Meta Learning combined with Federated Learning for Detecting Blackhole Attacks

Rathina Kumar, Getzi Jeba Lellipushpam Paulraj, Immanuel Johnraja Jebadurai

Karunya Institute of Technology and Sciences, rathinakumar@karunya.edu.in

Abstract.

Cybersecurity threats, such as blackhole attacks, pose a significant challenge to modern networks by disrupting data flow and compromising communication integrity. To deal with this, we propose a novel detection framework that blends Meta Learning (ML) and Federated Learning (FL) to enhance adaptability and privacy. ML enables models to rapidly adjust to new attack patterns with minimal data, while FL ensures decentralized training without exposing sensitive network information. At the core of this approach, Model Agnostic Meta Learning (MAML) improves detection accuracy by learning efficient initialization strategies. A comparative evaluation against Bayesian Meta Learning and Metric Based Meta Learning highlights MAML's effectiveness in managing class imbalances and adapting to evolving threats. Experimental results demonstrate promising detection accuracy, paving the way for further refinements to enhance performance, particularly in identifying rare attack scenarios.

Keywords. Meta-learning, Federated learning, Blackhole attacks, Network security.

1. INTRODUCTION

As cyberspace continues to expand, cybersecurity threats such as blackhole attacks are also becoming sophisticated, compromising network integrity. Traditional cybersecurity measures, while effective in controlled environments, struggle to implement detection mechanisms that ensure both accuracy and data privacy. Federated Learning (FL) has become a viable approach by enabling distributed model training across multiple devices without sharing raw data [1]. This decentralized approach reduces the risks associated with centralized data aggregation while maintaining model effectiveness. However, FL's reliance on large datasets poses challenges in scenarios where labeled attack samples are scarce, as seen in blackhole attack detection [2]. To address this issue, Meta Learning (ML), particularly Model Agnostic Meta Learning (MAML), is integrated with FL to enhance model adaptability. Model Agnostic Meta Learning (MAML) allows machine learning models to quickly adapt to new tasks using minimal data, making it highly effective for detecting rare cyber threats like blackhole attacks [3]. This study evaluates Federated Meta Learning (FML), combining FL and MAML, for improved blackhole attack detection while addressing data heterogeneity and model consistency challenges.

All Open Access articles are peer-reviewed, distributed under the Creative Commons Attribution-Non Commercial 4.0 International

2. LITERATURE REVIEW

The combination of Meta Learning and Federated Learning (FL) is gaining traction in cybersecurity. FL preserves data privacy by enabling decentralized learning, reducing security risks while enhancing adaptability against evolving threats [4]. One of the key challenges in FL-based attack detection is the need for adaptive learning models that can efficiently detect new and upcoming threats. Finn et al. [5] introduced Model Agnostic Meta Learning (MAML) to allow models to adapt quickly with limited training data, making it ideal for blackhole attack detection. Zhao et al. [6] highlight the effect of non-IID data on FL models, emphasizing the need for adaptive mechanisms to ensure consistent attack detection. Jiang et al. [7] enhance FL with personalized Meta Learning, improving blackhole attack detection by tailoring models to specific network patterns. Chen et al. [8] stress robust aggregation techniques to counter these risks, while Sun et al. [9] reinforce the need for secure learning mechanisms. Sohail et al. [10] introduced INFUSE, addressing class imbalances and limited attack data.

The literature suggests that combining meta-learning and federated learning provides a potential solution to the

obstacles of blackhole attack detection. The fusion of FL's privacy preserving architecture with meta-learning's rapid adaptation capabilities paves the way for more effective and resilient cybersecurity solutions in distributed network environments.

3. DATASET DESCRIPTION

For this study, we employ a tailored dataset designed to examine network traffic patterns for identifying and analyzing blackhole attacks. It is structured with 18 distinct attributes, as outlined in Table 3.1. The dataset is systematically structured as follows:

Table 3.1 Features of Custom Dataset

No.	Feature Name	Feature
F1	time	Timestamp of the packet transmission
F2	protocol type	Type of protocol associated with the packet
F3	destination	Destination of the packet network
F4	length	Packet Length in bytes
F5	info	More details about the packet and flags
F6	transmission_rate_per_1000_ms	Rate of transmission per 1000 milliseconds
F7	reception_rate_per_1000_ms	Rate of reception per 1000 milliseconds
F8	transmission_averageper_sec	Average transmission rate per second
F9	reception average_per_sec	Average reception rate per second
F10	transmission count per_sec	Count of transmissions per second
F11	reception_count_per_sec	Count of receptions per second.
F12	transmission total_duration_per_sec	Total duration of transmission per second
F13	reception total_duration_per_sec	Total duration of reception per second
F14	dao	Data Access Object specific to context
F15	dis	Data Integrity Status specific to context
F16	dio	Data Input/output specific to context
F17	category	Category of traffic: Normal or Blackhole
F18	label	Traffic type label: 0 (Normal), 1(Blackhole)

The dataset includes normal and blackhole attack traffic, where malicious nodes drop packets, causing data loss.

4. METHODOLOGY

This part describes the various models used in the research.

4.1 Federated Learning Framework

Federated Learning (FL) allows decentralized model training without sharing raw data, ensuring privacy in sensitive tasks like blackhole attack detection. Instead of centralizing data, clients train models locally and share updates, reducing security risks. Differences in client availability due to power constraints or connectivity issues can disrupt training [11].

4.2 Model Agnostic Meta Learning (MAML)

MAML is a meta-learning strategy designed to enhance the initial configuration of models, thereby enabling rapid adjustment to new tasks. In practical applications, MAML's inner loop helps a model adapt to specific network conditions, while the outer loop ensures it generalizes well across different scenarios.

1. Inner Loop (Task-Specific Updates*)*: The model is updated using a few gradient steps

based on specific task data.

2. Outer Loop (Meta-Updates*)*: The gradients from multiple tasks are aggregated to refine the shared initialization, ensuring better generalization across different tasks.

4.3 Bayesian Meta-Learning

Bayesian Meta-Learning enhances model reliability by quantifying uncertainty in predictions, making it effective in low-data scenarios. Using Bayesian Neural Networks (BNNs) and Gaussian Processes (GPs), it improves resilience to noisy data and generalization. FL, it enables probabilistic updates, ensuring robust blackhole attack detection in variable network conditions [12].

4.4 Metric Based Meta Learning (Siamese Network)

Metric Based Meta Learning, particularly through Siamese Networks, focuses on learning similarity metrics for comparing tasks or data points, making it well-suited for binary classification tasks such as blackhole attack detection. It maps data into an embedding space, applying Euclidean distance for classification. Optimized via contrastive or triplet loss, it enhances pattern recognition.

5. IMPLEMENTATION

This section includes the implementation of the models and a brief explanation of the metrics used to assess the outcomes.

5.1 Implementation of the models

The flowchart of the planned framework is displayed in Figure 5.1, which outlines the training and evaluation process.

1. **Data Splitting**: Divide the dataset into distinct sections for training, validation, and testing to ensure proper model evaluation.
2. **Model Initialization**: Establish a global model and distribute initial parameters to each client for localized training.
3. **Local Training**: Every client independently updates its local model using its assigned dataset while preserving data privacy.
4. **Weight Aggregation**: Aggregate weights from local models to update the global model.
5. **Model Evaluation**: Evaluate the global model on the validation set.
6. **Final Testing**: Evaluate the global model using the test dataset and generate results.

Figure 5.1. Federated Meta-Learning Workflow

The optimal hyperparameter settings for the models are as follows: For MAML, the learning rate is 0.0008, with 3 local epochs and 15 rounds. Bayesian Meta-Learning uses a learning rate of 0.001, a dropout rate of 0.5, 10 epochs, and 100 Monte Carlo samples. Metric Based Meta Learning (Siamese Network) has a learning rate of 0.001, hidden layer sizes of [64, 32], and 10 epochs.

5.2 Confusion Matrix

A confusion matrix is a systematic presentation of classification results between predicted labels and actual outcomes. Below is the confusion matrix representation (Table 5.1): The overall accuracy can be computed using the equation provided below equation:

$$\text{Accuracy} = TP + TN / TP + TN + FP + FN$$

Table 5.1 Confusion Matrix

	Predicted: Attack data	Predicted: Normal data
Actual: Attack data	True Positive (TP)	False Negative (FN)
Actual: Normal data	False Positive (FP)	True Negative (TN)

6. RESULT AND DISCUSSION

This part examines the experimental findings and compares the accuracy of three models. The performance of all three models is listed below.

Model Agnostic Meta Learning (MAML) - 0.9367 Bayesian Meta Learning – 1.00
Metric Based Meta Learning – 0.4500

Bayesian Meta-Learning achieved 100% accuracy but likely overfit, limiting generalization. MAML performed well at 93.67%, with 97.49% precision and 89.69% recall, making it suitable for dynamic environments. Siamese Networks struggled with 45% accuracy, requiring improved embeddings and refined metrics for better blackhole attack detection [13].

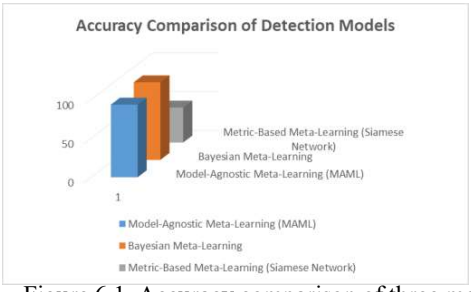

Figure 6.1. Accuracy comparison of three models

The graphical comparison in Figure 6.1 highlights MAML's superior accuracy, reinforcing its robustness across scenarios.

7. CONCLUSION

This study presents a framework combining Meta Learning and Federated Learning (FL) for enhanced blackhole attack detection in distributed networks. By integrating Model Agnostic Meta Learning (MAML) with FL's privacy preserving features, the system ensures accurate detection while maintaining data confidentiality. FL's decentralized structure prevents raw data from being shared, reducing the risks linked to centralized information storage. Evaluating scalability in real-world IoT and industrial networks will be crucial for broader adoption. Meta-learning techniques have been explored in network intrusion detection, highlighting their effectiveness in adapting to evolving threats [14].

8. REFERENCES

[1] Yang, Y., Liu, Y., & Zhang, Y., "Federated Learning for Cybersecurity: A Survey," IEEE Access, vol. 8, pp. 123456-123467, 2020.

[2] Huang, Y., & Wang, Y., "A Survey on Federated Learning: From Theory to Applications," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 2, pp. 1-18, 2021.

[3] Fallah, A., Mokhtari, A., & Ozdaglar, A., "Personalized Federated Learning with Theoretical Guarantees: A Model-Agnostic Meta-Learning Approach," Advances in Neural Information Processing Systems (NeurIPS), 2020.

[4] Kairouz, P., McMahan, B., Avent, B., et al., "Advances and Open Problems in Federated Learning," Foundations and Trends in Machine Learning, vol. 14, no. 1, pp. 1-210, 2019.

[5] Finn, C., Abbeel, P., & Levine, S., "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," Proc. 34th International Conference on Machine Learning (ICML), 2017.

[6] Zhao, Y., Li, M., Lai, Z., Suda, N., Civin, D., & Chandra, V., "Federated Learning with Non-IID Data," arXiv preprint, arXiv:1806.00582, 2018.

[7] Jiang, Y., Konecny, J., Rush, K., & Kannan, R., "Improving Federated Learning Personalization via Model Agnostic Meta Learning," Proc. NeurIPS Workshop on Federated Learning for Data Privacy and Confidentiality, 2019.

[8] Chen, D., Wu, H., Liu, H., et al., "Deep Federated Learning for Intrusion Detection in IoT Applications," IEEE Internet of Things Journal, vol. 8, no. 4, pp. 2291-2301, 2021.

[9] Sun, Y., Li, Q., Wang, X., et al., "Towards Federated Learning Models Robust to Poisoning Attacks," Proc. 35th AAAI Conference on Artificial Intelligence, vol. 35, no. 15, pp. 13601-13608, 2021.

[10] Sohail, A., Malik, M., & Khan, R., "Deep Neural Networks based Meta-Learning for Network Intrusion Detection," arXiv preprint, arXiv:2302.09394, 2023.

[11] Wang, J., Yurochkin, M., Sun, Y., Papailiopoulos, D., & Khazaeni, Y., "Federated Learning with Matched Averaging," Proc. 8th International Conference on Learning Representations (ICLR), 2020.

[12] Qu, Y., Li, Y., & Zhang, X., "Federated Sharpness-Aware Minimization (FedSAM): Towards Robust Federated Learning with Local Sharpness Control," Proc. IEEE International Conference on Data Science and Advanced Analytics (DSAA), 2022.

[13] Foret, P., Kleiner, A., Mobahi, H., & Neyshabur, B., "Sharpness-Aware Minimization for Efficiently Improving Generalization," Proc. International Conference on Learning Representations (ICLR), 2021.

[14] Zhang, Y., & Wang, J., "Meta-Learning for Network Intrusion Detection: A Review," *Journal of Network and Computer Applications*, vol. 178, 102973, 2021.

Biographies

K Rathina Kumar, a B.Tech Computer Science and Engineering student at Karunya Institute of Technology and Sciences (2025 batch), has experience in Android development, networking, and machine learning. He has worked on projects like an Online Election Management System and NutriGuide and interned at Cisco, IBM, and YBI Foundation. His interests include AI, networking, and web technologies.

Getzi Jeba Lellipushpam Paulraj received her B.E. degree in Electronics and Communication Engineering from M.S. University, Tirunelveli, India in the year 2004. She received her M.Tech. degree in Network and Internet Engineering from Karunya University, Coimbatore, India in the year 2009. She received her Ph.D. from Karunya Institute of Technology and Sciences, Coimbatore, India in the year 2018. Her areas of interest include IoT, Fog Computing and Data Analytics.

Immanuel Johnraja Jebadurai received his B.E. degree in Computer Science and Engineering from M.S. University, Tirunelveli, India in the year 2003. He received his
M.E. degree in Computer Science and Engineering from Anna University, Chennai India in the year 2005. He received his Ph.D. in Computer Science and Engineering from Karunya Institute of Technology and Sciences, Coimbatore, India in the year 2017. His areas of interest include Network Security, Vehicular Ad-Hoc Networking and IoT.