International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 November 2023*

Assessing Privacy-Preserved Federated Learning for Enhanced Cyber-Attack Detection in Edge-Based Iot Systems

Jitendra Singh Dodiya, Dr. Vijayalaxmi Biradar

Kalinga University, Naya Raipur, Chhattisgarh, India

ABSTRACT

A critical challenge is the equilibrium between harnessing the potential advantages of IoT and guaranteeing strong security and privacy for consumers. Intelligent Edge Computing (IEC) emerges as a crucial answer, providing a transformative approach to data processing and security. This research presents a privacy-preserving federated learning (FL) methodology for detecting cyber-attacks in an edge-based IoT ecosystem. A unique lightweight convolutional Transformer (LCT) network is developed to accurately detect cyber-attacks by learning attack patterns from IoT traffic on local edge devices, with the model customized by fine-tuning. We assess our proposed methodology using a real-world dataset of network traffic (NSL-KDD) that encompasses many attack types, and the experimental findings indicate that our customized federated learning technique surpasses conventional federated learning. Our technique is demonstrated to be successful in managing non-stationary data and adjusting to alterations in the network environment.

Keywords: Edge Computing, Data, Privacy, Training, Precision

I. INTRODUCTION

Instead of depending entirely on centralized cloud-based systems, the Internet of Things (IoT) may take use of Intelligent Edge Computing, a revolutionary method that moves computing capacity closer to the data source. This is achieved by deploying processing capabilities at the network's periphery. Traditional cloud-based IoT systems have a number of serious flaws that this decentralization fixes. Important for real-time applications like smart grids, industrial automation, and autonomous cars, Intelligent Edge Computing saves bandwidth utilization, increases reaction times, and decreases latency by processing data locally.

Moving to the edge offers many advantages in terms of privacy and security. The danger of interception, manipulation, and illegal access is increased since data transmission in traditional cloud-based systems generally involves large networks. Computing near the network's edge reduces the quantity of sensitive data that must travel across potentially unsecured networks, which in turn reduces the likelihood of these threats. Stronger and more situationally relevant security measures are made possible by this localized processing, which also enables finer-grained control over data access and use.

The overall security posture of IoT systems may be enhanced by integrating sophisticated security techniques at the edge. For instance, edge devices may provide safe key management and encryption using hardware-based security features like Trusted Platform Modules (TPMs) and Hardware Security Modules (HSMs). To further emphasize the importance of edge computing in recognizing and reacting to security issues quickly, it enables the adoption of threat mitigation measures and real-time anomaly detection.

Intelligent edge computing also successfully handles privacy issues related to the internet of things. Minimizing the danger of revealing personal or proprietary data is possible by keeping data processing local. This allows for the anonymization or aggregation of sensitive information before it is transferred to central servers. The CCPA and the General Data Protection Regulation (GDPR) are two examples of data protection laws that stress the need of limiting data collection and obtaining user agreement; this method is in line with both laws. With edge computing, data gathering procedures may be more precisely controlled, guaranteeing that only essential information is shared and processed in accordance with legal obligations.

Intelligent edge computing also encourages a security architecture that is more robust and flexible. It is possible to program edge devices such that they may run autonomously,

updating and implementing security rules without involving the main system. The distributed nature of the system makes it more resilient to assaults and failures; in the event that connection with the main server is interrupted, each edge node can keep protecting its local environment and continue to operate. Further increasing privacy and security is the use of federated learning methods at the edge, which enable collaborative construction of machine learning models without disclosing raw data.

New privacy and security measures are being made available by mixing edge computing with AI and blockchain, two examples of emerging technology. Blockchain technology's immutable records and decentralized trust mechanisms guarantee data authenticity and integrity, while AI algorithms installed at the edge can constantly scan data for any security risks. An all-encompassing answer to the complicated problems of data management and protection in an ever-more-connected world, this combination of edge computing with modern technologies strengthens the security posture of IoT devices.

