F-NIDS: Sistema de Detecção de Intrusão baseado em Aprendizado Federado

Jonathas Alves de Oliveira[1], Vinícius P. Gonçalves[1] (Coorientador)
Geraldo P. Rocha Filho[2] (Orientador)

[1]Universidade de Brasília (UnB)

[2]Universidade Estadual do Sudoeste da Bahia (UESB)

jhsalves@gmail.com, vpgvinicius@unb.br, geraldo.rocha@uesb.edu.br

Resumo. *Nesta dissertação, abordamos os desafios de escalabilidade e segurança enfrentados por Sistemas de Detecção de Intrusão de Rede (NIDS) distribuídos em cenários de IoT, decorrentes de preocupações com privacidade. Para superar esses desafios, propomos o F-NIDS (Federated Network Intrusion Detection System), um sistema modelado com base no Federated Learning (FL), integrando técnicas de comunicação assíncrona e Privacidade Diferencial (DP). Além disso, foram desenvolvidas três estratégias para equilibrar desempenho e privacidade. Os resultados demonstram que o F-NIDS alcançou alta robustez e desempenho, com valores de detecção superiores em cenários distribuídos. A estratégia de ajuste de parâmetros permitiu identificar o nível ideal de ruído gaussiano, garantindo a preservação da privacidade sem comprometer a acurácia. Portanto, o F-NIDS representa um avanço no estado da arte, combinando escalabilidade, segurança e privacidade em sistemas distribuídos, ratificando sua eficácia e potencial para redefinir os padrões de detecção de intrusões em ambientes de IoT.*

Abstract. *In this dissertation, we address the challenges of scalability and security faced by Network Intrusion Detection Systems (NIDS) in IoT scenarios, particularly those arising from privacy concerns. To overcome these challenges, we propose the F-NIDS (Federated Network Intrusion Detection System), a system designed based on Federated Learning (FL), integrating asynchronous communication techniques and Differential Privacy (DP). Additionally, three strategies were developed to balance performance and privacy. The results demonstrate that F-NIDS achieves high robustness and performance, with superior detection values in distributed scenarios. The parameter adjustment strategy allowed the identification of the optimal level of Gaussian noise, ensuring privacy preservation without compromising accuracy. Therefore, F-NIDS represents a significant advancement in the state of the art, combining scalability, security, and privacy in distributed systems, reaffirming its effectiveness and potential to redefine intrusion detection standards in IoT environments.*

1. Introdução

Nas últimas décadas, a crescente interconexão entre pessoas, dispositivos e serviços tem impulsionado o paradigma da IoT. Esse avanço trouxe desafios relacionados à escalabilidade, latência e privacidade das informações [Rahman and Asyhari 2019]. Arquiteturas

descentralizadas surgem como uma alternativa promissora para mitigar essas questões, oferecendo maior disponibilidade e escalabilidade [Roman et al. 2013, Cavalcante et al. 2022] No contexto da segurança, os NIDSs têm se destacado como ferramentas essenciais para monitorar e identificar atividades mal-intencionadas em redes de comunicação [Chaabouni et al. 2019]. No entanto, a natureza dinâmica e heterogênea das redes de IoT desafia a eficácia dos NIDSs [Bertino and Islam 2017], uma vez que os dispositivos heterogêneos, tais como sensores, câmeras e dispositivos médicos conectados, utilizam diferentes protocolos de comunicação, comportamentos normais variáveis e recursos computacionais limitados.

Tradicionalmente, os NIDSs operam de forma contínua durante o ciclo de vida das redes, mas em ambientes de IoT, em que os recursos são limitados e os dispositivos são diversos e altamente interconectados, esses sistemas enfrentam limitações significativas [Rahman et al. 2020, Bertino and Islam 2017]. Pesquisadores têm explorado o uso de técnicas de *Machine Learning* (ML) para aprimorar os NIDSs, melhorando a detecção de ataques cibernéticos [Chaabouni et al. 2019]. Contudo, questões relacionadas à privacidade em modelos de ML, especialmente em arquiteturas descentralizadas, permanecem um desafio crucial [Cabrero-Holgueras and Pastrana 2021]. Além disso, a necessidade de preservar a confidencialidade dos dados e modelos de treinamento enquanto se mantém a escalabilidade e a robustez é um campo aberto para pesquisa [Chen et al. 2020].

