Conf. Inf. Knowl. Manage., Oct. 2018, pp. 2077–2085.

[2] L. Tang and H. Liu, "Relational learning via latent social dimensions," in Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD), 2009, pp. 817–826.

[3] Y. Zhang et al., "Your style your identity: Leveraging writing and photography styles for drug trafficker identification in darknet markets over attributed heterogeneous information network," in Proc. World Wide Web Conf. (WWW), 2019, pp. 3448–3454.

[4] R. Ying, R. He, K. Chen, P. Eksombatchai, W. L. Hamilton, and J. Leskovec, "Graph convolutional neural networks for Web-scale recommender systems," in Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Jul. 2018, pp. 974–983.

[5] W. Fan et al., "Graph neural networks for social recommendation," in Proc. World Wide Web Conf. (WWW), 2019, pp. 417–426.

[6] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in Proc. Int. Conf. Learn. Represent., 2017, pp. 1–14.

[7] Yixin Liu, Zhao Li,Shirui Pan,Chen Gong, ChuanZhou , and George Karypis,"Anomaly Detection on Attributed Networks via Contrastive Self-Supervised Learning "(2020)

[8] T. N. Kipf and M. Welling, "Variational graph auto-encoders," 2016, arXiv:1611.07308. [Online]. Available: http://arxiv.org/abs/1611.07308

[9] G. Pang, C. Shen, L. Cao, and A. van den Hengel, "Deep learning for anomaly detection: A review," 2020, arXiv:2007.02500. [Online]. Available: http://arxiv.org/abs/2007.02500

[10] K. Ding, J. Li, R. Bhanushali, and H. Liu, "Deep anomaly detection on attributed networks," in Proc. SIAM Int. Conf. Data Mining. Philadel phia, PA, USA: SIAM, 2019, pp. 594–602.

[11] Y. Chen, X. Sean Zhou, and T. S. Huang, "One-class SVM for learning in image retrieval," in Proc. Int. Conf. Image Process., vol. 1, 2001, pp. 34–37.

[12] X. Xu, N. Yuruk, Z. Feng, and T. A. J. Schweiger, "SCAN: A structural clustering algorithm for networks," in Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD), 2007, pp. 824–833.

[13] B. Perozzi and L. Akoglu, "Scalable anomaly ranking of attributed neighborhoods," in Proc. SIAM Int. Conf. Data Mining, Jun. 2016, pp. 207–215.

[14] J. Li, H. Dani, X. Hu, and H. Liu, "Radar: Residual analysis for anomaly detection in attributed networks," in Proc. 26th Int. Joint Conf. Artif. Intell., Aug. 2017, pp. 2152–2158.

[15] Z. Peng, M. Luo, J. Li, H. Liu, and Q. Zheng, "ANOMALOUS: A joint modeling approach for anomaly detection on attributed networks," in Proc. 27th Int. Joint Conf. Artif. Intell., Jul. 2018, pp. 3513–3519.

[16] G. Pang, C. Shen, H. Jin, and A. van den Hengel, "Deep weakly supervised anomaly detection," 2019, arXiv:1910.13601. [Online]. Available: https://arxiv.org/abs/1910.13601