II. REVIEW OF LITERATURE

Abdel-Basset, Mohamed et al., (2022) When it comes to the digital transformation of traditional sectors into Industry 4.0, the Industrial Internet of Things (IIoT) is crucial. Data availability, enhanced analytics, and autonomous control are made possible via the IIoT by connecting various industrial devices such as sensors, actuators, appliances, and more. A variety of covert and ever-changing cyberattacks pose a serious danger to the reliability and safety of IIoT systems, made worse by their complex dispersed nature. This undermines the reliability of IIoT as a defence against cyberattacks, rendering conventional security measures useless. To address this issue, this paper introduces a BoEI framework that use blockchain technology and incorporates a novel decentralised federated learning method known as Fed-Trust. The goal of this framework is to identify cyberattacks in IIoT. A temporal convolutional generative network is used in the Fed-Trust to provide semi-supervised learning using semi-labeled data. BoEI incorporates a reputation-based blockchain to facilitate the decentralisation of transaction recording and verification, hence ensuring the privacy and security of data and gradients. To improve Fed-Trust's computation and communication speed, fog computing is used to offload the block mining process from the edge side. The Fed-Trust outperforms state-of-the-art cyberattack detection methods in proof-of-concept simulations conducted on two publicly available datasets.

Regan, Christopher et al., (2022) Devices that connect to the Internet of Things (IoT) are mass-produced and made available to the public in a short amount of time. Their intended uses range from environmental monitoring to on-demand electrical switches, among others. These Internet of Things devices tend to be diverse, get updates infrequently, and may 'hide' on a home or workplace network for long stretches of time. So, two major (research and operational) issues with IoT systems are privacy and security. Traditional threat detection methods have long struggled to identify potential dangers to Internet of Things devices, such as malware-based assaults and botnets. While many obstacles persist, there have been new efforts to develop deep learning-based solutions to address the shortcomings of traditional detection methods. In order to identify botnet assaults utilising decentralised traffic data collected from devices, this study suggests a federated-based method that makes use of a deep autoencoder. By preventing data transfers or relocations away from the network edge, the proposed federated method safeguards user privacy. The advantage of moving machine learning processing to the edge layer, where data is created, is that it improves data security. The results show that our suggested model can train using features such source IP, MAC-IP, and destination IP, etc., and obtain an anomaly detection accuracy rate of up to 98%. Our suggested decentralised method outperforms the centralised structure in terms of overall performance, particularly when it comes to accurately detecting attacks.

Pal, Geeta & Taqa, Amer. (2022) By doing computations close to the data source rather than in the cloud, edge computing helps keep costs down. In addition to boosting efficiency and cutting costs, processing data at the edge eliminates the need to store or transmit it. By rerouting data and traffic away from a single server, edge computing further improves data and traffic security. Instead of storing sensitive information in the cloud or another potentially vulnerable location, companies can process data locally, at the edge. Since the communications between the IoT device and the edge server are more difficult to detect and exploit, edge computing is also supposed to be more secure than cloud computing. Reduced latency, enhanced scalability and flexibility, enhanced security, and cost savings are the main advantages of edge computing. Edge computing eliminates the need for central servers in the cloud, allowing organisations to process data more effectively and obtain faster replies. Because no data transfers are required, organisations may save money on cloud computing expenses and take advantage of edge computing's improved security features. Two technologies that are playing an increasingly significant role in the current digital economy are the Internet of Things (IoT) and edge computing.