Nesse contexto, o FL surge como uma abordagem descentralizada que permite a execução de tarefas de ML com maior preservação da privacidade, ao evitar a troca direta de dados de treinamento [Zhu et al. 2021, Stergiou et al. 2021]. Embora o FL melhore a confidencialidade, a distribuição eficiente dos serviços de detecção de intrusão entre agentes na rede ainda exige avanços. Estratégias de comunicação assíncrona, que utilizam técnicas como enfileiramento de mensagens e modelos de *publish/subscribe*, oferecem soluções para atender à demanda por escalabilidade e resiliência [Hohpe and Woolf 2003]. Apesar dos benefícios, desafios como a complexidade de implementação e maior latência destacam a necessidade de soluções inovadoras nesse campo [Pautasso et al. 2008, Fowler 2002].

A implementação de NIDS desempenha um papel fundamental na segurança cibernética, protegendo a integridade, confidencialidade e disponibilidade dos dados em ambientes cada vez mais interconectados [Stallings 2017]. Salienta-se que, no entanto, a coleta e análise de dados para detecção de intrusões podem comprometer a privacidade dos usuários, gerando preocupações éticas e jurídicas [Potiguara Carvalho et al. 2020, Myers et al. 2008]. Essa tensão entre privacidade e eficácia destaca a importância de integrar mecanismos de proteção de dados, tais como anonimização e privacidade diferencial, para garantir a aceitação e eficácia contínua do NIDS [Shi et al. 2021]. É válido sublinhar que modelos de ML para NIDS enfrentam ameaças como ataques de inversão de modelo e inferência de membros, que comprometem a confidencialidade dos dados de treinamento [Fredrikson et al. 2015, Baluta et al. 2022]. Esses desafios reforçam a necessidade de desenvolver estratégias robustas que conciliem privacidade e desempenho na detecção de intrusões em cenários de IoT.

2. Objetivo

O principal objetivo desta dissertação é propor o F-NIDS, um sistema de detecção de intrusão distribuído modelado para superar as limitações de privacidade em cenários descentralizados, equilibrando desempenho e robustez, enquanto mantém uma acurácia eficiente. A proposta inclui uma arquitetura descentralizada inovadora que distribui tarefas entre agentes de detecção e emprega técnicas de privacidade diferencial para proteger os dados e modelos treinados de forma eficaz. O F-NIDS busca alcançar três metas principais: (i) garantir alta acurácia na classificação de tráfego em cenários de IoT distribuídos; (ii) implementar e validar uma arquitetura escalável e resiliente, capaz de se adaptar às variações de demanda; e (iii) assegurar a confidencialidade dos dados de treinamento e dos modelos, preservando a integridade e segurança das informações sensíveis.

3. Contribuições

As principais contribuições desta dissertação, que avançam o estado da arte, são as seguintes:

- **Proposta, Desenvolvimento e Validação do F-NIDS com Arquitetura**: O F-NIDS é um sistema de detecção de intrusão distribuído que utiliza técnicas de FL e Privacidade Diferencial (DP) para garantir escalabilidade e confidencialidade em redes IoT. Sua arquitetura escalável é baseada no algoritmo `FedAvg`, permitindo a agregação de modelos locais, reduzindo a largura de banda necessária e maximizando a eficiência em cenários distribuídos. Além disso, o sistema integra um mecanismo de comunicação assíncrona fundamentado no modelo *publish/subscribe*, suportado por protocolos como MQTT, proporcionando maior escalabilidade, resiliência e adaptabilidade a diferentes cenários de IoT.
- **Integração de Privacidade Diferencial**: Para assegurar a confidencialidade dos dados de treinamento, o F-NIDS implementa o algoritmo DP-SGD (*Differentially Private Stochastic Gradient Descent*). Esse mecanismo adiciona ruído gaussiano aos gradientes durante o treinamento local, protegendo os modelos contra ataques de inferência de membros e garantindo que o modelo global herde as propriedades de privacidade dos modelos locais. Essa contribuição permite que o F-NIDS alcance altos níveis de privacidade sem comprometer o desempenho.
- **Mecanismo de Detecção Distribuída**: A detecção de intrusão no F-NIDS é realizada por meio de um modelo de comunicação assíncrona baseado no paradigma *publish/subscribe*. Esse mecanismo permite que os agentes de detecção processem solicitações de forma escalável e isolada dos clientes, garantindo maior robustez e tolerância a falhas. Além disso, o desacoplamento entre clientes e agentes minimiza os riscos de exploração maliciosa dos modelos de detecção.

