PREDICTION OF CYBER ATTACKS USING MACHINE LEARNING ALGORITHMS

S. Arumai Shiney, AP/CSE DEPARTMENT, S.A ENGINEERING COLLEGE

P.Jayasri Archana Devi,AP/CSE DEPARTMENT, JAYA SAKTHI ENGINEERING COLLEGE

S.Selvakumaran, AP/AIDS DEPARTMENT,RAJALAKSHMI INSTIUTE OF TECHNOLOGY

M.Jayanthi, AP/CSE DEPARTMENT, JAYA SAKTHI ENGINEERING COLLEGE

ABSTRACT

One of the world's biggest issues nowadays is cyber-attacks. Every day, they wreak serious economic harm to both persons and nations. Cybercrime is also on the rise along with cyber-attacks. Understanding attack tactics and identifying cybercrime perpetrators are crucial in the fight against crime and criminals. Cyber-attack detection and prevention are challenging undertakings. Yet, academics have lately developed security models and made predictions using artificial intelligence techniques to solve these issues. There are many ways for predicting crimes that can be found in the literature. On the other hand, they struggle to foresee the strategies used in cybercrime and cyber-attacks. By using actual data to pinpoint an assault and its perpetrator, this issue can be solved. The information includes the kind of crime, the perpetrator's gender, the damage, and the attack techniques. Applications made by people who were subject to cyber-attacks can provide the forensic units with data. In this study, we use machine learning to examine two alternative models of cybercrimes and estimate the impact of defined variables on the identification of the attack vector and the perpetrator. In our methodology, we employed eight machine learning techniques and found that their accuracy rates were comparable. With an accuracy percentage of 95.02%, the Support Vector Machine Linear was proven to be the most effective cyber-attack technique. With a high degree of accuracy, the first model allowed us to forecast the types of attacks that the victims were most likely to experience. The most accurate method for identifying attackers, with a 65.42% accuracy rate, was logistic regression. In the second model, we forecasted whether the features of the perpetrators might be

compared to identify them. According to our findings, the likelihood of a cyber-attack diminishes as the victim's income and level of education rise. We anticipate that the suggested model will be applied by cybercrime units. Also, it will make it easier and more effective to detect cyber-attacks and defend against them.

Keywords: Machine learning, supervised learning, DOS attack,

1. INTRODUCTION

Machine learning is a branch of computer science and artificial intelligence (AI) that focuses on using data and algorithms to simulate human learning, gradually improving the model's accuracy. Machine learning algorithms create a model from sample data, commonly referred to as training data, without explicitly specifying the predictions or choices Machine learning algorithms are utilized in a wide range of applications, including speech recognition, email filtering, computer vision, medicine, when it is challenging or impractical to create traditional algorithms to carry out the required functions.

Machine learning includes deep learning as a subclass. Artificial neural networks, designed to mimic how people think and learn, are used in its operation. Prior to recently, the complexity of neural networks was constrained by processing capability. Without the assistance of a human, any pattern recognition issue can be resolved using deep learning techniques. More businesses are using deep learning, the fastest-growing subfield of machine learning, to create new business models.

The majority of firms use a defensive and retaliatory strategy to defend against cyber-attacks. Threats are only eliminated and analyzed after they have been identified, by which time the network has already been penetrated and valuable data stolen. Most firms use the same technologies and security measures for intrusion detection and prevention, such as firewalls and antivirus software, along with access controls like passwords. Yet, it's reasonable to argue that reflexive responses are, at best, damage control tactics and are generally useless given how complex and varied cybercrimes have become in recent years and how frequent the

attacks that don't make the tabloids are. Yet, the outlook for cyber security isn't as dire as we probably made it seem.

2. Related work

This section addresses the research carried out on image enhancement by various researchers in the literature.

The Cyber-Physical Systems (CPSs) constitute an emerging research field that has captured the interest of many scholars, according to the author of this paper [1]. The paper discusses the need for these systems to be implemented in various application domains, as well as the research challenges of defining an appropriate formalism that represents more than just networking and information technology—the information and knowledge will be integrated into physical objects. The paper begins by defining CPS. The presentation concludes with a brief state of the art on the key CPS research fields (generic architecture, design principles, modeling, dependability, and implementation), as CPSs are anticipated to play a significant role in the design and development of future engineering systems.