Mehmood, M. et al., (2021) The smart grid is an innovative paradigm of the traditional electricity system, aimed at

incorporating sustainable and renewable technology. The smart grid (SG) has emerged as a prominent study subject due to advancements in technologies such as the Internet of Things (IoT), edge computing, artificial intelligence, big data, and 5G. The efficiency of SG will be enhanced by intelligent embedded devices with decision-making capabilities. Diverse sensor kinds and data sources will gather high-resolution data. A significant problem for IoT is the management of the substantial data generated by sensors. Transmitting this substantial volume of data straight to the cloud would result in issues related to latency, security, privacy, and excessive bandwidth use. This concern is mitigated by edge computing (EC). In edge computing, data is processed at the periphery of the network, close to the embedded devices. This study presents a thorough assessment of smart grid technologies, grounded on IoT and EC. The article emphasises advancements in emerging technologies, the framework for EC-IoT-based smart grids, and the prerequisites for implementing the EC-IoT-based smart grid system. The framework for an EC-IoT-based smart grid is analysed, and essential prerequisites for implementing the EC-IoT-based smart grid system are delineated. Ultimately, some significant concerns and obstacles encountered in the deployment of EC-IoT-based smart grid systems are delineated. Several significant unresolved research difficulties are also mentioned.

Lu, Yunlong et al., (2020) Recent advancements in technologies like MEC and AI substantially enhance the acceleration of VCPS implementation. Methods such as dynamic content caching, optimal resource allocation, and data exchange are essential for improving service quality and user experience. Simultaneously, data leakage in VCPS might result in tangible repercussions, including jeopardising passenger safety and privacy, as well as inflicting significant property damage on data suppliers. The escalating data volume, fluctuating network architecture, and constrained resources render data leakage in VCPS a more formidable issue, particularly when it encompasses many users and transmission channels. This article presents a safe and intelligent architecture aimed at improving data privacy. We introduce our novel privacy-preserving federated learning technique and propose a two-phase mitigation system including intelligent data transformation and collaborative data leakage detection. Empirical findings derived from a real-world dataset illustrate the efficacy of our suggested strategy, indicating that it attains commendable accuracy, efficiency, and elevated security.

Dai, Minghui et al., (2020) The proliferation of IoT devices, along with advances in cloud and edge computing, is spurring the creation of new sectors and technologies. With this new integrated paradigm, the Internet of Things' (IoT) service capabilities might grow exponentially. But, due to their susceptibility to hostile assaults, edge-driven intelligent IoT devices pose significant security risks. For smart Internet of Things (IoT) systems, this article explores an edge-driven security architecture. To begin, we provide an overview of the architecture of edge-driven intelligent IoT and go over several common uses for this technology. Secondly, we highlight the dangers to edge-driven intelligent IoT security from the perspective of malicious attacker actions. Third, we take a look at intelligent IoT at the edge and create a case study about its security. Lastly, we go over several potential avenues for further study in this new field.

Shirisha, B. et al., (2018) With the proliferation of IoT (Internet of Things) applications, cloud computing is facing several difficulties. Low spectral efficiency (SE), non-adaptive machine communication, and excessive latency are the obstacles. The new computer paradigm known as "edge computing"—which requires relocating data to the periphery of the network—has emerged as a potential solution to these problems. Edge computing has the potential to address issues related to battery life, reaction speed, data security, bandwidth cost savings, privacy, and more. To bring the idea of edge computing to life, this article begins with an overview of the Internet of Things (IoT) and its answer, edge computing. It then moves on to present several case examples, covering topics such as cloud offloading, smart homes and cities, and collaborative edge.

III. EXPERIMENTAL SETUP

Materials

The NSL-KDD dataset is utilised for training and assessment purposes in order to conduct experiments with the suggested model. One of the files in the dataset is designated for training (KDDTrain+), while the other two are designated for testing. Each one is comprised of a total of 43 characteristics. Denial of service (DoS) assaults, normal traffic, probe traffic, remote to user (U2R) attacks, and root to local (R2L) attacks are the five primary types of traffic that are represented in the data. The data comprise a variety of other classes of communications. All of the different sets of NSL-KDD data are summarised in Table 1, which contains the class distribution.

