INCREASING THE DETECTION AND PRECISION OF CYBER SECURITY OPERATIONS FOR IMPROVING DATA GATHERING AND MODEL RENEWAL SYSTEM USING MACHINE LEARNING TECHNOLOGY

[1]M.Sangeetha, [2]Dr. Ivsj Rama Rao, [3]R Uttam Sai, [4]G Jhansi
[1,2,3]Assistant Professor, [4]UG Student, [1,2,3,4]Department of Computer Science Engineering, Visvesvaraya College of Engineering & Technology, Hyderabad, India.

Abstract
In order to safeguard an organization's Internet security, security information and event management (SIEM) has been implemented to streamline the numerous preventative technologies and signal alerts for security problems. Inspectors of the SOC (Standards of Conduct) are looking into this (SOC). However, the bulk of the warnings are inaccurate, and SCO is unable to handle all of the information. Therefore, it may not be accurate to talk about malicious assaults and compromised hosts. Machine learning might be used to decrease the amount of false positives and enhance the output of SOC analysts. We have created a framework for engineer learning that is user-centered for the Internet Safety Functional Center. The usage of popular data sources, their operation, and how to transform that data into something usable for machine learning are all topics covered in SOC. People who should read this essay fall into one of two categories: Intelligent researchers who have no prior experience with data scientists or computer safety should be the first to develop machine learning methods for machine safety. Currently there are no machine learning experiences available for Internet security experts who have a deep grasp and proficiency in Cyber Security. The last example is a computer built in Seyondike's SOC manufacturing to show all aspects of data collection, label production, feature engineering, and a machine learning algorithm, as well as a sample evaluation of performance.
Keywords: Cyber Security, user-centered, SOC, Internet security, Machine learning etc

I Introduction

Cybercrime detection and response is the process of guarding electronic systems and networks, including computers, servers, mobile devices, and other digital assets, from invasions (CDCR). Cyber and security are two separate fields that may be further separated into two separate groups. Computers, networks, programmers, and any data or information kept on them are all referred to as "cyber" resources. Contrarily, security is concerned with safeguarding data, applications, networks, and systems. It may also be referred to as IT or electronic information security, depending on who you ask.Everything in our life today is digital, including the network, computers, cell phones, and cameras. All critical infrastructures, including the banking system, hospitals, financial institutions, governments, and the industrial industry need the use of devices connected to the Internet. Unauthorized access or exposure to their intellectual property, financial information,or personal data could have serious repercussions. For nefarious purposes like extortion, political meddling, or simply vandalism, these intruders and threat actors can use this information to breach the system.

The global economy could be threatened by cyber-attacks and other security breaches, which are becoming a global concern. Protecting sensitive data against well-publicized security breaches, therefore, necessitates an effective cyber security plan. When it comes to securing sensitive information, businesses and organisations must develop efficient cyber security measures and policies to defend themselves from the increasing volume of cyberattacks.

Cyber Security Goals

The basic purpose of cyber security is to protect data. The security community provides a triangle of three interconnected principles in order to protect data from cyber-attacks. The acronym for this concept is the CIA trio. In order to ensure that an organization's information security infrastructure is properly protected, the CIA model is used. In the event that a breach of security is discovered, at least one of these principles has been broken.

It is the CIA's three-pillared methodology that sets it apart from other intelligence agencies. It's a way of thinking about a wide range of IT security issues. Each section will be examined in greater detail below.