Using the VERIS Community Database, this paper [2] examines the range of risk associated with a cyber-security incident occurring in a cyber-physically enabled environment. The results showed that most victims were from western states, most known actors were from the US and Russia, and geographic origin tended to reflect world politics. Information was the asset that was attacked the most frequently, and most assault strategies relied on abusing privileges. Large-scale internal security breaches were the most prevalently a result of human error. Due to a basic trade-off between usability and security in the design of computer systems, this tends to reveal that access in any form appears to be the source of vulnerability rather than event details.

The authors of this paper [3] propose the term "cyber-physical systems" to describe systems that interact with computers, communication channels, and physical devices to address a real-world issue. Cyber-Physical Systems are currently one of the top targets of hackers as the industry 4.0 revolution approaches, and any harm to them results in significant losses for a country. Some documented examples

involved security breaches on Cyber-Physical Systems, according to reliable sources. Understanding the theoretical underpinnings of security in the digital era was a topic of discussion everywhere. Security issues with the cyber-physical system, however, are largely understudied. In addition, a few technologies were made available to help with Cyber-Physical System security issues. For the cyber-physical system, better understanding and the introduction of more security measures are needed.

The author of this paper [4] discusses Cyber-Physical Systems (CPS), which are a collection of interconnected systems that can observe and control physical processes and objects. They share many similarities with Internet of Things (IoT) systems, but CPS focuses on how physical, networking, and computational processes interact. The Internet of Cyber-Physical Things, a new component of CPS as a result of their integration with IoT (IoCPT). A greater range of services and applications, such as e-Health, smart homes, and e-Commerce, are made possible by the rapid and significant evolution of CPS, which has an impact on many elements of people's way of life. Nevertheless, integrating the physical and virtual worlds creates new, risky security issues. As a result, both industries and researchers are interested in CPS security.

The authors of this paper [5] argue that, as in other domains with comparable characteristics, a human cyber defender's situation awareness (SA) is crucial to the effectiveness of decision outcomes in cyber security. The majority of current research on cyber situation awareness is focused on computers and information systems that combine diverse data. Cyberspace visualization projects, for instance, need the fusion of data from several sources. From this perspective, a successful cyber-SA is evaluated differently from one that is human-centered. In contrast, it's uncommon for us to evaluate human cognitive awareness in cyberspace. This partly reflects the requirement, based on prior theory, to specify key informational components that people must see before attempting to explain how people mix these components to understand the situation.

This paper [6] argues that enabling decision-making in cyberspace requires maintaining a high level of context awareness (CDSA). Only by mimicking actual occurrences as accurately as possible can this be trained. Thus, a Cyber Range is a crucial instrument. It makes it possible to incorporate a large number of people and

enables for the simulation of complicated networks. In this work, we first discuss the critical role that cyber ranges play in enhancing CDSA, and then we demonstrate how a cyber-range may be set up to enable such training.

This paper [7] In order to adapt to the changes in the physical environment and in the design of the system to carry out their intended activities, cyber-physical systems depend on good situation awareness. A cyber-physical system can modify its behavior in response to the real condition of the world as perceived by the cyber-physical system by being aware of the circumstances in the physical world. Understanding the state of the cyber-physical system allows for adaptation of the system's behavior in accordance with the current capabilities and state of the system, for example, by providing fewer features or features with limited functionality in the event that some of the system components are not functional. We must create systems that are capable of being robust in order to establish cyber-physical systems.

According to this paper [8], the Internet has swiftly advanced as a tool for global communication networks, enabling not only the sharing of information but also the completion of full tasks collaboratively using computational resources. Denial of service (DoS) attacks, in particular, are known to be challenging. Yet, over the past few decades, there has been a significant growth in criminal behaviors in networks, as well as in cunning and harmful content. The goal of the current study is to investigate the DoS flooding attack problem and make an effort to solve it by implementing classifiable countermeasures that stop, recognize, and react to DoS flooding attacks. Data were gathered from numerous online and offline sources using the study's secondary data gathering methodology. The study suggested a cyber-security TCP SYN, which is identified as the effective approach to reduce and block DDoS attacks. The application of these specific, cost-efficient defense mechanisms against these kinds of attacks allows business continuity to significantly improve their performance.