Table 1: Class distribution in training and testing sets of NSL-KDD data

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 11
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 November 2023*

Class label	Normal	Dos	Probing	R2L	U2R
Training Set (%)	53.3	36.48	9.21	0.82	0.05
Test Set (%)	53.07	36.59	9.51	0.90	0.02

Training of the suggested model is carried out on a Lenovo workstation that is outfitted with 64 gigabytes of memory and a central processing unit (CPU) that is an Intel® CoreTM i9-12900K. The Nvidia GeForce RTX 2080 is utilised for the purpose of expediting the process of training. Both the training and the creation of the model are coded in TensorFlow version 2.8. During the training phase, the AdamW algorithm is utilised for parameter optimisation over a period of sixty epochs. The initial learning rate is set at 0.001, and the batch size is set at 1024. Within our system, the number of communication cycles has been set to a total of 25.

Evaluation

To assess the efficacy of the presented models, the metrics employed in this study are specified as follows:

$$Precision = \frac{TP}{\mathrm{TP} + \mathrm{FP}}$$

$$Recall = \frac{TP}{\mathrm{TP} + \mathrm{FN}}$$

$$F1 - measure = 2 * \frac{\mathrm{Recall} \times \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}}$$

IV. RESULTS AND DISCUSSION

Table 2 presents the class-level performance of the proposed federated learning system compared to centralised training.

Table 2: Class-Level Performance of the Proposed FL Framework vs. Centralized Training

	Centralized			Proposed		
Methods	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
Normal	99.41	99.18	99.27	97.67	97.82	97.78
Dos	99.32	98.89	99.13	97.69	96.91	97.29
Probing	95.77	98.12	96.92	88.94	90.49	89.71
R2L	99.08	99.06	99.13	93.90	96.38	95.09
U2R	92.26	99.91	96.10	92.28	99.94	96.26

Recall and F1-score are rather similar, with the FL framework showing somewhat lower values (97.78) than the centralised technique (99.27), although the centralised method shows somewhat greater accuracy (99.41) than the suggested FL framework (97.67) for the Normal class. This suggests that both approaches are capable of accurately identifying typical traffic, although the centralised technique is somewhat more precise.

In terms of accuracy (99.32 vs. 97.69) and F1-score (99.13 vs. 97.29), the centralised method performs better than the FL framework when it comes to detecting and categorising Dos assaults. Even while it's not as noticeable, the centralised strategy also has better recall.

In comparison to the FL framework, which achieved a precision of 88.94, recall of 90.49, and F1-score of 89.71 for the Probing class, the centralised technique shows better performance across all metrics with a recall of 98.12 and an F1-score of 96.92. It seems that the centralised method works better for identifying and categorising probing attempts.

There is a clear distinction in the R2L class; the centralised technique achieves a high F1-score of 99.13 thanks to its 99.08 accuracy and 99.06 recall. On the other hand, the FL

framework's F1-score is 95.09, thanks to its lower accuracy (93.90) and greater recall (96.38). With respect to accuracy and total F1-score, the centralised approach stands head and shoulders above the competition in this class.

While the centralised approach has a little superior recall (99.91) and equivalent accuracy (92.28 vs. 92.26) and F1-score (96.26 vs. 96.10) for the U2R class, the FL framework shows comparable performance. This proves that FL framework can detect U2R assaults just as well as its competitors.

V. CONCLUSION

According to our findings, edge-based IoT systems can identify cyberattacks much better when using the privacy-preserved federated learning (FL) method. Our technology successfully detects various cyber-attack patterns in IoT data across local edge devices using a new lightweight convolutional Transformer network. This guarantees strong privacy protection and allows us to personalize our models. Our FL framework surpasses conventional centralized training methods in experimental evaluations conducted on the NSL-KDD dataset. This is especially true when it comes to handling non-stationary data and adjusting to changing network conditions. The findings demonstrate that our strategy successfully addresses real-world IoT security concerns by balancing privacy protection with high detection accuracy. All things considered, our research supports the idea that cyber-attack detection systems might benefit from federated learning methodologies implemented in edge computing settings.

REFERENCES: -

[1] M. Abdel-Basset, N. Moustafa, and H. Hawash, "Privacy-Preserved Cyberattack Detection in Industrial Edge of Things (IEoT): A Blockchain-Orchestrated Federated Learning Approach," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 1-1, Nov. 2022, doi: 10.1109/TII.2022.3167663.

[2] C. Regan, M. Nasajpour, R. Parizi, S. Pouriyeh, A. Dehghantanha, and K.-K. Choo, "Federated IoT security attack detection using decentralized edge data," *Machine Learning with Applications*, vol. 8, no. 02012, pp. 1-11, Dec. 2022, doi: 10.1016/j.mlwa.2022.100263.

[3] Q. Minh, V.-H. Nguyen, Q. Vu Khanh, L. Ngoc, A. Chehri, and G. Jeon, "Edge Computing for IoT-Enabled Smart Grid: The Future of Energy," *Energies*, vol. 15, no. 17, pp. 6140, Sep. 2022, doi: 10.3390/en15176140.

[4] E. Fazeldehkordi and T.-M. Grønli, "A Survey of Security Architectures for Edge Computing-Based IoT," *IoT*, vol. 3, no. 3, pp. 332-365, Jul. 2022, doi: 10.3390/iot3030019.

[5] G. Pal and A. Taqa, "Challenges and Future Research Directions in Next Generation Edge Computing and IoT," *International Journal of New Media Studies*, vol. 9, no. 2, pp. 2394-4331, Jun. 2022.

[6] M. Mehmood, A. Oad, M. Abrar, H. Munir, S. Hasan, K. A. Muqeet, and N. A. Golilarz, "Edge Computing for IoT-Enabled Smart Grid," *Security and Communication Networks*, vol. 2021, pp. 1-16, Apr. 2021, doi: 10.1155/2021/5524025.

[7] Z. Lyu, D. Chen, R. Lou, and Q. Wang, "Intelligent edge computing based on machine learning for smart city," *Future Generation Computer Systems*, vol. 115, pp. 90-99, Jan. 2021, doi: 10.1016/j.future.2020.08.037.

[8] M. Laroui, B. Nour, H. Moungla, A. Moussa, H. Afifi, and M. Guizani, "Edge and fog computing for IoT: A survey on current research activities & future directions," *Computer Communications*, vol. 180, pp. 1-12, Sep. 2021, doi: 10.1016/j.comcom.2021.09.003.

[9] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Federated Learning for Data Privacy Preservation in Vehicular Cyber-Physical Systems," *IEEE Network*, vol. 34, pp. 50-56, Jan. 2020, doi: 10.1109/MNET.011.1900317.

[10] M. Dai, Z. Su, R. Li, Y. Wang, J. Ni, and D. Fang, "An Edge-Driven Security Framework for Intelligent Internet of Things," *IEEE Network*, vol. 34, pp. 39-45, Jan. 2020, doi: 10.1109/MNET.011.2000068.

[11] M. Mukherjee, R. Matam, C. Mavromoustakis, H. Jiang, G. Mastorakis, and M. Guo, "Intelligent Edge Computing: Security and Privacy Challenges," *IEEE Communications Magazine*, vol. 58, no. 9, pp. 26-31, Sep. 2020, doi: 10.1109/MCOM.001.2000297.

[12] K. Sha, T. Yang, W. Wei, and S. Davari, "A survey of edge computing based designs for IoT security," *Digital Communications and Networks*, vol. 6, no. 2, pp. 1-8, Jun. 2019, doi: 10.1016/j.dcan.2019.08.006.

[13] B. Shirisha, D. D. Priya, M. R. K. Mahalakshmi, M. R. Chilukala, and B. Pravallika, "Emerging paradigm of edge computing in the context of IoT," *International Journal of Engineering and Technology (UAE)*, vol. 7, no. 3, pp. 122-125, Jun. 2018, doi: 10.14419/ijet.v7i3.3.14503.