EasyChair Preprint

№ 6051

Cyber Security: a Major Issues and Their Remedies

Sudarshan Bhalshankar

July 11, 2021

CYBER SECURITY : A MAJOR ISSUES AND THEIR REMEDIES

Sudarshan Dilip Bhalshankar.

M.E [CAD/CAM] , B.E. [Mechanical] , D. A. E.

Sudarshanbhalshankar111@gmail.com

PradnyaSurya Engineering Works. Pvt Ltd.

Solapur, Maharashtra, India.

pradnyasuryaengineerinworks@gmail.com

ABSTRACT

Machine learning is one among the widely used techniques in the field of ***COMPUTER SCIENCE***, and it finds wide applications in ***CYBERSECURITY***, natural language processing, image processing, pattern recognition and other fields. Despite successful applications of machine learning techniques in various cases like intrusion detection, ***DATA SECURITY***, network security and cyber security, the performance of such techniques significantly decreases because these machine learning algorithms and ***EQUIVALENT TRAINING*** data are susceptible to various ***SECURITY ATTACKS.*** This calls for in depth analysis of security threats and related defensive techniques of machine learning intensively which leads to the motivation of this review paper. An exploratory research on the existing methodology of machine learning techniques are studied and the research gap in the field of adaption of machine learning ***DEFENSIVE TECHNIQUES*** to cybersecurity is analyzed.

Keywords :- Defensive-Techniques , Data-Analysis , IT-Attacks , Social , Obsolete-Content , Cybersecurity , Machine-Learning , Algorithms.

❖ *INTRODUCTION.*

With the increase in the connectivity between human social lives with internet is making it to reveal the lifestyle of human beings to outside world which is prone to security attacks. To ensure personal computer, Networks, programs and information from attacks and modification, unapproved access, or devastation, Cybersecurity is proposed which is a combination of technology and procedures. Security includes at two steps namely at network and data. Security breach external intrusions and internal intrusions. Machine learning is a technique wherein the system gets better with experience. In legacy systems the system used.

To perform exactly the same task ‘n’ number of times without smartness. Cyber security involves adding security measures to prevent security breach of the system. Cyber security is significant as data and network is what it makes the communication possible in every field [7]. The components of cyber security consist of network security, application security, mobile security, data security, endpoint security [1]. The use of internet, computer application and android application has seen an immense growth and have turned out to be an integral part of younger generation people. Hence there the need to protect data has become significant.

Attackers are able to potentially use several paths by means of application to do havoc to your business or organization. According to the National Institute of Standards and Technology (NIST),

American companies as early as 2017 suffered losses of up to 65.6 billion dollars following IT attacks [6]. The exponential increase in percentage of attacks and security breaches has paved way for artificial intelligence and machine learning based methodologies in detecting cyber security threats[8]. In order to provide the best security applications be accepted and appropriate level of security be obtained, security-related benchmarks are very important [2]. In order to provide a self-learning based technique based machine learning technique is used. In this study, a general assessment of artificial intelligent and machine learning techniques are provided.

❖ *LITERATURE REVIEW.*

Saiyed Kashif Shaukat and Vinay J. Ribeiro have presented a cryptographic Ransomware called RansomWall which is a defense system. This defense system has layered architecture and is used for safeguarding against cryptographic ransomeware. The layers in the RansomWall are arranged in computational order that are produced in sample's execution. Every RansomWall level is depends upon a specific functionality. This technique is been executed on windows operating system A pattern is discovered which is common over various Cryptographic Ransomware upon investigation of broad data set .The Ransom Wall layered architecture is based on Static and Dynamic investigation hybrid approach to find out feature extracted. The layered approach monitors malicious actions by Cryptographic Ransomware to breakdown resistances. Of customer's system and tracks file operation performed over encrypting user data files. Cryptographic Ransomware generally uses complicated tactics such as polymorphic and metamorphic in order to avoid signature-dependent identification mechanism. The best popular known attack are zero-day attacks. A comprehensive model based on machine learning technique is used to extract the compressed feature that identifies zero-day samples. With the usage of Gradient Tree Boosting algorithm, the Machine Learning Layer offers top performance metrics. Using this model, a detection rate (TPR) of 98.25% with close to zero false positives is observed. A Virus Total connected to sixty security engines having 30 zero-day intrusion samples found to have less than 10% detection rate are gathered. All the samples in real-time is successfully detected by RansomWall. [1]

