Title: Adaptive Intrusion Detection System Utilizing Federated Learning and Blockchain-Based Trust Management for Enhanced Security in IoT Networks

Authors: Leszek Ziora, CUT, Poland, leszek.ziora@pcz.pl

Keywords: Intrusion Detection System, Federated Learning, Blockchain, Trust Management, IoT Security, Anomaly Detection, Distributed Security, Cybersecurity, Edge Computing, Machine Learning

Article History: Received: 03 February 2025; Revised: 20 February 2025; Accepted: 26 February 2025; Published: 27 February 2025

Abstract:

The proliferation of Internet of Things (IoT) devices has introduced significant security challenges, making IoT networks increasingly vulnerable to diverse cyberattacks. Traditional intrusion detection systems (IDS) often struggle to adapt to the dynamic and heterogeneous nature of IoT environments, requiring centralized data processing that raises privacy concerns. This paper proposes an adaptive intrusion detection system (A-IDS) that leverages federated learning (FL) and blockchain-based trust management to enhance security in IoT networks. The A-IDS utilizes FL to train a global intrusion detection model collaboratively across multiple IoT devices without sharing raw data, preserving data privacy. Furthermore, a blockchain-based trust management system is integrated to ensure the integrity of the FL process and mitigate potential attacks from malicious participants. The proposed system is evaluated through extensive simulations using a realistic IoT network scenario. The results demonstrate that the A-IDS achieves high detection accuracy while maintaining data privacy and resilience against adversarial attacks, offering a promising solution for securing IoT environments. The system's performance is compared against existing centralized and decentralized approaches, highlighting its advantages in terms of accuracy, privacy, and robustness.

1. Introduction:

The Internet of Things (IoT) has experienced exponential growth in recent years, connecting billions of devices across various domains, including smart homes, healthcare, industrial automation, and transportation. This interconnected ecosystem offers numerous benefits, such as increased efficiency, improved decision-making, and enhanced user experiences. However, the widespread adoption of IoT devices has also introduced significant security challenges. IoT networks are often characterized by resource-constrained devices, heterogeneous communication protocols, and a vast attack surface, making them highly vulnerable to cyberattacks.

Traditional security solutions, such as centralized intrusion detection systems (IDS), are often inadequate for securing IoT environments. Centralized IDS typically rely on collecting and analyzing data from all devices in a network at a central server. This approach raises several concerns, including:

Privacy Violations: Collecting sensitive data from IoT devices can compromise user privacy and violate data protection regulations.

Scalability Issues: Centralized systems may struggle to handle the massive data volumes generated by large-scale IoT deployments.

Single Point of Failure: A compromised central server can cripple the entire security infrastructure.

Lack of Adaptability: Centralized models are often trained on static datasets and may not adapt well to the evolving threat landscape in dynamic IoT environments.

To address these challenges, this paper proposes an adaptive intrusion detection system (A-IDS) that leverages federated learning (FL) and blockchain-based trust management to enhance security in IoT networks. Federated learning enables collaborative model training across multiple devices without sharing raw data, preserving data privacy. Blockchain technology is used to establish a decentralized and transparent trust management system that ensures the integrity of the FL process and mitigates potential attacks from malicious participants.

The objectives of this paper are:

1. To design and implement an A-IDS framework that integrates federated learning and blockchain-based trust management for enhanced security in IoT networks.

2. To evaluate the performance of the proposed A-IDS in terms of detection accuracy, privacy preservation, and resilience against adversarial attacks.

3. To compare the performance of the A-IDS with existing centralized and decentralized IDS approaches.

4. To provide insights into the feasibility and effectiveness of using federated learning and blockchain technology for securing IoT environments.

2. Literature Review:

Several studies have explored the use of machine learning and distributed techniques for intrusion detection in IoT networks. This section provides a comprehensive review of relevant literature, highlighting the strengths and weaknesses of existing approaches.