In this paper [9] In many different application sectors, sequence classification has been extensively used. Many different classification techniques are available, and they can be used using feature vectors. The difficulty of extracting feature vectors from sequences prevents these classification algorithms from being directly applied to the sequence classification problem. More specifically, the clustering problem in

sequences suffers from the "curse of dimensionality" because the characteristics in a sequence are sequential, making the task of classifying sequences more difficult than a conventional classification on feature vectors. We provide Seq2Image, a novel notion for converting sequences to images, in this research. Seq2Image is a straightforward but efficient technique for performing genomic sequence classification using convolutional neural networks (CNN). A given genomic sequence is first turned into a tensor, and the resulting tensor is subsequently converted into a picture. The produced images of the sequences are then classified using CNN deep learning-based image processing techniques. Our preliminary experimental study's results are quite encouraging, reaching 95.83% testing accuracy and 95.83% training accuracy for classifying 166 samples of human genomes from six different sequence families.

We describe many enhancements that enhance both the vector quality and training efficiency in this paper [10]. We significantly speed up the process and learn more regular word representations by subsampling the common words. Negative sampling, a straightforward substitute for the hierarchical SoftMax, is also described. The indifference to word order and inability to convey idiomatic expressions in word representations are intrinsic limitations. For instance, "Canada" and "Air" cannot simply be combined to become "Air Canada". By using this example as inspiration, we demonstrate a straightforward technique for detecting phrases in text and demonstrate that it is feasible to acquire effective vector representations for millions of phrases. High-quality distributed vector representations can be learned quickly and effectively using the skip-gram paradigm.

In this paper [11], we represent cyber-physical systems that are being attacked as linear time-invariant descriptor systems with unknowable inputs. This streamlined model disregards the existence of noise in the dynamics and measurements as well as system nonlinearities. To analyze stability, failures, and attacks in systems like electricity, sensor, and water networks, for example, such a simplified model has long been beneficial. We assume that the main findings of this work won't be much altered by more intricate models.

We examine a few of the emerging worries for CPS security in this paper [12]. We begin by outlining the necessity of creating adversary models for CPS. Then, we list some of the novel and fundamentally dissimilar issues that CPS presents in contrast to conventional IT security. In order to prevent, detect, respond to, survive, and deter computer attacks, we outline several potential topics for future research in the paper's conclusion. The author discusses three main obstacles to securing cyber-physical systems: (1) comprehending threats and potential effects of attacks, (2) identifying the special characteristics of these systems and how they differ from traditional IT security, and (3) discussing security measures appropriate for such systems. We examine security mechanisms in particular for prevention, detection, recovery, resilience, and deterrence.

The paper [13] suggests in this study that assessing the security of CPSs is presented. Using the suggested approach, one may examine the dynamic behavior of systems under security attacks and forecast the attacker's preferences for attacking CPSs. Attack tree structures are employed and parameterized with appropriate fuzzy data to address uncertainty in attacker decision-making to perform attacks. The model is then evaluated and attacker behavior is predicted using the fuzzy strategy for order of preference by similarity to ideal solution method. Moreover, the system's process model is used to explore the dynamic behavior of CPSs under attacks. The model's result is a relative assessment of the system's security level based on appropriate security metrics, such as the likelihood of attack scenarios, the amount of time the process can last before shutting down following an attack (time-to-shutdown), and security threats. By contrasting the strategy with another attack tree-based approach, we show its efficacy. We also estimate the defined quantitative security measures and provide two exemplary cases.

This paper [14] because the AMI (Advanced Metering Infrastructures) system receives and sends customers' billing and consumption data linked to fees and personal information, information security and system safety are crucial. Situation Awareness (SA) is a useful tool for maintaining system safety and security. Only physical or cyberspace was intended for the Conventional SA architecture. AMI, on the other hand, is a typical Cyber-Physical System (CPS) that should monitor both digital and physical assaults. A thorough SA framework for CPS is put

forward in this paper. A practical system case study with illustrations is offered. To validate the suggested framework, a simulation study is done. The reliability and safety of the energy system are two risks that are frequently present as technology advances. Failure of the system would likely be prohibitively expensive for everyone, and it would get more expensive over time.

This paper [15] In order to adapt to the changes in the physical environment and in the design of the system to carry out their intended activities, cyber-physical systems depend on good situational awareness. A cyber-physical system can modify its behavior in response to the real condition of the world as perceived by the cyber-physical system by being aware of the circumstances in the physical world. Understanding the state of the cyber-physical system allows for adaptation of the system's behavior in accordance with the current capabilities and state of the system, for example, by providing fewer features or features with limited functionality in the event that some of the system components are not functional. We must create systems that are capable of being robust in order to establish cyber-physical systems.