4. Publicação

Como resultado da dissertação, foram publicados dois trabalhos. O primeiro foi apresentado e publicado nos Anais do XLI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC 2023) [Oliveira et al. 2023]. O segundo foi publicado na revista internacional, *Elsevier - Computer Networks* [Oliveira et al. 2023]. A Tabela 1 apresenta um resumo das publicações.

Tabela 1. Resumo das publicações.

Trabalho	Tipo de Publicação	Local	Qualis	Impact Factor	Citações
[Oliveira et al. 2023]	Periódico	Computer Networks	A1	4.4	14
[Oliveira et al. 2023]	Conferência	SBRC	A4	–	2

5. Principais Resultados

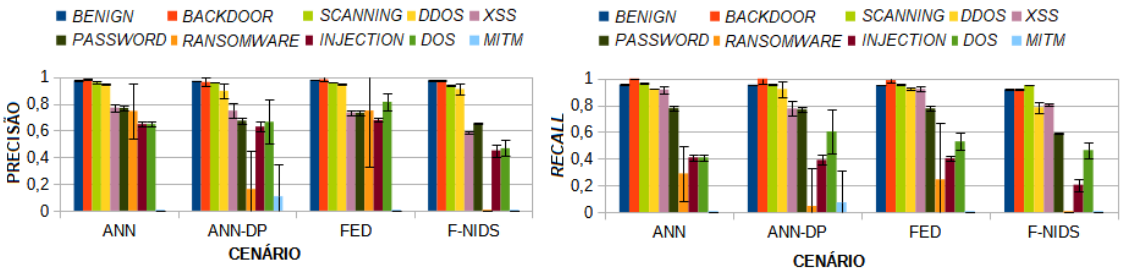
Para avaliar o F-NIDS, foram modelados cenários de classificação binária (tráfego normal versus malicioso), multiclasse (diferentes tipos de ataques) e simulação de pacotes provenientes de ambientes adversos. As métricas de predição binária e multiclasse (acurácia, precisão e *recall*) foram escolhidas para comparar o F-NIDS com três métodos amplamente utilizados em NIDS: ANN, ANN-DP e FED. A acurácia foi empregada para analisar a convergência durante as rodadas de treinamento, enquanto a precisão e o *recall* avaliaram o desempenho por classe.

Os experimentos foram realizados utilizando o conjunto de dados `NF-ToN-IoT-v2`, processado com a ferramenta `NetFlow`, contendo 41 atributos após a remoção de informações sensíveis, como endereços IP. O dataset, composto por 2,5 milhões de registros, foi dividido em 80% para treinamento e 20% para teste. No treinamento federado, os dados de treinamento foram igualmente distribuídos entre 100 agentes. A implementação do F-NIDS foi realizada com `Tensorflow` para o treinamento de redes neurais, `TF Privacy` para a aplicação de privacidade diferencial e `Flower` para aprendizado federado. Os hiperparâmetros do modelo foram ajustados utilizando técnicas de *hyperparameter tuning* para maximizar o desempenho e a eficiência.

5.1. Análise dos Principais Resultados

Na Figura 1, são apresentados os resultados de precisão e recall obtidos para cada classe utilizando o método de classificação F-NIDS, comparados aos métodos ANN, ANN-DP e FED. Para o F-NIDS, a precisão das previsões nas classes *Benign*, *Backdoor*, *Scanning* e *DDoS* não apresentou diferenças significativas em relação aos outros métodos. Destaca-se uma redução na precisão de classificação nas classes *Ransomware* e *MITM* ao usar o método F-NIDS. Esse comportamento estar relacionado ao número limitado de exemplos disponíveis para essas classes, menor do que o tamanho do minilote utilizado. De maneira geral, tais resultados permitem concluir que a aplicação conjunta do algoritmo `FedAvg` com o `DP-SGD` teve um impacto que não afetou a precisão e no recall de diversas classes. Esse impacto foi mais evidente em classes com menor quantidade de exemplos de treinamento, refletindo a natureza não identicamente distribuída do conjunto de dados original.

A Tabela 2 apresenta os resultados das métricas de classificação binária para analisar a eficácia do F-NIDS na classificação de tráfego benigno e malicioso. Os resultados indicam que o F-NIDS, embora tenha mostrado desempenho ligeiramente inferior em termos de precisão e acurácia em comparação aos métodos centralizados (ANN e ANN-DP) e federados sem privacidade (FED), ainda mantém níveis competitivos de recall. Essa diferença é atribuída à introdução do ruído gaussiano pelo algoritmo `DP-SGD`, necessário para garantir privacidade diferencial. Esses resultados destacam o equilíbrio alcançado pelo F-NIDS entre robustez na proteção da privacidade e desempenho na classificação



(a) Precisão (b) *Recall*

Figura 1. Resultados de precisão e *recall* multiclasse.