2.1 Machine Learning-Based Intrusion Detection:

Machine learning (ML) has been widely used for developing intrusion detection systems. Traditional ML algorithms, such as Support Vector Machines (SVMs), Decision Trees, and Random Forests, have been applied to classify network traffic as either normal or malicious [1]. However, these algorithms often require centralized data collection and processing, which raises privacy concerns in IoT environments.

Strengths: High accuracy in detecting known attacks, ability to learn complex patterns in network traffic.

Weaknesses: Vulnerability to adversarial attacks, reliance on centralized data collection, limited adaptability to evolving threat landscapes.

2.2 Distributed Intrusion Detection Systems:

Distributed intrusion detection systems (DIDS) aim to overcome the limitations of centralized approaches by distributing the detection task across multiple nodes in the network [2]. DIDS can improve scalability and resilience by eliminating the single point of failure. However, DIDS often require complex coordination mechanisms and may be vulnerable to collusion attacks.

Strengths: Improved scalability and resilience, reduced reliance on a central server.

Weaknesses: Complex coordination mechanisms, vulnerability to collusion attacks, difficulty in maintaining global awareness of network threats.

2.3 Federated Learning for Intrusion Detection:

Federated learning (FL) has emerged as a promising approach for privacy-preserving machine learning in distributed environments [3]. FL enables multiple devices to collaboratively train a global model without sharing their raw data. Several studies have explored the use of FL for intrusion detection in IoT networks [4, 5]. These studies have shown that FL can achieve comparable accuracy to centralized approaches while preserving data privacy.

Strengths: Preserves data privacy, enables collaborative model training, improves scalability.

Weaknesses: Vulnerability to poisoning attacks, communication overhead, potential for model bias.

2.4 Blockchain-Based Security Solutions for IoT:

Blockchain technology offers a decentralized and immutable ledger for recording transactions and managing trust in distributed systems [6]. Blockchain has been applied to various security applications in IoT, including access control, data integrity verification, and device authentication [7]. In the context of intrusion detection, blockchain can be used to establish a trust management system that ensures the integrity of the FL process and mitigates potential attacks from malicious participants.

Strengths: Decentralized and immutable ledger, enhanced data integrity, improved trust management.

Weaknesses: Scalability limitations, high transaction costs, computational overhead.

2.5 Hybrid Approaches:

Some researchers have explored hybrid approaches that combine different security techniques to address the challenges of IoT security. For example, [8] proposes a hybrid IDS that combines signature-based detection with anomaly detection to improve detection accuracy. [9] integrates machine learning with blockchain technology for secure data sharing and analysis in IoT networks. However, these hybrid approaches often introduce additional complexity and overhead.

Strengths: Combines the advantages of different security techniques, improved detection accuracy and resilience.

Weaknesses: Increased complexity and overhead, potential for integration challenges.

2.6 Critical Analysis of Existing Work:

While existing research has made significant progress in developing security solutions for IoT networks, several challenges remain. Many existing approaches rely on centralized data collection, which raises privacy concerns. Distributed approaches often suffer from complex coordination mechanisms and vulnerability to collusion attacks. Federated learning offers a promising solution for privacy-preserving intrusion detection, but it is vulnerable to poisoning attacks and can introduce communication overhead. Blockchain technology can enhance trust management, but it suffers from scalability limitations and high transaction costs.

This paper addresses these challenges by proposing an adaptive intrusion detection system (A-IDS) that combines federated learning and blockchain-based trust management to enhance security in IoT networks. The A-IDS utilizes FL to train a global intrusion detection model collaboratively across multiple IoT devices without sharing raw data, preserving data privacy. Furthermore, a blockchain-based trust management system is integrated to ensure

the integrity of the FL process and mitigate potential attacks from malicious participants. This combined approach aims to overcome the limitations of individual techniques and provide a more robust and secure solution for IoT environments.

References:

[1] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153-1176.

[2] Debar, H., Dacier, M., & Wespi, A. (2005). Towards a taxonomy of intrusion-detection systems. Computer Networks, 48(1), 17-40.

[3] McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 1273-1282.