This paper [16] Cyber-physical systems (CPSs) are intricate systems that incorporate control, communication, and computing technology. CPSs are used widely today in smart grids, smart manufacturing, smart cities, and intelligent transportation. The integration of industrial control systems with contemporary communication technologies would, however, necessarily expose CPSs to greater security risks, which might seriously impair system performance or even result in CPS annihilation. This article provides an overview of recent developments in industrial cyber-physical system security (ICPSs). We specifically examine the Denial-of-Service (DoS) assault and the Deception attack, two common types of attacks, and give current findings regarding attack detection, estimation, and control of ICPSs. Based on various system modeling and analytic techniques, classifications of recent studies are analyzed and summarized. The benefits and drawbacks of different approaches are also highlighted. The report ends with a few possible routes for secure ICPS research in the future.

This paper [17] The new methods for tackling prediction issues in time series include machine and deep learning-based algorithms. It has been demonstrated that

using these strategies yields more accurate findings than using traditional regression-based modeling. Long Short-Term Memory (LSTM), a type of artificial Recurrent Neural Network (RNN) with memory, has reportedly been found to outperform Autoregressive Integrated Moving Average (ARIMA) by a significant margin. More "gates" are incorporated into the LSTM-based models in order to memorize longer input data sequences. The main query is whether the LSTM architecture's built-in gates already provide a strong prediction and whether further data training is required to further enhance the forecast.

This paper [18] It can be difficult to identify zero-day threats and vulnerabilities. For network managers, accurately identifying them is of highest importance. The protection system will be more effective the better the accuracy. In an ideal world (i.e., with 100% accuracy), the system may identify zero-day malware without worrying about incorrectly classifying innocent files as dangerous ones or allowing disruptive malicious code to run as innocent ones. The ability of several machine learning methods to recognize zero-day malware is examined in this article. 34 machine/deep learning classifiers were examined, and we discovered that the random forest classifier had the highest accuracy. In order to detect zero-day malware with zero rates for false positive and false negative, the study offers a number of research issues on how well machine learning and deep learning algorithms perform.

In this paper [19], we give a general overview of the system, outline its operational principle, and discuss the logistics of its deployment. We demonstrate the system's utility for unsupervised detection as well as the role that a human analyst may play in creating an active learning feedback loop. The analyst can improve the detection model that underlies their decision-making and significantly lower the false positive rate by using an accept or reject approach. We propose Corporate Insider Threat Detection (CITD), a multi-disciplinary research project that combines technical and behavioral activities to evaluate the threat posed by persons. CITD is an anomaly detection system. To determine the degree to which users stray from their reported behaviors, the system recognizes user and role-based profiles and measures deviations in order to determine the threat that a set of actions might represent.

In this paper [20], the authors discuss Vehicular Ad-hoc Networks (VANETs), in which moving automobiles communicate with one another to improve safety, comfort, and driving effectiveness. Particularly for applications that aim to improve traffic efficiency, such traffic information systems, the constrained wireless channel capacity is a problem. In these systems, a greater geographic region than required for traffic safety applications must be covered in order to swiftly distribute a high number of traffic or road status observations to interested vehicles, frequently via multi-hop forwarding. A practical tool to increase the scalability of such applications is in-network aggregation protocols. Nevertheless, from a security standpoint, they create additional entry points for insider attackers since vehicles collaborate to merge and modify messages while disseminating information.

TABLE 1 TABULAR SKETCH OF THE LITERATURE REVIEW

Reference	Algorithm/methodology	Advantages	Disadvantages
N. I. of Standard Technology. "Cyber-physical systems."	Meta-modeling technique.	Reduces the computational burden.	Time consuming.
] S. Walker-Roberts, M. Hammoudeh, O. Aldabbas, M. Aydin, and A. Dehghantanha, "Threats on the horizon: Understanding security threats in the era of	VERIS Community Database.	Can filter the dataset using the size.	Complex functionality.

cyber-physical systems,"			
M. N. Al-Mhiqani, R. Ahmad, W. Yassin, A. Hassan, Z. Z. Abidin, N. S. Ali, and K. H. Abdulkareem, "Cyber-security incidents: A review cases in cyber-physical systems,"	Statistical analysis of the collected data using a quantitative approach.	Provides accurate results.	Privacy concerns.
J.-P. A. Yaacoub, O. Salman, H. N. Noura, N. Kaaniche, A. Chehab, and M. Malli, "Cyber-physical systems security: Limitations, issues and future trends,"	IOT LED	Adaptable and less cost.	Low power.
R. S. Gutzwiller, S. M. Hunt, and D. S. Lange, "A task analysis toward characterizing cyber-cognitive situation awareness (ccsa) in cyber defense analysts,"	cognitive task analysis (CTA)	Increases scalability.	Time-intensive.
T. Debatty and W. Mees, "Building a cyber range for	CDSA	Flexibility in detection.	Low quality.