Ositis Anastasia Petrovna, Kalashnikov Evgeniy Alexandrovich and Kondybayeva Almagul Baurzhanovna have proposed a neural network called MP KK hybrid network (Multilayer Perceptron - Kohonen-KK map). The layered framework of immediate connection with a network of MP-KK network is connected with Kohonen map.The possibility of identifying known attacks was 91%, the likelihood of recognizing unknown attacks was 86%.This hybrid network has two steps. First is to train the network with self-organization with input network. After the training is finished, the neurons have a static weights. The neurons output indicators are planned and using a standard way a multilayer perceptron learns with the teacher. The MP KK hybrid network provides flexibility to classify attacks which is the primary benefit of utilizing neural network. The neural network is proficient for examining information from the network, irrespective of whether the statistics is partial or inaccurate and also this network is capable of performing analysis with non-linear form data. The use of neural network has a benefit that it can "learn" the indications of attacks and recognize scenarios that are not characteristic of previously observed scenarios. A neural network with high of accuracy identifies known suspicious attacks when trained with proper data set. The Drawback of MP KK hybrid network is that the volume of an artificial neural network to diagnose intrusion features is completely reliant on the exact knowledge of the system, the statistics used for training

and the training systems are important. In order to ensure statistically accurate outcomes a lot of data is needed for training procedure. Neural network for intrusion detection preparation needs a huge number distinct attacks, and this quantity of facts is tough to get. [2] Bilgehan Arslan, Mehtap Ulker and Seref Sagiroglu discuss about usage of machine learning for securing biometric systems. The usage of machine learning for biometric systems can be defined as determining characteristics and corresponding units for biometric recognition, verification, identification processes. Detection of attacks on biometric structures and modules. Biometric information as a feature or customer template of fingerprints is commonly kept in a database.

throughout this procedure, the data on message channels used in the acknowledgement procedure can be taken or the records containing features can be prone to attacks.

In order to avoid such attacks, the biometric properties are slightly transformed to a user template instead of recording the biometric properties. There are four types of methods which are used for biometric template namely transformation of the characteristics, biometric cryptosystems, watermarking and haptic-biometric template protection in transportation layer the replay attack can be stopped using (Genetic and Evolutionary Feature extraction-Machine Learning) GEFE machine learning technique.To generate disposable Feature Extractors, a technique known as GEFE ML can be utilized in order to avoid biometric replay attacks. This technique evolves feature extractors that are distinct and high performing by using Genetic & Evolutionary Computations. Fuzzy extraction method can be used to prevent Dos Attack in physical layer in physical layer, the spoofing attack can be prevent by means of SVM/HMM/Bayesian Classifier.

Support vector machines (SVMs) are used for classification and regression and it is associated to supervised learning approaches. Bayesian classifier is a classification technique based on Bayes' Theorem with a hypothesis of independence among predictors. Mapping a classification of observations to a classification of labels is done through HMM (Hidden Markov model). In application layer, the variation attack can be prevented using (Scale Invariant Feature Transform) SIFT. The fraud pattern is being analyzed on network using machine learning techniques The training of the Sift system is done by the Sift system itself by learning patterns of fraud which are detected on its user locations, and this become part of the machine-learning network. Although in pattern security phase, Machine Learning techniques are used widely, but the Machine learning techniques are not used widely in the biometric system attack and prevention mechanisms phases. Instead Statistical methods and filtering methods finds usage more widely. [3] Anna L. Buczak and Erhan Guven discuss about (Machine Learning and Data Mining) ML and DM algorithms used for cyber security. According to the authors, the training data should be clean and should have statistical properties to apply machine learning and data mining algorithms. Also information whether system is in offline or online mode is important. All these queries will assist in shaping the most