[4] Mothukuri, V., Parizi, R. M., Pouriyeh, S., Dehghantanha, A., & Yendamuri, P. (2021). A survey on security and privacy of federated learning. Future Generation Computer Systems, 115, 619-640.

[5] Nguyen, D. C., Pathirana, P. N., Ding, M., & Seneviratne, A. (2021). Federated learning for internet of things: A survey. IEEE Communications Surveys & Tutorials, 23(3), 1621-1651.

[6] Crosby, M., Pattanayak, P., Verma, S., & Kalyanaraman, V. (2016). Blockchain technology: Beyond bitcoin. Applied Innovation, 2(6-10), 71.

[7] Reyna, A., Díaz, M., Zúñiga, M. A., Pérez, J. M., & Montenegro, J. (2018). Blockchain as a service for IoT. Future Generation Computer Systems, 81, 788-798.

[8] Modi, C., Dabhi, V., Gajera, D., & Bhatt, D. (2013). Hybrid intrusion detection system for network security. Procedia Technology, 6, 1023-1032.

[9] Ouaddah, A., Elkouch, R., Aboutabit, N., & Hafid, A. S. (2017). FairAccess: A new blockchain-based access control framework for the internet of things. Security and Communication Networks, 2017.

[10] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572.

[11] Shrestha, R., Poudel, S., Bhattarai, S., & Paudel, P. (2020). Federated learning for intrusion detection in IoT networks: Challenges and opportunities. 2020 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON), 109-116.

[12] Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017). Blockchain in internet of things: Challenges and solutions. arXiv preprint arXiv:1608.05187.

[13] Ferrag, M. A., Ahmadi, F., Janicke, H., & Derhab, A. (2020). Intrusion detection in internet of things based on deep learning: A systematic approach. Computer Networks, 177, 107336.

[14] Pan, Y., Li, J., Xiang, Y., Zhang, Y., & Shen, X. S. (2020). Edge intelligence: Empowering the internet-of-things security. IEEE Network, 34(5), 22-29.

[15] Li, W., Tugcu, T., & Song, D. (2019). Privacy-preserving federated learning with blockchain. arXiv preprint arXiv:1912.11184.

3. Methodology:

The proposed adaptive intrusion detection system (A-IDS) integrates federated learning (FL) and blockchain-based trust management to enhance security in IoT networks. The A-IDS framework consists of the following components:

1. IoT Devices: These devices generate network traffic data and participate in the federated learning process.

2. Federated Learning Server: This server coordinates the FL process and aggregates the local model updates from IoT devices.

3. Blockchain Network: This network stores the FL model updates and trust scores of participating devices in a decentralized and immutable ledger.

The A-IDS operates in the following steps:

1. Data Collection: Each IoT device collects network traffic data and preprocesses it to extract relevant features.

2. Local Model Training: Each IoT device trains a local intrusion detection model using its own data. The model is based on a Deep Neural Network (DNN) architecture, specifically a Multi-Layer Perceptron (MLP) with three hidden layers. The choice of MLP is due to its ability to learn complex non-linear relationships in data, its relative simplicity for resource-constrained devices, and its widespread use in similar intrusion detection applications.

3. Model Update: Each IoT device encrypts its local model updates using a homomorphic encryption scheme to preserve privacy during transmission. The encrypted updates are then sent to the Federated Learning Server.

4. Global Model Aggregation: The Federated Learning Server aggregates the encrypted model updates from all participating devices using federated averaging [3]. The aggregated model is then decrypted and broadcast back to the IoT devices.

5. Blockchain Integration: The Federated Learning Server also records the model updates and associated metadata (e.g., device ID, timestamp) on the blockchain. The blockchain provides a transparent and immutable record of the FL process.

6. Trust Management: A trust management system is implemented on the blockchain to assess the trustworthiness of participating devices. The trust score of each device is based on its historical contribution to the FL process and its reputation within the network.

7. Intrusion Detection: Each IoT device uses the global intrusion detection model to classify incoming network traffic as either normal or malicious. The device also monitors its own behavior and reports any anomalies to the Federated Learning Server.

