Explorning The Role Of Machine Learning In Enhancing Cloud Security

Raguvijay SP[1], Srishyam S[2], Kishore R[3], Ms.M.Amshavalli[4]
[1, 2, 3, 4] Erode Sengunthar engineering college

Abstract- *The popularity and usage of Cloud computing is increasing rapidly. Several companies are investing in this field either for their own use or to provide it as a service for others. One of the results of Cloud development is the emergence of various security problems for both industry and consumer. One of the ways to secure Cloud is by using Machine Learning (ML). ML techniques have been used in various ways to prevent or detect attacks and security gaps on the Cloud. In this paper, we provide a Systematic Literature Review (SLR) of ML and Cloud security methodologies and techniques. We analyzed 63 relevant studies and the results of the SLR are categorized into three main research areas: (i) the different types of Cloud security threats, (ii) ML techniques used, and (iii) the performance outcomes.*

Moreover, distributed denial-of-service (DDoS) and data privacy are the most common Cloud security areas, with a 16% level of use and 14%respectively. On the other hand, we found 30 ML techniques used some used hybrid and others as standalone. The most popular ML used is SVM in both hybrid and standalone models. Furthermore, 60% of the papers compared their models with other models to prove the efficiency of their proposed model. Moreover, 13 different evaluation metrics were enumerated. The most applied metric is true positive rate and least used is training time. Lastly, from 20 datasets found, KDD and KDD CUP data set are the most used among relevant studies.

Keywords- Cloud security, machine learning, DDos, privacy, security

2.1 An Efficient DDoS TCP Flood Attack Detection and Prevention System in a Cloud Environment

Although the number of cloud projects has dramatically increased over the last few years, ensuring the availability and security of project data, services, and resources is still a crucial and challenging research issue. Distributed denial of service (DDoS) attacks are the second most prevalent cybercrime attacks after information theft. DDoS TCP flood attacks can exhaust the cloud's resources, consume most of its bandwidth, and damage an entire cloud project within a short period of time. The timely detection and prevention of such attacks in cloud projects are therefore vital, especially for eHealth clouds. In this paper, we present a new classifier system for detecting and preventing DDoS TCP flood attacks (CS_DDoS) in public clouds. The proposed CS_DDoS system offers a solution to securing stored records by classifying the incoming packets and making a decision based on the classification results. During the detection phase, the CS_DDOS identifies and determines whether a packet is normal or originates from an attacker.

2.2 A Systematic Literature Review on Cloud Computing Security: Threats and Mitigation Strategists

Cloud computing has become a widely exploited research area in academia and industry. Cloud computing benefits both cloud services providers (CSPs) and consumers. The security challenges associated with cloud computing have been widely studied in the literature. This systematic literature review (SLR) is aimed to review the existing research studies on cloud computing security, threats, and challenges. This SLR examined the research studies published between 2010 and 2020 within the popular digital libraries. We selected 80 papers after a meticulous screening of published works to answer the proposed research questions. The outcomes of this SLR reported seven major security threats to cloud computing services. The results showed that data tampering and leakage were among the highly discussed topics in the chosen literature.

2.3 Data Security and Privacy Protection for Cloud Storage A Survey

The new development trends including Internet of Things (IoT), smart city, enterprises digital transformation and world's digital economy are at the top of the tide. The continuous growth of data storage pressure drives the rapid development of the entire storage market on account of massive data generated. By providing data storage and management, cloud storage system becomes an indispensable part of the new era. Currently, the governments, enterprises and individual users are actively migrating their data to the cloud. Such a huge amount of data can create magnanimous wealth. However, this increases the possible risk, for instance,

unauthorized access, data leakage, sensitive information disclosure and privacy disclosure. Although there are some studies on data security and privacy protection, there is still a lack of systematic surveys on the subject in cloud storage system. In this paper, we make a comprehensive review of the literatures on data security and privacy issues, data encryption technology, and applicable countermeasures in cloud storage system.

2.4 MSCryptoNet: Multi-Scheme Privacy-Preserving Deep Learning in Cloud Computing

Privacy in the Internet of Things is a fundamental challenge for the Ubiquitous healthcare systems that depend on the data aggregated and collaborative deep learning among different parties. This paper proposes the MSCryptoNet, a novel framework that enables the scalable execution and the conversion of the state-of-the-art learned neural network to the MSCryptoNet models in the privacy-preservation setting.

2.5 Next-Generation Neural Networks: Capsule Networks With Routing-by-Agreement for Text Classification

These days, neural networks constantly prove their high capacity for nearly every application case and are considered as key technology for learning systems. However, neural networks need to continuously evolve for managing new arising challenges like increasing task complexity, explainability of decision making processes, expanded problem domains, providing resilient and robust systems etc. One possible enhancement of traditional neural networks constitutes the innovative Capsule Network (CapsNet) technology, which combines the expressiveness of distributed entity representations with an intelligent and interpretable signal propagation, named as routing-by-agreement. Since Caps Nets represent a relatively young acquirement, further research is essential for gaining profound knowledge about Caps Net theory and best practices for diverse application areas.

2.6 Cloud-Based Cyber-Physical Intrusion Detection for Vehicles Using Deep Learning

Detection of cyber attacks against vehicles is of growing interest. As vehicles typically afford limited processing resources, proposed solutions are rule-based or lightweight machine learning techniques. We argue that this limitation can be lifted with computational offloading commonly used for resource constrained mobile devices. The increased processing resources available in this manner allow access to more advanced techniques. Using as case study a small four-wheel robotic land vehicle, we demonstrate the practicality and benefits of offloading the continuous task of intrusion detection that is based on deep learning. This approach achieves high accuracy much more consistently than with standard machine learning techniques and is not limited to a single type of attack or the in-vehicle CAN bus as previous work. As input, it uses data captured in real-time that relate to both cybersex and physical processes, which it feeds as time series data to a neural network architecture. We use both a deep multi layer perceptron and recurrent neural network architecture, with the latter benefiting from a long-short term memory hidden layer, which proves very useful for learning the temporal context of different attacks.

2.7 Hyper band Tuned Deep Neural Network With Well Posed Stacked Sparse Auto Encoder for Detection of D Dos Attacks in Cloud

Cloud computing has very attractive features like elastic, on demand and fully managed computer system resources and services. However, due to its distributed and dynamic nature as well as vulnerabilities in virtualization implementation, the cloud environment is prone to various cuber-attacks and security issues related to cloud model. Some of them are inability to access data coming to and from cloud service, theft and misuse of data hosted, no control over sensitive data access, advance threats like malware injection attack, wrapping attacks, virtual machine escape, distributed denial of service attack (DDoS) etc. DDoS is one of the notorious attack. Despite a number of available potential solutions for the detection of DDoS attacks, the increasing frequency and potency of recent attacks and the constantly evolving attack vectors, necessitate the development of improved detection approaches. This article proposes a novel architecture that combines a well posed stacked sparse AutoEncoder (AE) for feature learning with a Deep Neural Network (DNN) for classification of network traffic into benign traffic and D Dos attack traffic.

2.8 Secure Data Storage and Sharing Techniques for Data Protection in Cloud Environments: A Systematic Review, Analysis, and Future Directions.

A large number of researchers, academia, government sectors, and business enterprises are adopting the cloud environment due to the least upfront capital investment, maximum scalability, and several other features of it. Despite the multiple features supported by the cloud environment, it also suffers several challenges. Data protection is the primary concern in the area of information security and cloud computing. Numerous solutions have been developed to address this challenge. This article presents a comparative and systematic study, and in-depth analysis of leading techniques

for secure sharing and protecting the data in the cloud environment.

2.9 A Virtual Machine Consolidation Algorithm Based on Ant Colony System and Extreme Learning Machine for Cloud Data Center

The energy consumption issue of large-scale data centers is attracting more and more attention. Virtual machine consolidation can significantly reduce energy consumption by migrating virtual machines from one physical machine to another. However, excessive virtual machine consolidation can lead to dangerous Service Level Agreement (SLA) violations. Therefore, how to balance between effective energy consumption and SLA violations avoidance effectively is a paradox to be mediated. The virtual machine consolidation problem is NP-hard. The traditional heuristic algorithm is easy to fall into the local optimal and some meta-heuristic algorithms can help to avoid it. However, the existing meta-heuristic algorithms are with high complexity.

2.10 Distributed Machine Learning Oriented Data Integrity Verification Scheme in Cloud Computing Environment

Distributed Machine Learning (DML) is one of the core technologies for Artificial Intelligence (AI). However, in the existing distributed machine learning framework, the data integrity is not taken into account. If network attackers forge the data, modify the data, or destroy the data, the training model in the distributed machine learning system will be greatly affected, and the training results are led to be wrong. Therefore, it is crucial to guarantee the data integrity in the DML. In this paper, we propose a distributed machine learning oriented data integrity verification scheme (DML-DIV) to ensure the integrity of training data. Firstly, we adopt the idea of Provable Data Possession (PDP) sampling auditing algorithm to achieve data integrity verification so that our DML-DIV scheme can resist forgery attacks and tampering attacks. Secondly, we generate a random number, namely blinding factor, and apply the discrete logarithm problem (DLP) to construct proof and ensure privacy protection in the TPA verification process.

2.11 COMPARATIVE TABLE

No	Authors	Title	Journal& Year	Methology	Demerits
1	AqeelSahi1,2,DavidLai2 YanLi2 , (Member, Ieee),And Mohammed Diykh1,2	An Efficient DDoS TCP Flood Attack Detection and Prevention System in a Cloud Environment	IEEE 2017	Data collection and preprocessing, Feature selection Attack detection and prevention.	The system is complex to implement and maintain. False positives
2	Bader Alouffi Muhammad Hasnain Abdullahalharbi,Wael Alosaimi Hashem Alyami1	A Systematic Literature Review Threats and Mitigation Strategies	IEEE 2021	Once the research have been identified, the next step is to identify the relevant literature. This by searching academic databases, such as Google Scholar, Scopus, and Web of Science.	Conducting an SLR consuming process. It can involve identifying and screening hundreds of studies, from the selected studies, and analyzing the data

3	Pan Yang 1 , Naixue Xiong2 , (Senior Member, Ieee), And Jingli Ren 1	Data Security and Privacy Protection for Cloud Storage: A Survey	IEEE 2020	• Developing more efficient and scalable data algorithm developing more effective access control mechanisms data auditing techniques	A survey typically existing literature and available data sources. It may lack access primary,firsthand data, could provide more detailed insights
4	Owusu-Agyemangkwabenazhen Qin Tianming Zhuang, And Zhiguang Qin	MSCryptoNet: Multi-Scheme Privacy-Preserving Deep Learning in Cloud Computing	IEEE 2019	The first step is to convert the neural network model to a homomorphic encryption-compatible form. This is done by approximating the activation functions of the neural network with low-degree polynomials. This can be done using a variety	MSCryptoNet can introduce significant computational overhead, especially when used to train and evaluate large neural networks on encrypted data. ..
5	Nikolai a. k. Steur and Friedhelm Schwenker , (member, ieee)	Next-Generation Neural Networks: Capsule Networks With Routing-by-Agreement for Text Classification	IEEE 2021	• Capsule networks are able to learn hierarchical relationships between entities in data, which is important for text classification tasks. • Capsule networks are more robust to noise and variations in the data than traditional neural networks.	Capsule networks can be computationally train, especially for large datasets. This is because agreement mechanism requires iteratively updating the activation vectors of all the capsules in the network.
6	George Loukas, Tuan Vuong, Ryan a Sakellar Yongpil Yoon	Cloud-Based Cyber-Physical Intrusion Detection for Vehicles Using Deep Learning	IEEE 2018	Cloud-based cyber-physical intrusion detection (CPID) from cyber attacks. anomalies in vehicle that may indicate cyber attack.	Vehicle data collected CPID confidential. Important properly unauthorized access
7	Aanshi Bhardwaj, Veenu Mangat , (member, ieee), and Renu Vig	Hyper band Tuned Deep Neural Network With Well Posed Stacked	IEEE 2020	This may involve cleaning the data, removing outliers, and converting the data to a format that is compatible with the deep learning	Ensemble methods often create complex models that can be difficult to interpret
8		Secure Data Storage and	IEEE 2022	. A variety of search engines and databases can be used, such	The selection of papers for inclusion

	Ishu Gupta Ashutosh Kumar Singh Chung-Nan Lee And Rajkumar Buyya	Sharing Techniques for Data Protection in Cloud Environments: A Systematic Review, Analysis, Future Directions		as Pub Med, Google Scholar, and Scopus.	in the review, as well as the analysis and interpretation of the findings, is based on the judgment of the reviewers. This subjectivity can introduce bias into the review.
9	Fagui liu, Zhenjiang ma , bin wang , and Weiwei lin	A Virtual Machine Consolidation Algorithm Based on Ant Colony System and Extreme Learning Machine for Cloud Data Center	IEEE 2019	This information includes the resource requirements of VMs, the capacity of PMs, and the current state of PMs (e.g., idle, running, overloaded). The next step is to generate candidate migration plans. A candidate migration plan is a set of VM migrations that can be performed to consolidate the VMs a smaller number of PMs	• The ACS-ELM algorithm is a complex algorithm and can be computationally expensive to implement. • Hyper parameters: The ACS-ELM algorithm has several hyper parameters that need to be tuned to achieve the best results.
10	Xiao-ping zhao and Rui Jiang	Distributed Machine Learning Oriented Data Integrity Verification Scheme in Cloud Computing Environment	IEEE 2020	The next step is to generate a data signature for each share. The data signature is a cryptographic hash of the shard that can be used to verify the integrity of the shard.	The DML-DIV scheme generates a challenge for each shard periodically, which can lead to significant communication overhead.

III. CONCLUSION

We carried out a systematic literature review to analyze ML techniques used in Cloud security. The review investigated relevant studies that answered 3 RQs; Cloud security area, type of ML techniques used, and the accuracy estimation of the ML model. Overall, we obtained 60 research papers after applying our selection criteria.

Moreover, we noticed that little work has been done using deep learning techniques in cloud security. We encourage researchers to take advantage of the deep learning in this regard.

REFERENCES

[1] OWUSU-AGYEMANG. KWABENA, ZHEN QIN, TIANMING ZHUANG, AND ZHIGUANG QIN MSCryptoNet: Multi-Scheme Privacy-Preserving Deep Learning in Cloud Computing, publication February 25, 2019,

[2] NIKOLAI A. K. STEUR AND FRIEDHELM SCHWENKER, (Member, and IEEE) Next-Generation Neural Networks: Capsule Networks with Routing-by-Agreement for Text Classification publication September 7, 2021
[3] GEORGE LOUKAS, TUAN VUONG, RYAN HEARTFIELD, GEORGIA SAKELLARI, YONGPIL YOON, AND DIANE GAN Cloud-Based Cyber-Physical Intrusion Detection for Vehicles Using Deep Learning Publication December 11, 2017
[4] AANSHI BHARDWAJ, VEENU MANGAT, (Member, IEEE), AND RENU VIG Hyper band Tuned Deep Neural Network with Well Posed Stacked Sparse Auto Encoder for Detection of DDoS Attacks in Cloud Publication October 5, 2020
[5] ISHU GUPTA, ASHUTOSH KUMAR SINGH , CHUNG-NAN LEE1 , AND RAJKUMAR BUYYA , Secure Data Storage and Sharing Techniques for Data Protection in Cloud Environments: A Systematic Review, Analysis, and Future Directions publication 4 July 2022
[6] FAGUI LIU, ZHENJIANG MA, BIN WANG , AND WEIWEI LIN A Virtual Machine Consolidation Algorithm Based on Ant Colony System and Extreme Learning Machine for Cloud Data Center publication December 23, 2019
[7] XIAO-PING ZHAO AND RUI JIANG Distributed Machine Learning Oriented Data Integrity Verification Scheme in Cloud Computing Environment publication February 4, 2020
[8] ÁLVARO LÓPEZ GARCÍA, JESÚS MARCO DE LUCAS, MARICA ANTONACCI, WOLFGANG ZU CASTELL, MARIO DAVID , MARCUS HARDT, LARA LLORET IGLESIAS, GERMÁN MOLTÓ, MARCIN PLOCIENNIK , VIET TRAN A Cloud-Based Framework for Machine Learning Workloads and Applications publication January 6, 2020
[9] MUHAMMAD MEHMOOD, RASHID AMIN, MUHANA MAGBOUL ALI MUSLAM, (Member, IEEE), JIANG XIE, AND HAMZA ALDABBAS Privilege Escalation Attack Detection and Mitigation in Cloud Using Machine Learning publication 8 May 2023
[10] KAUSHIK SEKARAN, MOHAMMAD S. KHAN, RIZWAN PATAN , AMIR H. GANDOMI, PARIMALA VENKATA KRISHNA, AND SURESH KALLAM Improving the Response Time of M-Learning and Cloud Computing Environments Using a Dominant Firefly Approach publication February 12, 2019