appropriate ML approach. According to this paper [4], a simple distribution like Gaussian distribution cannot be used to model the network data because, in training a single network packet may have a payload which is connected to loads of network protocols and user activities. In such a case, the techniques such as HMMs or Bayesian networks is not the best approach as the data is not possessing the properties which are most suitable for them. The Evolutionary computation techniques are time consuming to execute and this technique cannot be applied for systems that train online as it is not most suitable. Random Forests finds advantageous usage when the training data is unusual. Decision trees, evolutionary computation, and association rules finds advantageous usage when the attack signature capture is significant. Machine Learning and Data Mining techniques require straining data and these techniques cannot work deprived of such data , and it is tough and tedious to capture those training data. This paper also throws light on the need for the approaches that are fast incremental learning and might be applied for day-today updates of replicas for anomaly and misuse detection as the models needs to be retrained. [4] Qiang Liu, Pan Li,]Wentao Zhao, Wei Cai , Shui Yu and Victor C.M. Leung describe defensive techniques of Machine Learning namely Reject on Negative Impact (RONI) which effectively removes adversarial samples that are injected into training data. It scales to a variety of classifiers. Adversarial training which is easy to understand and implement and it scales to a variety of classifiers. Defense distillation is used to obtain a smoother DNN model by reducing its sensitivity regarding input perturbations. It improves the generalization capability of a DNN. It effectively mitigates adversarial samples crafted by fast gradient sign method (FGSM).Ensemble method - It is flexible to integrate multiple classifiers or different defensive methods. Differential privacy - 1) it preserves the privacy of training data. 2) It preserves the privacy of learning algorithms. [5]

❖ *DISADVANTAGES OF DEFENSIVE TECHNIQUES.*

Reject on Negative Impact (RONI) has a lack of extensive performance evaluation in a variety of application scenarios. In Adversarial training effectiveness depends on the adversarial samples in the training phase. In Defense distillation it is weak to defend against adversarial samples crafted by Jacobian-based saliency map approach (JSMA).In Ensemble method it is not robust to adversarial samples with transferability. In Differential privacy it influences the performance of classifiers on legitimate data.

❖ *RESEARCH GAP IDENTIFIED.*

Need for new machine learning technique to mitigate evolving security threats assessing the right machine learning technique based o decision system which works under adversarial environments.

❖ *CONCLUSION.*

This paper presents the overview of different machine learning defensive techniques for cybersecurity. Each technique used for implementing an intrusion detection system has its own particular points of interest and impediments. Hence it is problematic to pick a specific technique to execute an intrusion detection system over the others. Datasets for network intrusion detection are very vital for preparing and testing systems. The ML methods do not work without training data and getting such a dataset is troublesome and tedious. However, there are numerous issues with the current public dataset for example uneven data, obsolete content. These issues have constrained the advancement of research here. Network information update very fast, which leads to the training and the usage of ML model difficult, model should be retrained long-term and rapidly. Consequently incremental learning and lifelong learning will be the emphasis in the investigation of this arena in the future.

❖ *REFERENCES.*

1. S. Shaukat and V. Ribeiro, "RansomWall: A layered defense system against cryptographic ransomware attacks using machine learning", 2018 10th International Conference on Communication Systems & Networks (COMSNETS), 2018.
2. O. Petrovna, K. Baurzhanovna and K. Alexandrovich, "Optimized data protection and network attacks' analysis throughout security classes identification with machine learning methods in radiofrequency applications: Data protection and network attacks' analysis in radiofrequency applications", 2018 Systems Of Signals Generating and Processing in the Field of on Board Communications, 2018.
3. B. Arslan, M. Ulker and S. Sagiroglu, "Machine Learning Methods Used in Evaluations of Secure Biometric System Components", 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), 2017.
4. Buczak and E. Guven, "A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection", IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153-1176, 2016.
5. Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu and V. Leung, "A Survey on Security Threats and Defensive Techniques of Machine Learning: A Data Driven View", IEEE Access, vol. 6, pp. 12103-12117, 2018.
6. Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou and C. Wang, "Machine Learning And Deep Learning Methods for Cybersecurity", IEEE Access, pp. 1-1, 2018.
7. Barkly, "WannaCry Ransomware Statistics: The Numbers Behind the Outbreak," May 2017. [Online]. Available: https://blog.barkly.com/ wannacry-ransomware-statistics-2017.
8. B. Eren, "A New Proposal for the Effective Design and Implementation of Biometric Technologies: Multimodal Technology", Mimar Sinan University Institute of science, Master Thesis, Istanbul, pp. 5-28, 2009.

❖ *Author Details.*

Name :- Sudarshan Dilip Bhalshankar.

Educational Qualification :- M.E [CAD/CAM] ,

B.E. [Mechanical] , D. A. E.

Mail :- Sudarshanbhalshankar111@gmail.com

Founder Of :- PradnyaSurya Engineering Works. Pvt Ltd.

Address :- Solapur, Maharashtra, India.

Pin code :- 413005

Mail :- pradnyasuryaengineerinworks@gmail.com