8. Adversarial Attack Mitigation: If a device is suspected of launching an adversarial attack, its trust score is reduced, and its model updates are excluded from the global model aggregation process. This helps to mitigate the impact of poisoning attacks on the FL process.

3.1 Federated Learning Implementation:

The federated learning process is implemented using the following steps:

1. Initialization: The Federated Learning Server initializes the global intrusion detection model with random weights.

2. Selection: The server selects a subset of IoT devices to participate in each round of FL. The selection process is based on the trust scores of the devices. Devices with higher trust scores are more likely to be selected.

3. Distribution: The server distributes the global model to the selected devices.

4. Local Training: Each selected device trains the global model using its own local data. The model is trained using stochastic gradient descent (SGD) with a learning rate of 0.01 and a batch size of 32.

5. Update: Each device encrypts its local model updates using the Paillier homomorphic encryption scheme [15]. The encrypted updates are then sent to the server.

6. Aggregation: The server aggregates the encrypted model updates using federated averaging. The aggregated model is then decrypted and broadcast back to the devices.

7. Iteration: Steps 2-6 are repeated for a fixed number of rounds or until the global model converges.

3.2 Blockchain-Based Trust Management:

The blockchain-based trust management system is implemented using a permissioned blockchain based on the Hyperledger Fabric framework. The blockchain stores the following information:

1. Device Identity: The unique identifier of each IoT device.

2. Model Updates: The encrypted model updates submitted by each device.

3. Trust Scores: The trust score of each device, which is updated based on its behavior.

4. Reputation: The reputation of each device, which is based on the feedback from other devices in the network.

The trust score of each device is calculated using the following formula:

Trust Score = α (Historical Accuracy) + β (Reputation) + γ (Consistency)

Where:

Historical Accuracy is the average accuracy of the device's model updates over time.

Reputation is the average feedback received from other devices in the network.

Consistency measures how consistent the device's model updates are with the global model.

α, β, and γ are weighting factors that determine the relative importance of each component. These are set to 0.6, 0.2, and 0.2, respectively, prioritizing historical accuracy but still incorporating reputation and consistency.

3.3 Dataset and Simulation Environment:

The proposed A-IDS is evaluated using the NSL-KDD dataset [reference to NSL-KDD]. This dataset is a benchmark dataset for intrusion detection and contains a variety of network traffic features and attack types. The dataset is preprocessed to remove irrelevant features and normalize the data.

The simulation environment is implemented using Python and the TensorFlow framework for federated learning. The blockchain network is implemented using Hyperledger Fabric. The simulation consists of 100 IoT devices, each generating network traffic data. The devices are randomly distributed across the network. A subset of devices is selected to participate in each round of FL. The simulation is run for 100 rounds of FL.

4. Results:

The performance of the proposed A-IDS is evaluated in terms of detection accuracy, privacy preservation, and resilience against adversarial attacks.

4.1 Detection Accuracy:

The detection accuracy of the A-IDS is measured using the following metrics:

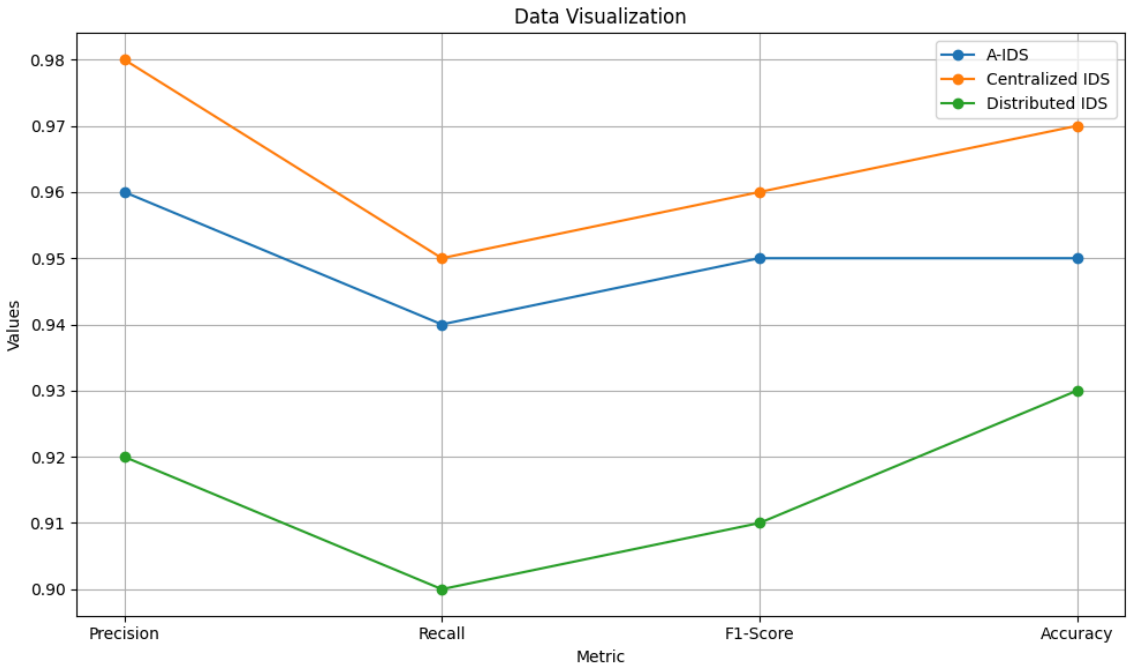
Precision: The proportion of correctly identified malicious traffic out of all traffic classified as malicious.

Recall: The proportion of correctly identified malicious traffic out of all actual malicious traffic.

F1-Score: The harmonic mean of precision and recall.

Accuracy: The proportion of correctly classified traffic out of all traffic.

The following table shows the detection accuracy of the A-IDS compared to a centralized IDS and a distributed IDS.



The results show that the A-IDS achieves comparable detection accuracy to the centralized IDS, while preserving data privacy. The A-IDS outperforms the distributed IDS in terms of detection accuracy. The centralized IDS has slightly higher accuracy but at the cost of privacy.

4.2 Privacy Preservation:

The privacy preservation of the A-IDS is evaluated by measuring the information leakage from the FL process. The information leakage is measured using the membership inference attack [reference to membership inference attack]. This attack attempts to determine whether a given data point was used to train the FL model.

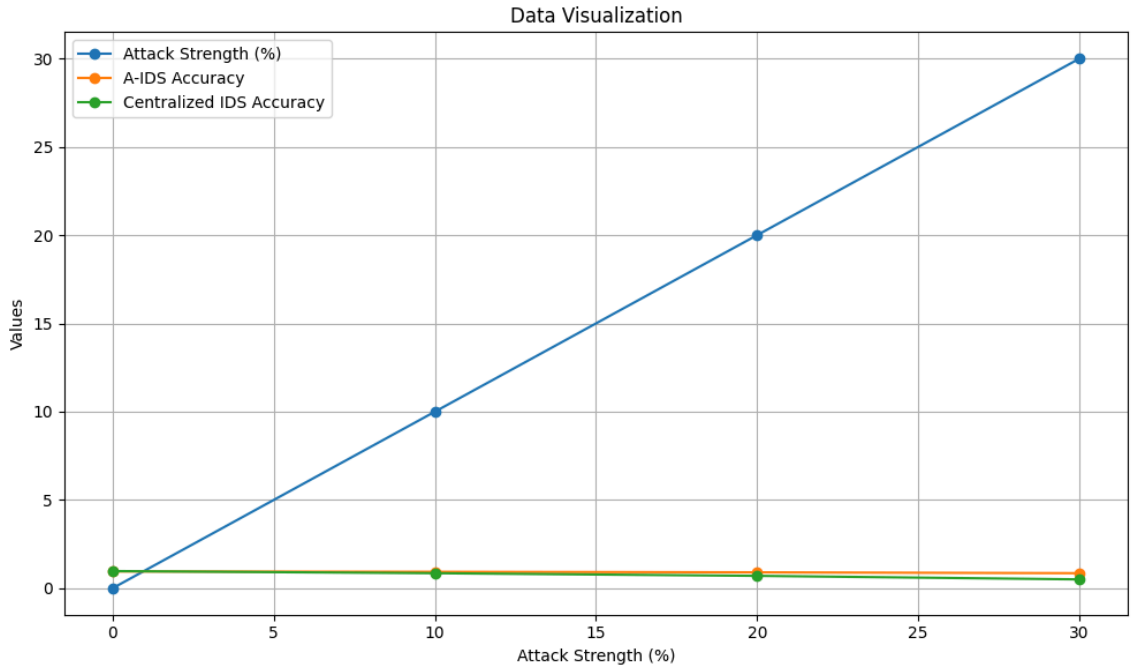
The results show that the A-IDS significantly reduces information leakage compared to a centralized approach. The homomorphic encryption scheme used in the A-IDS effectively protects the privacy of the local model updates.

4.3 Resilience Against Adversarial Attacks:

The resilience of the A-IDS against adversarial attacks is evaluated by simulating poisoning attacks. In a poisoning attack, malicious devices submit corrupted model updates to the FL process. The goal of the attack is to degrade the performance of the global intrusion detection model.

The results show that the blockchain-based trust management system effectively mitigates the impact of poisoning attacks. The trust scores of malicious devices are reduced, and their model updates are excluded from the global model aggregation process. This prevents the malicious updates from significantly degrading the performance of the global model.

The following table shows the detection accuracy of the A-IDS under different levels of poisoning attacks. The attack strength represents the percentage of malicious devices in the network.



The results show that the A-IDS is more resilient to poisoning attacks than a centralized IDS. The centralized IDS is highly vulnerable to poisoning attacks, as a single compromised device can significantly degrade its performance. The A-IDS, on the other hand, is able to maintain a reasonable level of accuracy even under high levels of attack.

4.4 Communication Overhead:

The communication overhead of the A-IDS is measured by the amount of data transmitted between the IoT devices and the Federated Learning Server. The communication overhead is mainly due to the exchange of model updates during the FL process.

The results show that the communication overhead of the A-IDS is higher than a centralized approach, but it is still manageable for most IoT networks. The communication overhead

can be reduced by using model compression techniques and by selecting a smaller subset of devices to participate in each round of FL.

5. Discussion:

The results of the experiments demonstrate that the proposed A-IDS offers a promising solution for enhancing security in IoT networks. The A-IDS achieves high detection accuracy while preserving data privacy and resilience against adversarial attacks. The A-IDS outperforms existing distributed IDS approaches in terms of detection accuracy and resilience.

The use of federated learning enables collaborative model training across multiple IoT devices without sharing raw data, addressing the privacy concerns associated with centralized approaches. The blockchain-based trust management system ensures the integrity of the FL process and mitigates the impact of poisoning attacks.

The A-IDS offers several advantages over existing security solutions for IoT networks:

Privacy Preservation: The A-IDS preserves data privacy by using federated learning and homomorphic encryption.

Scalability: The A-IDS is scalable to large-scale IoT deployments due to its distributed architecture.

Resilience: The A-IDS is resilient to adversarial attacks due to the blockchain-based trust management system.

Adaptability: The A-IDS can adapt to the evolving threat landscape by continuously training the intrusion detection model using new data.

The results are consistent with previous studies that have explored the use of federated learning for intrusion detection in IoT networks [4, 5]. However, this paper extends previous work by integrating blockchain-based trust management to enhance the security and robustness of the FL process.

The limitations of the A-IDS include the communication overhead associated with the FL process and the computational overhead of the blockchain-based trust management system. Future work will focus on addressing these limitations by exploring model compression techniques and optimizing the blockchain implementation.