training cyberdefense situation awareness,"			
J. Preden, "Generating situation awareness in cyber-physical systems: Creation and exchange of situational information,"	Mediated interaction concept	Disregards time dependence.	Lack of confidentiality.
Q. V. Le and T. Mikolov, "Distributed representations of sentences and documents"	Skip-gram model	It is unsupervised learning hence can work on any raw text.	It cannot capture the polysemy.
S. Rao, "Denial of service attacks and mitigation techniques: Real time implementation with detailed analysis"	Prevention technique	Protection for data and network	High cost.
N. Tavakoli, "Sequence analysis using visualization and deep convolutional neural network"	Convolutional Neural Network (CNN)	Automatically detects the important features without any human supervision	Lots of training data is required.
Q. V. Le and T. Mikolov, "Distributed representations of sentences and documents"	Skip-gram model	It is unsupervised learning hence can work on any raw text.	It cannot capture the polysemy.
Pasqualetti, F., Dörfler, F.	Time-	Data is more	Less flexibility

and Bullo, F. (2013). "Attack detection and identification in cyber-physical systems"	invariant descriptor systems	protected	
A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems"	Network Dynamics	Accuracy in detection	Lack of trust
H. Orojloo and M. Abdollahi Azgomi, "Predicting the behavior of attackers and the consequences of attacks against cyber-physical systems"	Fuzzy technique	Manage vulnerability	Mistaken for likelihood hypothesis.
Z. Yang, T. Li, and W. Jiang, "Situation awareness for cyber-physical system: A case study of advanced metering infrastructure,"	Advanced Metering Infrastructure	Can report leaks or service issues.	Excessive monitoring.
Zhang, D., Wang, Q.G., Feng, G., Shi, Y. and Vasilakos, A.V. (2021). "A survey on attack detection, estimation and control of industrial cyber–physical systems"	ICPSs	ICP plasma include short analysis time	ICP plasma are spectral complexity

S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of LSTM and bilstm in forecasting time series,"	LSTM	Better at handling long-term dependencies	Require more training data in order to learn effectively.
F. Abri, S. Siami-Namini, M. A. Khanghah, F. M. Soltani, and A. S. Namin, "Can machine/deep learning classifiers detect zero-day malware with high accuracy?"	Super vector machine (svm)	Used to avoid the difficulties of using linear functions	Long training time for large datasets
Legg, P.A.; Buckley, O.; Goldsmith, M.; Creese, S. Caught in the act of an insider attack: Detection and assessment of insider threat.	Corporate Insider Threat Detection (CITD)	Known dangers, protected easily	Sell intellectual property, customer data
Dietzel, S.; Gürtler, J.; van der Heijden, R.; Kargl, F. Redundancy-based statistical analysis for insider attack detection in VANET aggregation schemes.	vehicular ad hoc networks (VANETs)	Prevention of collision, safety	High mobility

3. Proposed System Model enhanced

Dataset collection: If there are any missing values in the data, this could cause inconsistent results. If the data gathered contain missing values that could cause inconsistency, the model is trained using a custom dataset. The model is trained using a unique dataset made up of typical cybercrimes.

Preparing and processing a data set: If the data are missing values that could provide inconsistent results. Raw data will be prepared and used in the deep learning model during data preprocessing. For the purpose of creating a machine learning model, this is a crucial first step. The data is not always organized and clean while developing a machine learning project. As a result, a data pretreatment activity is carried out for this.

3.1 Proposed method

The user's data is stored in the dataset which is to be considered as the past data. The past data is pre-processed in the proposed system and the preprocessed dataset is compared with the existing Machine Learning algorithms. The predicted result is then tested for its accuracy. Generally, a cyberattack seeks to compromise the integrity of data or steal managed information, as well as to disrupt, disable, damage, or maliciously alter the computing environment or infrastructure employed by an organization. Cyberspace's present condition portends uncertainty for the future of the Internet and its expanding user base. Big data collected by device sensors reveals vast amounts of information that can be utilized for targeted attacks, and new paradigms pose new issues as a result. Even if several different existing techniques, models, and algorithms have been the basis for predictions of cyber-attacks, it is vital to take into account new models and algorithms that are based on data representations other than task-specific methodology. It can, however, be changed to learn the different data forms. The result shows the effectiveness of the proposed machine literacy algorithm fashion that can be compared with accuracy with entropy calculation.