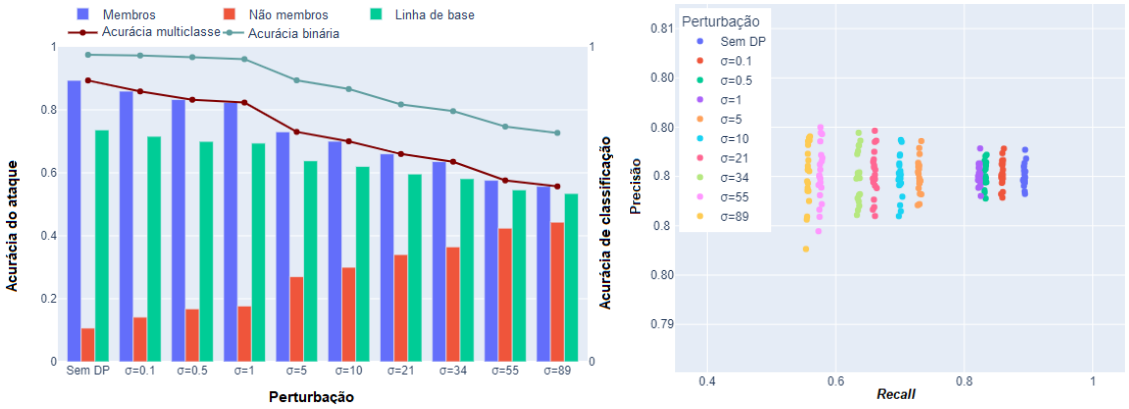
binária, demonstrando que ele é uma solução viável para cenários em que a confidencialidade dos dados é uma prioridade.

Tabela 2. Avaliação das métricas de classificação binária dos métodos. Cada resultado é seguido por seu desvio padrão $\pm$.

Método	Precisão	*Recall*	Acurácia
ANN	$0{,}975 \pm 0{,}005$	$0{,}955 \pm 0{,}005$	$0{,}974 \pm 0$
ANN-DP	$0{,}971 \pm 0{,}002$	$0{,}954 \pm 0{,}003$	$0{,}973 \pm 0{,}001$
FED	$0{,}981 \pm 0$	$0{,}980 \pm 0$	$0{,}974 \pm 0$
F-NIDS	$0{,}885 \pm 0{,}28$	$0{,}970 \pm 0{,}264$	$0{,}962 \pm 0$

A Figura 2a apresenta os resultados de acurácia de um ataque de inferência de membros, juntamente com a acurácia de classificação do modelo original. Três comportamentos principais podem ser observados na Figura 2a. Primeiro, à medida que o valor do ruído gaussiano aumenta, a linha de base do ataque diminui, reduzindo a probabilidade de sucesso na inferência de não membros. Segundo, a capacidade de detectar não membros aumenta com o crescimento de σ, mas sem ultrapassar 50% de acurácia. Por fim, destaca-se que a acurácia da classificação do modelo permanece significativamente superior à linha de base do ataque, especialmente na detecção binária, evidenciando o equilíbrio entre utilidade de classificação e robustez. Com base nesses resultados, o valor de ruído escolhido para o F-NIDS foi $\sigma = 21$, pois atende aos critérios de robustez contra ataques de inferência de membros, mantendo a acurácia de detecção de intrusões acima de 80%. Além disso, a Figura 2b apresenta a relação entre precisão e recall nos resultados do ataque. Observa-se que a precisão do ataque não sofre impacto significativo com o aumento do nível de ruído. Entretanto, o recall diminui consideravelmente à medida que o nível de ruído aumenta. Esses resultados indicam que, embora a probabilidade de um não membro ser erroneamente classificado como membro não seja substancialmente afetada por σ, a probabilidade de um membro real ser classificado como não membro aumenta significativamente com níveis mais altos de ruído.

A Tabela 3 apresenta os resultados do teste de desempenho do ataque adversário baseado em regras, considerando a linha de base (*L. Base*) do ataque, a acurácia do ataque de inferência adversário (*Acr. adv membros*), a acurácia da classificação binária (*Acr. binária*) e o *F1 score* obtido. Esses resultados foram calculados para diferentes valores de σ, que representa um parâmetro de variabilidade ou ruído no modelo. Observa-se que,



(a) Performances de classificação e robustez

(b) Precisão e recall do ataque

Figura 2. Análise de acurácia e robustez para uma inferência de membros baseada em regra.

à medida que σ aumenta, há uma redução gradual na linha de base do ataque, indicando menor eficácia do ataque adversário em condições de maior incerteza. Da mesma forma, a acurácia do ataque de inferência e a precisão binária diminuem conforme o valor de σ cresce, refletindo a dificuldade de identificar corretamente os membros e não membros no conjunto adversário. O *F1 score*, que combina precisão e recall, também sofre uma leve degradação, especialmente em valores mais altos de σ, embora mantenha um desempenho consistente em cenários de baixo ruído. Esses resultados indicam que o ataque adversário é mais eficaz em cenários com menor variabilidade, representados pelos menores valores de σ. Por outro lado, conforme o ruído aumenta, a eficácia do ataque diminui, impactando diretamente todas as métricas avaliadas.

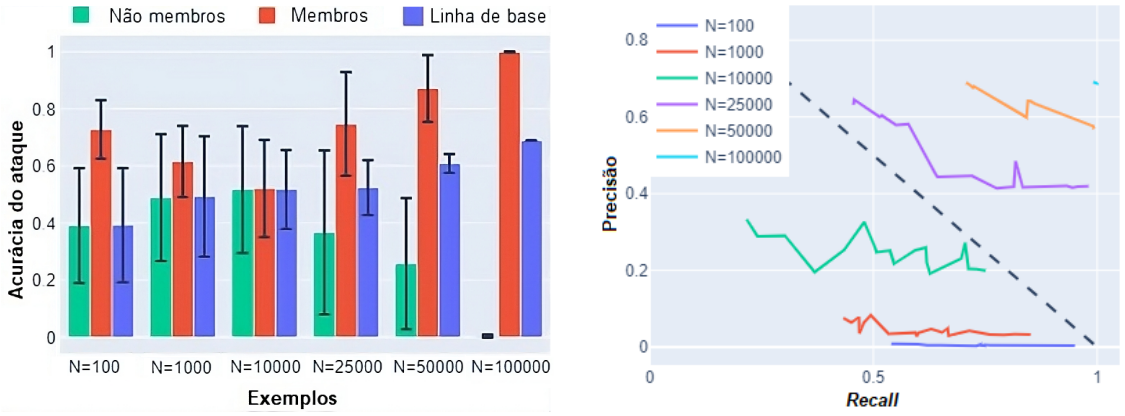
Tabela 3. Desempenho do ataque adversário baseado em regras.

σ	L. Base	Acr. adv membros	Acr. binária	F1 Score adv
0,1	$0{,}72 \pm 0{,}0010$	$0{,}86 \pm 0{,}0010$	$0{,}97 \pm 0{,}0010$	$0{,}83 \pm 0{,}0007$
1	$0{,}69 \pm 0{,}0008$	$0{,}82 \pm 0{,}0011$	$0{,}96 \pm 0{,}0014$	$0{,}81 \pm 0{,}0006$
21	$0{,}60 \pm 0{,}0013$	$0{,}66 \pm 0{,}0015$	$0{,}82 \pm 0{,}0029$	$0{,}72 \pm 0{,}0012$
55	$0{,}55 \pm 0{,}0016$	$0{,}58 \pm 0{,}0016$	$0{,}75 \pm 0{,}0030$	$0{,}71 \pm 0{,}0012$
89	$0{,}53 \pm 0{,}0016$	$0{,}56 \pm 0{,}0015$	$0{,}73 \pm 0{,}0030$	$0{,}66 \pm 0{,}0013$

Embora o classificador modelado no F-NIDS seja protegido com privacidade diferencial, se uma amostra de tamanho excessivo for armazenada em alguns dos DAs, isso ainda pode representar um risco a ser considerado. Esse problema ocorre porque, dependendo do seu tamanho, uma amostra pode conter propriedades estatísticas semelhantes às do conjunto de dados original. O aprendizado federado desempenha um papel essencial ao permitir a divisão de grandes volumes de dados em amostras menores distribuídas em um número maior de agentes, fazendo com que frações menores de dados confidenciais sejam armazenadas com maior confidencialidade. Para essa investigação, seis amostras do conjunto de dados de treinamento, de tamanhos distintos, foram selecionadas e usadas em um ataque adversário do *Membership Inference Black Box* — Inferência de Membros em Caixa Preta —, também contido na biblioteca ART.

A Figura 3a apresenta a eficácia dos ataques obtidos por classificadores treinados com diferentes valores de N. É possível observar que a eficácia na detecção de membros e não membros diminui à medida que o valor de N diminui. Quando o valor de $N = 100$ é atingido, ele coincide com o valor mais baixo registrado para a linha de base da precisão do ataque, pois os valores de precisão da detecção para membros e não membros permaneceram abaixo de 40%. Entretanto, um classificador treinado com um tamanho de exemplo tão baixo provavelmente não seria muito útil para a detecção de intrusões. Portanto, um número de exemplos de treinamento $10K \leq N \leq 25K$ atinge o nível razoavelmente aceitável para os três indicadores presentes na figura, em que ambos os indicadores têm valores próximos a 50% de eficiência de ataque.

Na Figura 3b, a precisão e o recall obtidos nos resultados do ataque, foram comparados. Por meio do gráfico, é possível concluir que a precisão tem um aumento significativo à medida que o número de exemplos de treinamento é aumentado. Isso indica que quanto maior o número de exemplos capturados por um agente malicioso, maior a chance de construir um modelo capaz de inferir membros. No entanto, o recall não mostrou grande sensibilidade ao número de exemplos usados no treinamento do *ShadowModel*, influenciando pouco a capacidade desse modelo de categorizar corretamente os não membros.



(a) Acurácia do ataque para cada *ShadowModel*

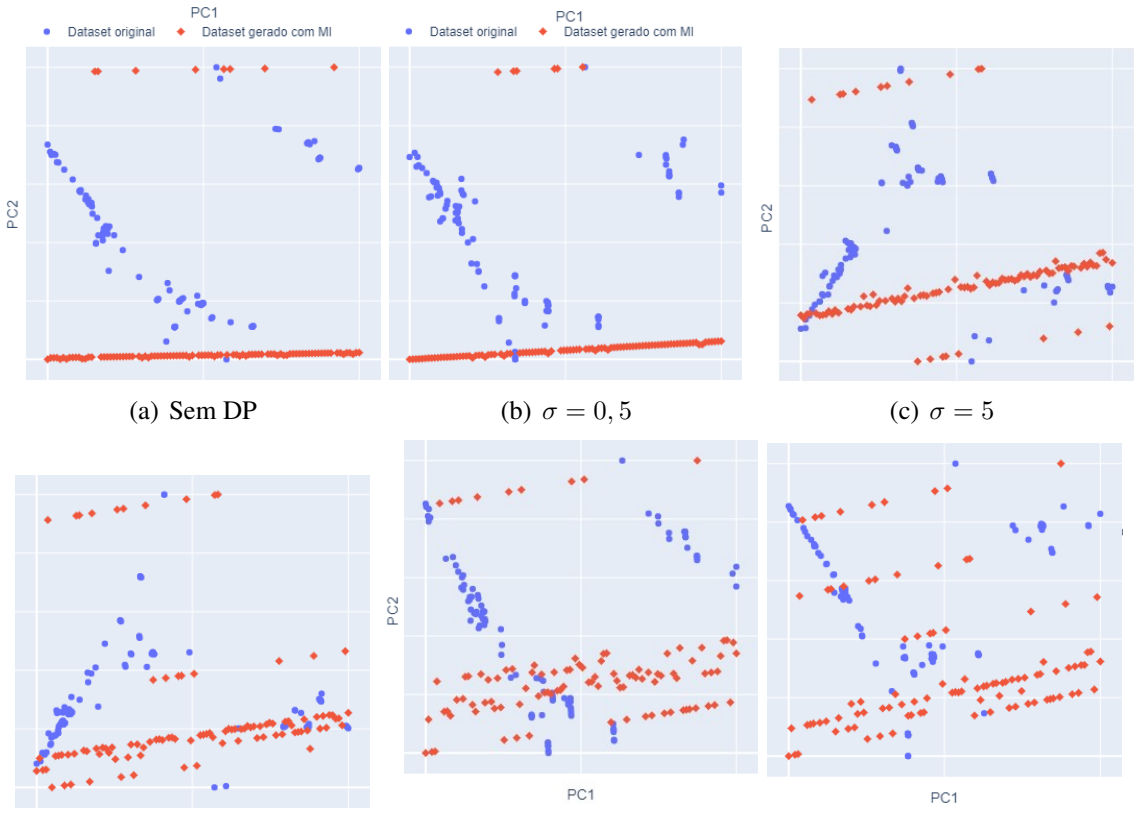
(b) Precisão e *Recall* curves for each *ShadowModel*

Figura 3. Análise de acurácia e robustez num cenário de ataque de inferência de membros.

O último teste de robustez realizado no classificador F-NIDS foi o MI (*Model Inversion* — Inversão de Modelo). Esses ataques à privacidade envolvem tentativas de um adversário em reconstruir informações confidenciais sobre indivíduos ao explorar as saídas de um modelo de aprendizado de máquina. O adversário utiliza consultas estratégicas ao modelo, combinadas com conhecimento prévio, para inferir atributos privados ou pontos de dados usados na geração das previsões do modelo.

A Figura 4 apresenta uma visualização dos conjuntos de dados originais e adversários gerados pelos seis modelos treinados com diferentes níveis de ruído, reduzidos a dois componentes principais. Na Figura 4a, sem o uso do algoritmo DP no modelo-alvo, o conjunto gerado pelo MI exibe um padrão semelhante ao original, com pontos próximos e distribuídos de forma parecida. É possível traçar um limite de decisão similar

para distinguir regiões em ambos os conjuntos. No entanto, à medida que o nível de ruído aumenta, os padrões dos exemplos gerados tornam-se progressivamente mais distintos dos dados originais. Esse comportamento é especialmente notável quando $\sigma = 89$, como mostrado na Figura 4f. Nesse ponto, os dados gerados apresentam dispersão significativamente diferente, tornando inviável traçar um limite de decisão comum que funcione para ambos os conjuntos. Além disso, observa-se que, com $\sigma \geq 5$ (Figura 4c), não é mais possível traçar regiões de decisão lineares para separar os conjuntos de treinamento e os adversários com limites semelhantes. Isso indica que, à medida que o ruído gaussiano é aumentado, os conjuntos de dados originais e adversários tornam-se substancialmente diferentes. Com níveis de ruído elevados, como $\sigma \geq 5$, os dados adversários tornam-se inadequados para o treinamento de um modelo adversário capaz de reproduzir amostras confidenciais do conjunto de treinamento original.



(a) Sem DP (b) $\sigma = 0,5$ (c) $\sigma = 5$

(d) $\sigma = 21$ (e) $\sigma = 55$ (f) $\sigma = 89$

Figura 4. Dimensões reduzidas para dois componentes dos conjuntos de dados originais e adversários.

6. Considerações Finais

Esta pesquisa demonstrou que um dos principais desafios na IoT é a implementação de um sistema de detecção de intrusão descentralizado e escalável, capaz de enfrentar as crescentes preocupações com privacidade e segurança. Para superar esses desafios, foi desenvolvido o F-NIDS, um sistema baseado em FL que integra técnicas de comunicação assíncrona e DP.

Os resultados indicaram que, com níveis mais baixos de ruído gaussiano, o F-NIDS apresenta métricas de acurácia, precisão e recall comparáveis às estratégias tradicionais baseadas em aprendizado centralizado. Além disso, os testes demonstraram que um valor de $\sigma = 21$ é eficaz para proteger contra ataques de inferência de membros baseados em regras de caixa-preta e ataques de inversão de modelo. Adicionalmente, manter o tamanho da amostra nos agentes de detecção entre $N = 10K$ e $N = 25K$ contribui para salvaguardar a confidencialidade dos agentes, mesmo em cenários adversos.

Com sua arquitetura adaptável e resultados consistentes, o F-NIDS não apenas alcançou alta robustez e desempenho em cenários distribuídos, mas também demonstrou sua eficácia em atender às demandas específicas de redes IoT. O sistema provou ser capaz de preservar a privacidade sem comprometer a acurácia, ao mesmo tempo que se adapta a diferentes cenários e flutuações de carga. Portanto, o F-NIDS se destaca como uma solução que avança o estado da arte por ser capaz de redefinir os padrões de detecção de intrusões em redes de IoT e estabelecer um novo paradigma para sistemas distribuídos com alta segurança e escalabilidade.

Agradecimentos

Esta pesquisa foi financiada pela FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), Brasil, nº 2021/06210-3.

Referências

Baluta, T., Shen, S., Hitarth, S., Tople, S., and Saxena, P. (2022). Membership inference attacks and generalization: A causal perspective. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, CCS '22, page 249–262, New York, NY, USA. Association for Computing Machinery.

Bertino, E. and Islam, N. (2017). Botnets and internet of things security. *Computer*, 50(2):76–79.

Cabrero-Holgueras, J. and Pastrana, S. (2021). Sok: Privacy-preserving computation techniques for deep learning. *Proceedings on Privacy Enhancing Technologies*, 2021(4):139–162.

Cavalcante, I. C., Meneguette, R. I., Torres, R. H., Mano, L. Y., Gonçalves, V. P., Ueyama, J., Pessin, G., Amvame Nze, G. D., and Rocha Filho, G. P. (2022). Federated system for transport mode detection. *Energies*, 15(23):9256.

Chaabouni, N., Mosbah, M., Zemmari, A., Sauvignac, C., and Faruki, P. (2019). Network intrusion detection for iot security based on learning techniques. *IEEE Communications Surveys Tutorials*, 21(3):2671–2701.

Chen, H., Hussain, S. U., Boemer, F., Stapf, E., Sadeghi, A. R., Koushanfar, F., and Cammarota, R. (2020). Developing privacy-preserving ai systems: The lessons learned. In *2020 57th ACM/IEEE Design Automation Conference (DAC)*, pages 1–4.

Fowler, M. (2002). *Patterns of Enterprise Application Architecture*. Addison-Wesley Longman Publishing Co., Inc., USA.

Fredrikson, M., Jha, S., and Ristenpart, T. (2015). Model inversion attacks that exploit confidence information and basic countermeasures. In *Proceedings of the 22nd*

ACM SIGSAC Conference on Computer and Communications Security, CCS '15, page 1322–1333, New York, NY, USA. Association for Computing Machinery.

Hohpe, G. and Woolf, B. (2003). *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley Longman Publishing Co., Inc., USA.

Myers, J., Frieden, T., Bherwani, K., and Henning, K. (2008). Ethics in public health research: Privacy and public health at risk: Public health confidentiality in the digital age. *American journal of public health*, 98:793–801.

Oliveira, J., Meneguette, R., Gonçalves, V., Jr., R. S., Guidoni, D., Oliveira, J., and Filho, G. R. (2023). F-nids – sistema de detecção de intrusão descentralizado com base em aprendizado federado. In *Anais do XLI Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 29–42, Porto Alegre, RS, Brasil. SBC.

Oliveira, J. A., Gonçalves, V. P., Meneguette, R. I., de Sousa, R. T., Guidoni, D. L., Oliveira, J. C., and Rocha Filho, G. P. (2023). F-nids — a network intrusion detection system based on federated learning. *Computer Networks*, 236:110010.

Pautasso, C., Zimmermann, O., and Leymann, F. (2008). Restful web services vs. "big"' web services: Making the right architectural decision. In *Proceedings of the 17th International Conference on World Wide Web*, WWW '08, page 805–814, New York, NY, USA. Association for Computing Machinery.

Potiguara Carvalho, A., Potiguara Carvalho, F., Dias Canedo, E., and Potiguara Carvalho, P. H. (2020). Big data, anonymisation and governance to personal data protection. In *The 21st Annual International Conference on Digital Government Research*, dg.o '20, page 185–195, New York, NY, USA. Association for Computing Machinery.

Rahman, M. A. and Asyhari, A. T. (2019). The emergence of internet of things (iot): Connecting anything, anywhere. *Computers*, 8(2).

Rahman, M. A., Asyhari, A. T., Leong, L., Satrya, G., Hai Tao, M., and Zolkipli, M. (2020). Scalable machine learning-based intrusion detection system for iot-enabled smart cities. *Sustainable Cities and Society*, 61:102324.

Roman, R., Zhou, J., and Lopez, J. (2013). On the features and challenges of security and privacy in distributed internet of things. *Computer Networks*, 57(10):2266–2279. Towards a Science of Cyber Security Security and Identity Architecture for the Future Internet.

Shi, J., Ge, B., Liu, Y., Yan, Y., and Li, S. (2021). Data privacy security guaranteed network intrusion detection system based on federated learning. In *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6.

Stallings, W. (2017). *Network Security Essentials: Applications and Standards*. Pearson.

Stergiou, C. L., Psannis, K. E., and Gupta, B. B. (2021). Infemo: Flexible big data management through a federated cloud system. *ACM Trans. Internet Technol.*, 22(2).

Zhu, H., Zhang, H., and Jin, Y. (2021). From federated learning to federated neural architecture search: a survey. *Complex & Intelligent Systems*, 7.