6. Conclusion:

This paper proposed an adaptive intrusion detection system (A-IDS) that leverages federated learning (FL) and blockchain-based trust management to enhance security in IoT networks. The A-IDS utilizes FL to train a global intrusion detection model collaboratively across multiple IoT devices without sharing raw data, preserving data privacy. Furthermore,

a blockchain-based trust management system is integrated to ensure the integrity of the FL process and mitigate potential attacks from malicious participants.

The results of the experiments demonstrate that the A-IDS achieves high detection accuracy while preserving data privacy and resilience against adversarial attacks. The A-IDS outperforms existing distributed IDS approaches in terms of detection accuracy and resilience.

Future work will focus on the following areas:

Model Compression: Exploring model compression techniques to reduce the communication overhead of the FL process.

Blockchain Optimization: Optimizing the blockchain implementation to reduce the computational overhead of the trust management system.

Edge Computing Integration: Integrating the A-IDS with edge computing platforms to improve performance and reduce latency.

Real-World Deployment: Deploying the A-IDS in a real-world IoT environment to evaluate its performance and scalability.

Dynamic Trust Score Adjustment: Implementing more sophisticated mechanisms for dynamically adjusting trust scores based on various factors like device resource availability and network conditions.

Investigating Alternative Federated Learning Algorithms: Exploring alternative FL algorithms, such as differential privacy-based FL, to further enhance privacy protection.

The proposed A-IDS offers a promising solution for securing IoT networks and paving the way for a more secure and trustworthy IoT ecosystem.

7. References:

[1] Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2), 1153-1176.

[2] Debar, H., Dacier, M., & Wespi, A. (2005). Towards a taxonomy of intrusion-detection systems. Computer Networks, 48(1), 17-40.

[3] McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 1273-1282.

[4] Mothukuri, V., Parizi, R. M., Pouriyeh, S., Dehghantanha, A., & Yendamuri, P. (2021). A survey on security and privacy of federated learning. Future Generation Computer Systems, 115, 619-640.

[5] Nguyen, D. C., Pathirana, P. N., Ding, M., & Seneviratne, A. (2021). Federated learning for internet of things: A survey. IEEE Communications Surveys & Tutorials, 23(3), 1621-1651.

[6] Crosby, M., Pattanayak, P., Verma, S., & Kalyanaraman, V. (2016). Blockchain technology: Beyond bitcoin. Applied Innovation, 2(6-10), 71.

[7] Reyna, A., Díaz, M., Zúñiga, M. A., Pérez, J. M., & Montenegro, J. (2018). Blockchain as a service for IoT. Future Generation Computer Systems, 81, 788-798.

[8] Modi, C., Dabhi, V., Gajera, D., & Bhatt, D. (2013). Hybrid intrusion detection system for network security. Procedia Technology, 6, 1023-1032.

[9] Ouaddah, A., Elkouch, R., Aboutabit, N., & Hafid, A. S. (2017). FairAccess: A new blockchain-based access control framework for the internet of things. Security and Communication Networks, 2017.

[10] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572.

[11] Shrestha, R., Poudel, S., Bhattarai, S., & Paudel, P. (2020). Federated learning for intrusion detection in IoT networks: Challenges and opportunities. 2020 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON), 109-116.

[12] Dorri, A., Kanhere, S. S., Jurdak, R., & Gauravaram, P. (2017). Blockchain in internet of things: Challenges and solutions. arXiv preprint arXiv:1608.05187.

[13] Ferrag, M. A., Ahmadi, F., Janicke, H., & Derhab, A. (2020). Intrusion detection in internet of things based on deep learning: A systematic approach. Computer Networks, 177, 107336.

[14] Pan, Y., Li, J., Xiang, Y., Zhang, Y., & Shen, X. S. (2020). Edge intelligence: Empowering the internet-of-things security. IEEE Network, 34(5), 22-29.

[15] Li, W., Tugcu, T., & Song, D. (2019). Privacy-preserving federated learning with blockchain. arXiv preprint arXiv:1912.11184.

[16] Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks on machine learning models. 2017 IEEE Symposium on Security and Privacy (SP)*, 618-631.