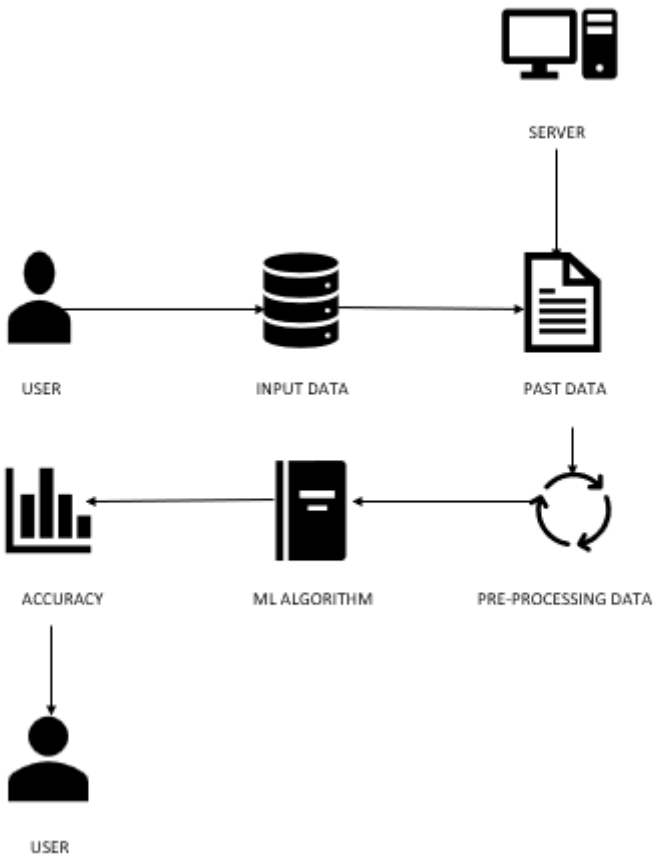

Fig.1 System Architecture

4. RESULT AND DISCUSSION

The study's objectives include accurate incident data analysis, crime prevention, and the capture of offenders. This paper's main focus is on using data analysis to draw conclusions about crime prevention. These findings will surface any hidden information and provide light on the investigations conducted by law enforcement officials. Machine learning techniques may be used to assess if the same culprit was responsible for the cyberattack based on the victim's information, the cyberattacks methodology, and whether the perpetrator has been identified or not. Over the years, many techniques have been used to determine the losses incurred by the victims of cyber events. Each victim's total financial losses were calculated by adding up their cumulative losses throughout the years. It is believed that the deterrence provided by the legislation and awareness campaigns is the cause of the decrease in such events, which was particularly evident after 2017. According to

Fig. 2, the level of economic damages brought on by cyberattacks is really substantial.