Fig1: Cybersecurity goals

Types of Cyber Security Threats

Data corruption or theft, gaining access to a network, or general disruption of digital life is allexamples of threats in the field of cyber security. The following are examples of current cyber risks, according to experts:

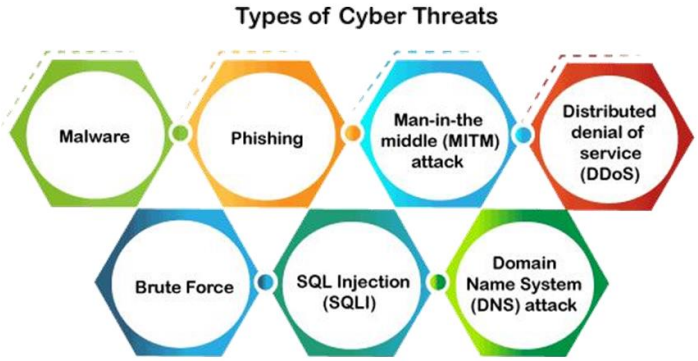


Fig2: Cyber Threat Types

Machine Learning

Generally, tasks in machine learning can be divided into broad groups. These classifications are based on the way in which the system designed receives or provides feedback on the learning it has received. There are two common ways to supervised machine learning:unsupervised learning and supervised learning, in which samples of input and output data are tagged by humans. Let's take a closer look at these strategies.

Supervised Learning: To train a computer, an instructor provides the computer with examples of inputs and outputs that are labelled with their anticipated outcomes. The algorithm can "learn" by comparing its actual output to the "learned" outputs and then changing the model as necessary using this method. Unlabeled data can be predicted using pattern recognition, which is the basis of supervised learning. Supervised learning can be used to feed images of sharks and waters into an algorithm. Using thisdata, an algorithm will be trained to recognise unlabeled shark and ocean photographsas fish and as water in the future. Sustained learning can be used to anticipate future events based on historical data. It can be used to detect impending movements in the stock market or to weed out spam emails. Untagged photos of dogs can be used as input data in supervised learning to classify the tagged images.

Unsupervised Learning: Learning algorithms are not given labels for their incoming data in unsupervised learning. In the absence of labelled data, machine learningsystems that can learn from unlabeled data are very useful. If you're looking for patterns in your data, unsupervised learning can help you find them. But if you're looking for the representations needed to classify your data, unsupervised learning canhelp you find them too. For transactional data, unsupervised learning is the most popular method. Customers and their purchases may be in your database, but as humans, it will be difficult to discern what commonalities can be inferred from their profiles and the purchases they make. For women in a given age range who use unscented soaps, an unsupervised learning algorithm may be used to determine whether or not they are pregnant. This market might therefore be specifically targeted by a marketing campaign for pregnancy and baby products.

Existing System

There has been a lack of attention paid to end-user security in most enterprise security techniques. Consequently, typical security functions and accompanying devices such as firewalls and intrusion detection systems (IDS) are primarily concerned with protecting the network level In light of the additional security problems discussed in the preceding section,

such an approach has its limits.

In order to stop or swiftly identify harmful behaviour, Data Analysis for Network Cyber- Security monitors and analyses network traffic data. In order to undertake a comprehensive risk assessment, risk values were added to an ISMS and quantitative evaluation was carried out. Risk can be reduced to a certain extent, as the quantitative evaluation revealed. The effectiveness of the planned countermeasures in terms of cost will need to be examined in thefuture. It includes information such as the type of attack, the frequency, the target host ID,and the source host ID for each assault. An attack tree-based methodology and mitigation measures were proposed by Ten et al. to protect the SCADA system as a vital infrastructure from cyber attacks.

Proposed System

By bringing security closer to end users, user-centric cyber security helps businesses mitigate the risk associated with rapidly changing end-user realities. There is a difference between user-centric cyber security and user security. Cyber security that is focused on the demandsof the user does so while still protecting the integrity of the company's network and its assets. It's easy to mistakenly believe that safeguarding the network against vulnerabilities introduced by users is the same as protecting the network itself from them. Enterprisesbenefit more from user-centric security. Independent, real-time systems with high performance requirements make up cyber security systems. In addition to criticalinfrastructures like the national power grid and transportation, they are also used in the medical and defence industries. It is only possible to meet the stringent demands of these applications if the computer, telecommunications and control technology systems are tightly integrated. Key infrastructures have long been a target for criminals because of their complexity and interconnection with other critical infrastructures. People, processes, technology, and other components of these CPSs can be attacked if risk management systems aren't in place or aren't working properly. Confidential information has been sought by the hackers. As part of this project, the major purpose is to remove unneeded data from the dataset

System Architecture

The following figure shows the system architecture of the proposed system.

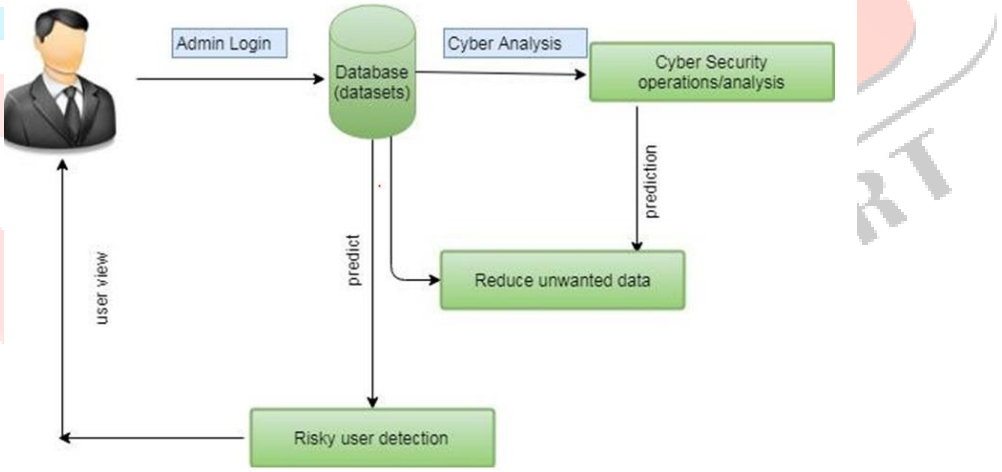


Fig3: System Architecture

Cyber Analysis

In a cyber-threat analysis, vulnerabilities in an organization's internal and external information are compared to actual cyber-attacks. A smooth transition from reactive security to proactive security is made possible by this threat-oriented approach to cyber attack defence.

Best practises for the implementation of protective mechanisms that maximise availability, confidentiality, and integrity while not compromising usability or functionality are also a goalof threat assessments. Analysis of Cyper. It could be anything that disrupts, interferes with or destroys any valued service or item in the company's repertory. A threat Regardless matter whether the source is "human" or "nonhuman," the study must evaluate every possiblesecurity risk.

Dataset Modification

The Datasets panel allows you to conceal individual dataset objects if your dashboard contains a lot of dataset objects. A good example of this is when importing vast amounts of data from a file, but not removing every undesirable data column before doing so, you can conceal the unwanted attributes and metrics. In the Datasets panel, click to make dataset items invisible. To make the Datasets panel's hidden objects visible, To changethe name of an object in a dataset, If you want

to measure an attribute, you can use a metric to do so To use a metric to construct an attribute, An attribute's geospatial role can be specified by defining the attribute's geolocation. In order to add more time-related datato an attribute, to replace a dashboard dataset object.

Data Reduction

Data redundancy, compression, snapshots, and thin provisioning can all be used to maximize storage capacity while reducing data usage. The most efficient technique to minimize storage's data is to simply delete undesired or unnecessary data.

Risky User Detection

To avoid customer shame, false alarm immunity is needed. Theft-detection efficiency thatis high enough to safeguard all kinds of commodities Having a wide range of exits allows you more options when it comes to the layout of the entrance and departure points. The variety of patterns available makes it easy to match any store's decor. System performance is maximised by the use of advanced digital controller technology.

System Design

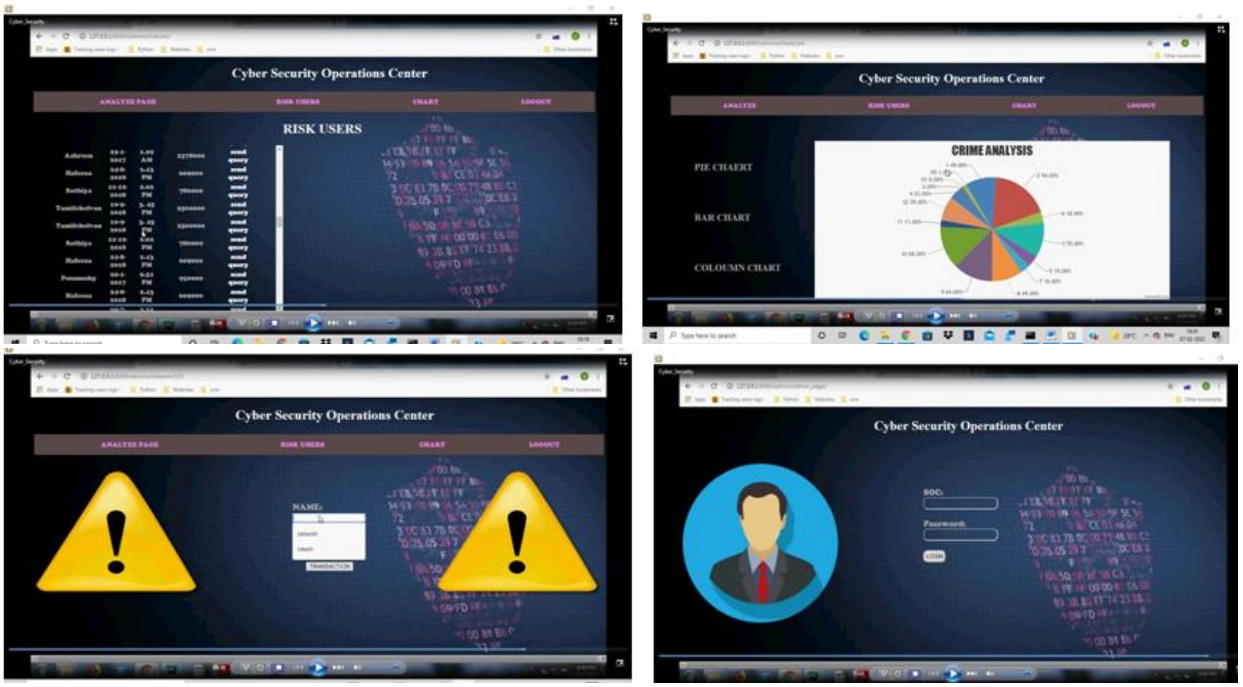
Requirement Analysis

To make the application more user-friendly, the project analysed the design of a few applications. As a result, it was critical to maintain a logical flow of movement from one screen to the next while also minimising the amount of input required of the user. The browser version had to be chosen so that it would be compatible with the majority of browsers in order to make the application more accessible.

Outputs

Analyze page

Receive alert page



Conclusion

Some of the services we provide include computer learning systems that are user-centered and draw on a variety of security logs, awareness data, and inspector intelligence. This allows for extensive customization and a solution for the detection of potentially dangerous users by the Enterprise System Operating Center. By assessing efficiency, IO, host, and users, machine learning techniques may be employed in the SOC product context to produce user-centric features. We show that the learning system can extract more insights from the labels that are most unbalanced and limited using straightforward mechanical techniques. The neurological model of modeling, which represents more than 20% of the total, accounts for more than a fifth of the current rule-based system. Other learning methods will be examined to improve data gathering and model renewal, real-time estimation, and risk management in order to increase the detection precision of the current scenario. Let's have a look at different methods of learning to increase detection accuracy in the future.

References

1. Aghajani, G.; Ghadimi, N. Multi-objective energy management in a micro-grid. Energy Rep. 2018, 4, 218–225.
2. Alhakami .W, ALharbi .A, Bourouis .S, Alroobaea .R, and Bouguila .N (2019),"Network anomaly intrusion detection using a nonparametric bayesian approach and feature selection", IEEE Access, Vol. 7, pp. 52181– 52190.
3. AlkeshBharati and Sarvanaguru RA (2018), "Crime Prediction and Analysis Using Machine Learning", International Research Journal of Engineering and Technology, Vol. 5, Issue. 9,pp. 1037 - 1042.
4. Al-Shaer E.S, Wei .J, Hamlen K.W, and Wang .C (2019), "Autonomous Cyber Deception: Reasoning, Adaptive Planning, and Evaluation of Honey Things", Springer, pp. 1 - 235.
5. Chaojun, G.; Jirutitijaroen, P.; Motani, M. Detecting false data injection attacks in ac state estimation. IEEE Trans. Smart Grid 2015, 6, 2476–2483
6. Daniel S. Berman, Anna L. Buczak, Jeffrey Chavis, Cherita L. Corbett (2019), "A Survey of Deep Learning Methods for Cyber Security", Information, Vol. 10, No. 122,pp. 1 - 35.
7. Dehghani, M.; Khooban, M.H.; Niknam, T.; Rafiei, S.M.R. Time-varying sliding mode control strategy for multibus low-voltage microgrids with parallel connected renewable power sources in islanding mode. J. Energy Eng. 2016, 142, 05016002.
8. Dehghani, M.; Khooban, M.H.; Niknam, T. Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations. Int. J. Electr. Power Energy Syst. 2016, 78, 455–462.
9. Ding, D.; Han, Q.-L.; Xiang, Y.; Ge, X.; Zhang, X.-M. A survey on security control and attack detection for industrial cyber-physical systems. Neurocomputing 2018, 275,1674–1683.
10. Esmalifalak, M.; Liu, L.; Nguyen, N.; Zheng, R.; Han, Z. Detecting stealthy false datainjection using machine learning in smart grid. IEEE Syst. J. 2014, 11, 1644–1652.

11. Foroutan, S.A.; Salmasi, F.R. Detection of false data injection attacks against state estimation in smart grids based on a mixture Gaussian distribution learning method. IET Cyber-Phys. Syst. Theory Appl. 2017, 2, 161–171.
12. Giraldo, J.; Urbina, D.; Cardenas, A.; Valente, J.; Faisal, M.; Ruths, J.; Tippenhauer, N.O.; Sandberg, H.; Candell, R. A survey of physics-based attack detection in cyber- physical systems. ACM Comput. Surv. 2018, 51, 1–36.
13. Habibi, M.R.; Baghaee, H.R.; Dragi˘cevi, T.; Blaabjerg, F. Detection of false data injection cyber-attacks in DC microgrids based on recurrent neural networks. IEEE J. Emerg. Sel. Top. Power Electron. 2020.
14. Hassen Mohammed Alsafi, Wafaa Mustafa Abduallah and Al-Sakib Khan Pathan (2019), "IDPS: An Integrated Intrusion Handling Model for Cloud ",Networking and International Architecture,arxiv.
15. He, Y.; Mendis, G.J.; Wei, J. Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. IEEE Trans. Smart Grid 2017, 8, 2505–2516.
16. K. Shaukat, S. Luo, S. Chen and D. Liu, "Cyber Threat Detection Using Machine Learning Techniques: A Performance Evaluation Perspective," 2020 InternationalConference on Cyber Warfare and Security (ICCWS), 2020, pp. 1-6, doi: 10.1109/ICCWS48432.2020.9292388.
17. Khraisat A, Gondal I, and Vamplew P (2019),"An anomaly intrusion detection systemusing C5 decision tree classifier", Trends and applications in knowledge discovery and data mining, Springer, pp 149–155.
18. Lakshmanaprabu .S.K, Shankar .K, Ilayaraja .M, Abdul Wahid Nasir, Vijayakumar Naveen Chilamkurti.V (2019), "Random forest for big data classification in the internet of things using optimal features", International Journal of Machine Learning and Cybernetics, Vol. 10,pp. 2609 - 2618.