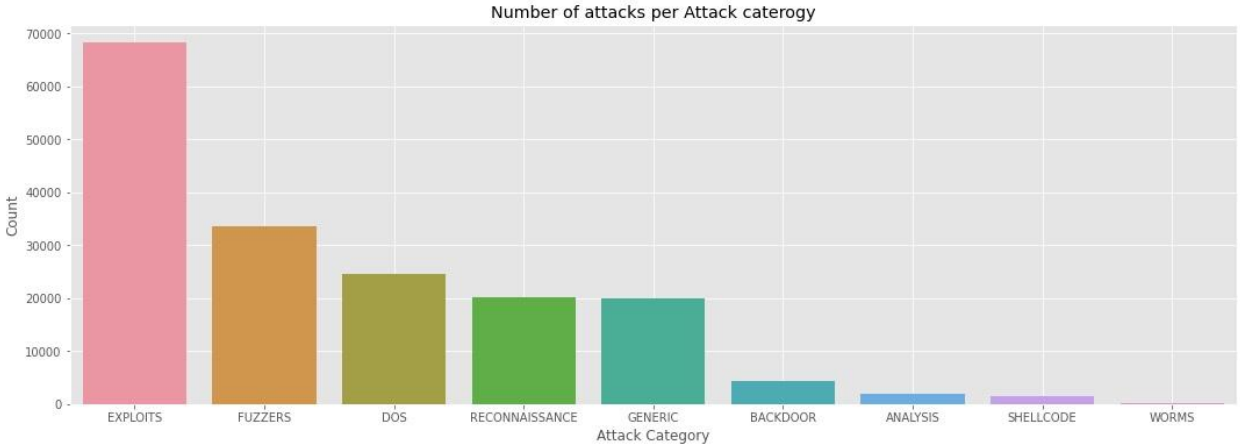

Fig.2 Damage caused by cyber-attacks

FREQUENCY OF ATTACKS

It is clear from the other forecasts that the information on the crime's kind, timing, perpetrator, demographics, and regional characteristics is mostly used as a feature. In our study, predictions are based on start and finish times, source and destination IP addresses, attack name, attack reference, source and destination ports, protocol, attack category, and attack subcategory. Many studies have forecasted the locations and times of future crimes. Many of them, however, have not discussed how crimes are committed, how to avoid them, or what characteristics the criminal exhibits. One of the study's main benefits is the use of actual data, which is a first step towards profiling individuals who share characteristics with the victims of attacks. Predicting the cyberattack method and if the offender can be identified is another benefit of the proposed study. Our findings indicate that any exposure to cybercrime decreases as wealth and educational attainment levels rise.

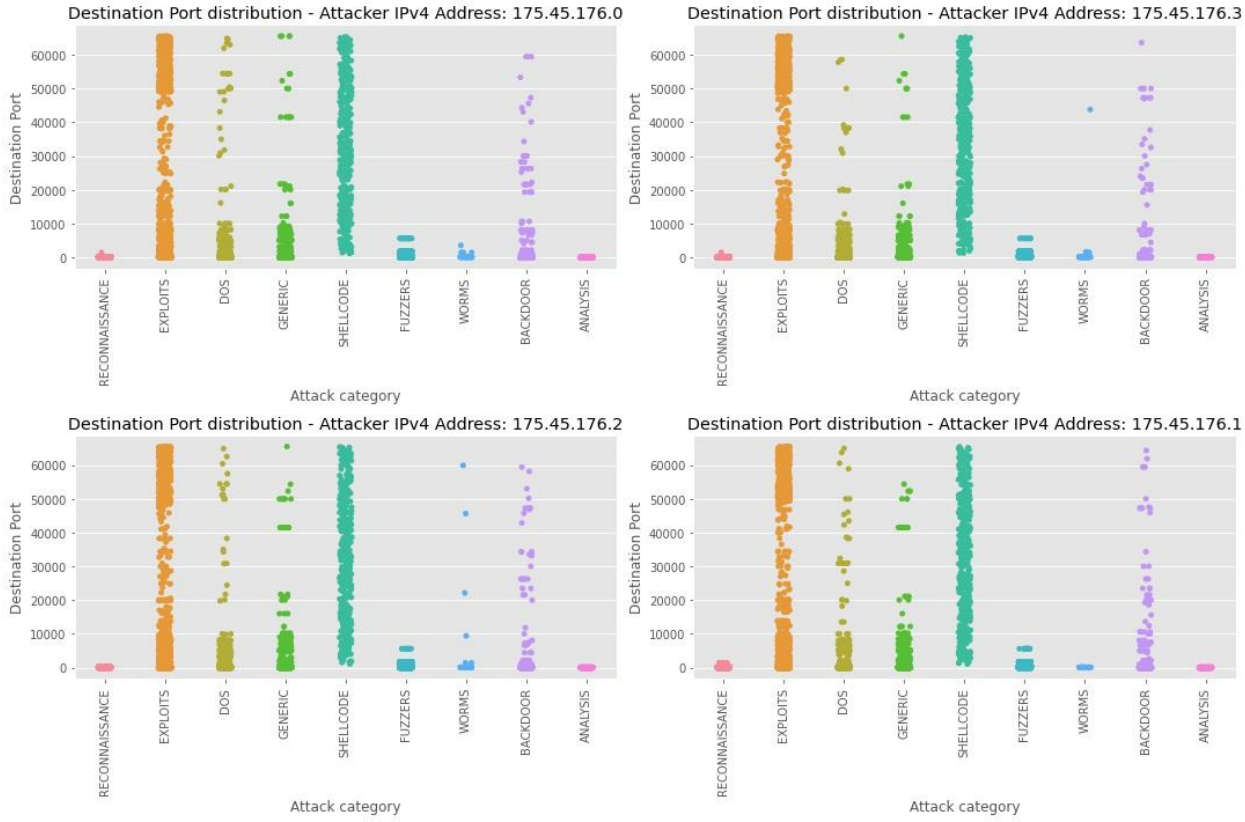

Fig 3 Frequency of attacks

5. CONCLUSION

A technique that uses data from past cybercrime instances along with machine learning algorithms to forecast and detect cyberattacks. The characteristics of those who might be assaulted and the types of attacks they would experience are predicted by the model. Machine learning techniques have shown to be effective enough. The most effective of these techniques is the linear SVMs approach. About 60% of the time, the model is successful in identifying the attacker who will launch a cyberattack. It might be possible to try to raise this ratio using other artificial intelligence techniques. According to our methodology, it is vital to call attention to social engineering and malware attacks in particular. It was shown that the likelihood of a cyberattack decreased as the victim's income and educational level increased. This study's main objective is to guide law enforcement organizations in the battle against cybercrime and to offer quicker and more

efficient options for tracking down criminals. The examination of the characteristics of the attack victims revealed in our analysis study can be used to develop new training and warning systems for people with comparable traits. Machine learning algorithms can be used to anticipate crime, criminal activity, victim profiling, and cyberattacks in future works, and the outcomes can be compared. Cybercrime data from other provinces may also be obtained based on discussions with other authorized entities with crime databases to be used for comparison with this study.

6. REFERENCES

[1] N. I. of Standard Technology. Cyber-physical systems. https://www.nist.gov/el/cyber-physical-systems.

[2] S. Walker-Roberts, M. Hammoudeh, O. Aldabbas, M. Aydin, and A. Dehghantanha, "Threats on the horizon: Understanding security threats in the era of cyber-physical systems," The Journal of Supercomputing, vol. 76, no. 4, pp. 2643–2664, 2020.

[3] M. N. Al-Mhiqani, R. Ahmad, W. Yassin, A. Hassan, Z. Z. Abidin, N. S. Ali, and K. H. Abdulkareem, "Cyber-security incidents: A review cases in cyber-physical systems," International Journal of Advanced Computer Science and Applications, vol. 9, 2018.

[4] D. Serpanos, "The cyber-physical systems revolution," Computer, vol. 51, no. 3, pp. 70–73, 2018.

[5] J.-P. A. Yaacoub, O. Salman, H. N. Noura, N. Kaaniche, A. Chehab, and M. Malli, "Cyber-physical systems security: Limitations, issues and future trends," Microprocessors and Microsystems, vol. 77, p. 103201, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0141933120303689

[6] M. R. Endsley, "Design and evaluation for situation awareness enhancement," Proceedings of the Human Factors Society Annual Meeting, vol. 32, no. 2, pp. 97–101, 1988. [Online]. Available: https://doi.org/10.1177/154193128803200221

[7] R. S. Gutzwiller, S. M. Hunt, and D. S. Lange, "A task analysis toward characterizing cyber-cognitive situation awareness (ccsa) in cyber defense analysts," in 2016 IEEE International Multi-Disciplinary Conference on

Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2016, pp. 14–20.

[8] T. Debatty and W. Mees, "Building a cyber range for training cyberdefense situation awareness," in 2019 International Conference on Military Communications and Information Systems (ICMCIS), 2019, pp. 1–6.

[9] J. Preden, "Generating situation awareness in cyber-physical systems: Creation and exchange of situational information," in 2014 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), 2014, pp. 1–3.

[10] Z. Yang, T. Li, and W. Jiang, "Situation awareness for cyber-physical system: A case study of advanced metering infrastructure," in 2018 IEEE International Conference on Prognostics and Health Management (ICPHM), 2018, pp. 1–6.

[11] H. Orojloo and M. Abdollahi Azgomi, "Predicting the behavior of attackers and the consequences of attacks against cyber-physical systems," Security and Communication Networks, vol. 9, no. 18, pp. 6111–6136, 2016. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/sec.1761

[12] A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems," in Workshop on Future Directions in Cyber-physical Systems Security. DHS, July 2009. [Online]. Available: http://chess.eecs.berkeley.edu/pubs/601.html

[13] C. D. Manning, H. Schutze, and P. Raghavan, ¨ Introduction to information retrieval. Cambridge university press, 2008.

[14] Q. V. Le and T. Mikolov, "Distributed representations of sentences and documents," CoRR, vol. abs/1405.4053, 2014. [Online]. Available: http://arxiv.org/abs/1405.4053

[15] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in Advances in Neural Information Processing Systems, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds., vol. 26. Curran Associates, Inc., 2013, pp. 3111–3119. [Online]. Available: https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf