Detecção de ataques em Internet das Coisas utilizando Mineração de Fluxos Contínuos de Dados

Isabela Hara Bando[1], **Bruno Bogaz Zarpelão**[1]

[1]Departamento de Computação – Universidade Estadual de Londrina (UEL)
Caixa Postal 10.011 – CEP 86057-970 – Londrina – PR – Brasil

isabela.hara.bando@uel.br, brunozarpelao@uel.br

Abstract. *The development of technology in recent years has significantly increased the use of interconnected smart devices, that is, the Internet of Things (IoT). Due to the large flow of shared and stored information, ensuring the security of IoT networks is essential, and also a huge challenge, given the several forms of cyberattacks that threaten these networks today. Most of the solutions studied for detecting cyber attacks use batch learning algorithms, which are trained from a static database, which leads to loss of effectiveness as new behaviors emerge in the network. The monitored data is constantly undergoing many natural changes and therefore it is necessary that the algorithms are able to adapt more easily. In this context, this project aims to study and implement different continuous data stream mining algorithms for intrusion detection in IoT networks, using incremental learning. The detection effectiveness of each of these algorithms will be evaluated on multiple publicly available datasets.*

Resumo. *O desenvolvimento da tecnologia nos últimos anos aumentou significativamente o uso de dispositivos inteligentes interconectados, isto é, a Internet das Coisas (Internet of Things - IoT). Devido ao grande fluxo de informações compartilhadas e armazenadas, garantir a segurança das redes IoT é essencial, e também um enorme desafio, tendo em vista os diversos tipos de ciberataques que ameaçam essas redes atualmente. Grande parte das soluções estudadas para detecção de ataques cibernéticos utilizam algoritmos de aprendizado em lote, cujo treinamento é feito a partir de uma base de dados estática, o que leva à perda de eficácia com o surgimento de novos comportamentos na rede. Os dados monitorados passam, constantemente, por muitas mudanças naturais e por isso, é necessário que os algoritmos sejam capazes de se adaptar com uma maior facilidade. Levando em conta esse contexto, este projeto visa estudar e implementar diferentes algoritmos de mineração de fluxo de dados contínuos para detecção de intrusões em redes IoT, utilizando o aprendizado incremental. A eficácia de detecção de cada um destes algoritmos será avaliada em múltiplos conjuntos de dados disponibilizados publicamente.*

1. Introdução

Com o avanço tecnológico e o aumento constante da conectividade, a Internet das Coisas (*Internet of Things* - IoT) emergiu como um campo de grande potencial de desenvolvimento. Além disso, a IoT está se tornando cada vez mais integrada à vida cotidiana das pessoas, estando presente em uma ampla variedade de dispositivos inteligentes, desde eletrodomésticos e sistemas de segurança até veículos autônomos e equipamentos médicos.

Portanto, garantir a segurança das redes IoT e dos sistemas é um desafio relevante, que tem sido objeto de estudos ao longo dos anos [1, 8, 15, 20].

Devido à diversidade de ameaças existentes, à falta de conscientização dos usuários e à ausência de atualizações regulares de software, torna-se imperativo desenvolver modelos capazes de se adaptar às mudanças e depender cada vez menos da intervenção humana.

Embora os algoritmos de aprendizado em lote (*batch*) tenham sido amplamente utilizados na detecção de ataques em redes [13, 16, 18], eles possuem uma limitação significativa: são treinados com um conjunto de dados estático e têm dificuldade em lidar com mudanças na rede, como mudanças de conceito (*concept drift*), exigindo a atualização ou retreinamento do modelo [7, 14]. Como as redes de computadores estão em constante evolução e os padrões de tráfego podem variar com o tempo, modelos treinados com dados do passado podem perder sua eficácia rapidamente, resultando em taxas de detecção de ataques mais baixas.

Dada essa limitação, outra abordagem que pode ser explorada é o uso de algoritmos de aprendizado em fluxo contínuo de dados. Esses algoritmos são projetados para lidar com situações que exigem aprendizado incremental, permitindo que o modelo seja ajustado à medida que novos dados são recebidos em um fluxo contínuo, sem a necessidade de processar novamente todo o conjunto de dados. Estudos recentes já começaram a investigar essa abordagem [2, 5, 17], embora abranjam apenas uma parte dos algoritmos potenciais que podem ser utilizados, indicando que ainda há muito a ser explorado nesse campo.

Neste projeto, serão estudados e implementados diferentes algoritmos de mineração de fluxo de dados contínuos, com o objetivo de avaliar o potencial de cada um deles na identificação de ataques nas redes IoT. Para tanto, primeiramente, será realizado um levantamento de conjuntos de dados públicos que contenham tráfego de rede IoT e atendam dois requisitos: a presença de mudanças naturais de comportamento, para avaliar a capacidade de adaptação dos algoritmos, e diversidade de ataques, para compreender melhor o potencial de detecção frente a diferentes ameaças. Na sequência, os algoritmos serão implementados e testados sobre os conjuntos selecionados, buscando avaliar principalmente métricas de desempenho preditivo. Por fim, serão investigadas técnicas para diminuição da demanda por exemplos rotulados para o treinamento dos modelos de aprendizado.

2. Fundamentação Teórico-Metodológica e Estado da Arte

2.1. Internet das Coisas

A Internet das Coisas (do inglês *Internet of Things*) é um paradigma tecnológico que se refere à interconexão de dispositivos por meio da Internet. Esses dispositivos estão equipados com sensores, atuadores e tecnologia de comunicação que permitem a coleta, troca e análise de dados, possibilitando assim, a automação de tarefas e a tomada de decisões em tempo real [3].

Para entender a melhor a respeito dessa rede interconectada, é imperativo compreender sua arquitetura, que é fundamentada em camadas ou fases, que podem variar em

nome e quantidade dependendo da abordagem ou modelo específico, mas uma arquitetura típica pode incluir três camadas principais, conforme a apresentada na figura 1 [6].

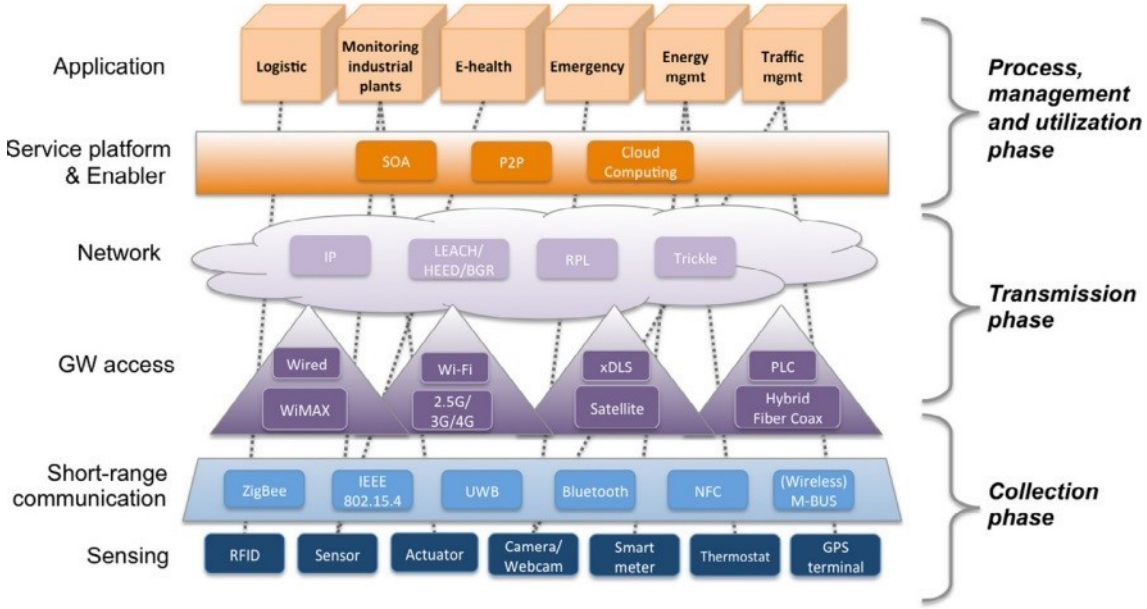

Figura 1. Representação horizontal para aplicações de IoT

As camadas são divididas da seguinte forma:

- Camada de Coleta (ou Camada de Percepção): composta pelos dispositivos sensores e atuadores distribuídos fisicamente no ambiente. Os sensores são responsáveis por captar dados essenciais como luz, temperatura, umidade e outros, em tempo real, enquanto os atuadores executam ações em resposta a essas informações, gerando uma perspectiva geral do ambiente. Essa camada serve como ponto de entrada para os dados, representando o elo entre o mundo físico e o ambiente digital.
- Camada de Transmissão (ou Camada de Rede): responsável pela comunicação e transmissão eficiente dos dados coletados pelos sensores para a camada superior. Estão inclusas nessa camada algumas tecnologias heterogêneas, que realizam métodos de endereçamento, roteamento e permitem o acesso a rede e a entrega dos dados a aplicações e servidores externos.
- Camada de Aplicação (ou Camada de gerenciamento e utilização): onde ocorre o processamento, análise e aplicação dos dados em soluções específicas, adaptadas de acordo com as diferentes necessidades. Esta camada também representa a interface com os usuários e as operações que se beneficiam com as informações provenientes dos dispositivos IoT.

2.2. Sistemas de Detecção de Intrusão

Os sistemas de detecção de intrusão (*Intrusion Detection Systems* - IDS) são ferramentas essenciais na segurança da informação, projetadas para identificar atividades suspeitas ou maliciosas em redes ou sistemas. Eles podem ser predominantemente um *software* ou um *hardware*, ou uma combinação de ambos e desempenham um papel crucial na

detecção e prevenção de intrusões, auxiliando na proteção, manutenção da integridade, confidencialidade e disponibilidade dos sistemas [12].

A arquitetura geral de um IDS inclui os seguintes componentes [12, 19]:

- Sensores: responsáveis pela coleta de dados
- Mecanismo de Análise: processa os dados coletados pelos sensores e identifica padrões suspeitos ou maliciosos.
- Base de Conhecimento: contém informações sobre padrões de ataques conhecidos e perfis de comportamento típicos. Essa base é usada para comparar os padrões identificados durante a análise.
- Resposta a Incidentes: após a detecção de um ataque, o sistema pode desencadear uma resposta, como bloquear o acesso do usuário ou realizar o envio de alertas.

Um IDS pode ser classificado em *Network-based Intrusion Detection Systems* (NIDS) ou em *Host-based Intrusion Detection Systems* (HIDS). No primeiro, a implantação ocorre no perímetro da rede e os pacotes que a atravessam são examinados em tempo real, em busca de atividades maliciosas. Já no segundo, a instalação é feita diretamente nos hospedeiros (computadores ou servidores) e monitoram as atividades e eventos que ocorrem no próprio sistema operacional, incluindo chamadas de sistema, processos em execução, entre outros. [19]

Para determinar se um ataque está de fato ocorrendo, os padrões de tráfego, atividade ou código podem ser comparados com uma base de conhecimento que contém assinaturas conhecidas de ataques previamente identificados. Esse método é conhecido como baseado em assinatura (*signature-based*) e se mostra altamente eficaz para detectar ataques bem conhecidos, visto que ao encontrar uma correspondência com a base, o evento é considerado uma intrusão [4, 19].

Em contrapartida, outra possível abordagem é o método baseado em anomalias (*anomaly based*), mais eficiente na identificação de ataques menos conhecidos, uma vez que envolve a criação de um perfil do comportamento normal do sistema ou usuário. Qualquer desvio significativo desse padrão é considerado uma anomalia e pode indicar uma possível atividade maliciosa. [4, 19].

2.3. Mineração de Fluxos Contínuos de Dados

A mineração de fluxos contínuos de dados é um campo da ciência de dados e aprendizado de máquina que lida com a análise e extração de padrões e informações relevantes a partir de dados que estão em constante evolução, chegando em grande quantidade, em alta velocidade e de forma contínua [10, 9, 11]. O objetivo é compreender e extrair informações valiosas dos dados em tempo real, permitindo tomadas de decisão rápidas e eficazes.

Nesse cenário, os modelos e algoritmos são projetados e possuem algumas características mais específicas como:

- Aprendizado Incremental: cada nova amostra é processada de forma incremental e, em seguida, é realizado o aprendizado e a atualização de suas estatísticas e parâmetros.
- Descarte ou Retenção Limitada: após o processamento de uma amostra, dependendo da aplicação, a amostra pode ser descartada para economizar recursos ou

pode ser armazenada em um histórico limitado para análises futuras ou para a atualização de modelos.
- Adaptabilidade e Mudança de Conceito: deve ser capaz de se adaptar a mudanças nas características dos dados ao longo do tempo, conhecidas como mudanças de conceito.

Além disso, existem duas principais técnicas relacionadas ao aprendizado de máquina: o supervisionado e o não-supervisionado.

No aprendizado supervisionado, os modelos são treinados utilizando um conjunto de dados rotulados, onde as entradas e suas correspondentes saídas desejadas são fornecidas. Essa técnica é frequentemente empregada em tarefas de classificação, na qual é possível categorizar novos dados em classes predefinidas, e regressão, sendo útil para prever respostas contínuas e valores numéricos.

Por outro lado, o aprendizado não supervisionado é uma técnica em que o modelo é treinado em um conjunto de dados desprovido de rótulos associados. O modelo identifica padrões e estruturas nos dados por conta própria, sem orientação externa, buscando agrupar instâncias semelhantes ou reduzir a dimensionalidade dos dados. Sendo assim, é valioso quando se utiliza abordagens de agrupamento (ou *Clustering*), no qual os dados são agrupados em conjuntos distintos, e os membros de cada grupo compartilham características comuns.

3. Objetivos

O principal objetivo deste trabalho é realizar um estudo aprofundado e a implementação prática de uma variedade de algoritmos de mineração de fluxo de dados contínuos com o propósito específico de detectar ataques em redes IoT por meio da análise do tráfego de rede. Este estudo é fundamentado em metas específicas que delineiam a direção e o escopo da pesquisa:

1. Identificar conjuntos de dados públicos que sejam representativos das redes IoT e que contenham diversos tipos de ataques conhecidos. Além disso, esses conjuntos devem apresentar mudanças naturais de comportamento ao longo do tempo, proporcionando um desafio real para os algoritmos selecionados. Essa etapa é crucial para avaliar a capacidade dos algoritmos de adaptação a situações diversas e em constante evolução.
2. Implementar os algoritmos escolhidos e a avaliar seus respectivos desempenhos preditivos. Isso implica em verificar como esses algoritmos se comportam diante das mudanças naturais no comportamento da rede e de diferentes tipos de ataques. A análise comparativa entre os algoritmos permitirá determinar quais deles são mais eficazes em identificar e responder a ameaças em um ambiente IoT dinâmico.
3. Investigar estratégias e técnicas que possam reduzir ou eliminar a necessidade de exemplos rotulados no treinamento dos algoritmos. Isso é fundamental, uma vez que a obtenção de dados rotulados pode ser custosa e trabalhosa, e a escassez desses exemplos pode limitar a eficácia dos modelos de aprendizado. Portanto, esta etapa busca aprimorar a eficiência e a praticidade dos algoritmos de detecção de ataques em redes IoT.

4. Procedimentos metodológicos/Métodos e técnicas

Inicialmente, será realizada uma revisão bibliográfica a fim de estabelecer uma base sólida de conhecimento com foco na mineração de fluxos contínuos de dados e como ela pode ser utilizada para a detecção de intrusões em redes IoT. Após esse mapeamento envolvendo as mais recentes abordagens, técnicas e algoritmos empregados, será possível identificar os principais desafios enfrentados atualmente e as soluções mais promissoras disponíveis na literatura, fornecendo assim um alicerce sólido para a implementação de um método eficaz de detecção de ataques em fluxos contínuos de dados.

O próximo passo é fazer a escolha de conjuntos de dados públicos que sejam pertinentes e representativos ao experimento, isto é, devem apresentar características como diferentes tipos de ataques e variações naturais de comportamento na rede ao longo do tempo. Posteriormente, será feito o pré-processamento dos dados, isso inclui: conversões de arquivos PCAP para o formato CSV, extração dos fluxos e de recursos específicos, rotulação dos ataques e outros tratamentos que se mostrarem necessários.

Em seguida, será realizada a seleção dos algoritmos de aprendizado que melhor se encaixam no contexto dos fluxos contínuos de dados. É fundamental considerar a natureza dinâmica e constante evolução dos dados, optando por algoritmos que possuam capacidade de processamento incremental, permitindo a atualização contínua do modelo à medida que novos dados chegam.

Uma vez implementados, os algoritmos serão conduzidos a experimentos práticos para avaliação de suas eficácias, através de métricas de desempenho que serão coletadas, incluindo *precision, recall* e *F1-score*, provenientes de matrizes de confusão. Utilizando exemplos rotulados, serão realizados testes para detecção de ataques em tempo real nos fluxos de dados.

Por fim, será investigada a possibilidade de reduzir ou eliminar a dependência de dados rotulados durante o treinamento dos algoritmos. Isso pode envolver técnicas de aprendizado semi-supervisionado ou não supervisionado, visando utilizar de maneira mais eficaz os dados disponíveis.

Após a realização de novos testes, será feita a análise dos resultados, a comparação da eficácia dos modelos e a conclusão em relação ao uso da mineração de fluxos de dados contínuos na detecção de ataques em redes IoT.

5. Cronograma de Execução

Atividades:

1. Revisão bibliográfica com foco em mineiração de fluxo de dados contínuos;
2. Escolha e preparação do conjunto de dados;
3. Implementação dos algoritmos selecionados;
4. Realização de testes e coleta de métricas para avaliar os algoritmos
5. Estudar e implementar técnicas para diminuição do uso de exemplos rotulados;
6. Análise e comparação dos resultados obtidos;
7. Escrita do Trabalho de Conclusão de Curso

Tabela 1. Cronograma de Execução

	set	out	nov	dez	jan	fev	mar	abr
Atividade 1	X	X						
Atividade 2		X						
Atividade 3			X	X				
Atividade 4				X	X			
Atividade 5					X	X		
Atividade 6							X	
Atividade 7				X	X	X	X	X

6. Contribuições e/ou Resultados esperados

Os resultados esperados deste projeto incluem a avaliação da eficácia dos algoritmos de mineração de fluxo de dados contínuos supervisionados na detecção de ameaças cibernéticas. Além disso, pretende-se propor novas soluções que não apenas minimizem a necessidade do uso de rótulos, mas também tenham uma capacidade maior de se adaptar a mudanças naturais no comportamento das redes de comunicação, diminuindo assim, a carga de trabalho dos operadores humanos.

7. Espaço para assinaturas

Londrina, *18 de setembro de 2023*.

_____ _____

Aluno Orientador

Referências

[1] Mohamed Abomhara and Geir M. Køien. Security and privacy in the internet of things: Current status and open issues. In *2014 International Conference on Privacy and Security in Mobile Systems (PRISMS)*, pages 1–8, 2014.

[2] Gustavo Vitral Arbex, Kétly Gonçalves Machado, Michele Nogueira, Daniel M. Batista, and Roberto Hirata. Iot ddos detection based on stream learning. In *2021 12th International Conference on Network of the Future (NoF)*, pages 1–8, 2021.

[3] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: A survey. *Computer Networks*, 54(15):2787–2805, 2010.

[4] Stefan Axelsson. Intrusion detection systems: A survey and taxonomy. 2000.

[5] Sylvia Worlali Azumah, Nelly Elsayed, Victor Adewopo, Zaghloul Saad Zaghloul, and Chengcheng Li. A deep lstm based approach for intrusion detection iot devices network in smart home. In *2021 IEEE 7th World Forum on Internet of Things (WF-IoT)*, pages 836–841, 2021.

[6] Eleonora Borgia. The internet of things vision: Key features, applications and open issues. *Computer Communications*, 54:1–31, 2014.

[7] Gregory Ditzler, Manuel Roveri, Cesare Alippi, and Robi Polikar. Learning in nonstationary environments: A survey. *IEEE Computational Intelligence Magazine*, 10(4):12–25, 2015.

[8] Mario Frustaci, Pasquale Pace, Gianluca Aloi, and Giancarlo Fortino. Evaluating critical security issues of the iot world: Present and future challenges. *IEEE Internet of Things Journal*, 5(4):2483–2495, 2018.

[9] Mohamed Medhat Gaber, Arkady Zaslavsky, and Shonali Krishnaswamy. Mining data streams: A review. 34(2), 2005.

[10] Geoff Hulten, Laurie Spencer, and Pedro Domingos. Mining time-changing data streams. New York, NY, USA, 2001. Association for Computing Machinery.

[11] Jeff Ullman Jure Leskovec, Anand Rajaraman. *Mining of Massive Datasets*. Cambridge University Press, 2021.

[12] Aleksandar Lazarevic, Vipin Kumar, and Jaideep Srivastava. *Intrusion Detection: A Survey*, volume 5, pages 19–78. 01 2005.

[13] Zhipeng Liu, Niraj Thapa, Addison Shaver, Kaushik Roy, Xiaohong Yuan, and Sajad Khorsandroo. Anomaly detection on iot network intrusion using machine learning. In *2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, pages 1–5, 2020.

[14] Alexandra L'Heureux, Katarina Grolinger, Hany F. Elyamany, and Miriam A. M. Capretz. Machine learning with big data: Challenges and approaches. *IEEE Access*, 5:7776–7797, 2017.

[15] Rwan Mahmoud, Tasneem Yousuf, Fadi Aloul, and Imran Zualkernan. Internet of things (iot) security: Current status, challenges and prospective measures. In *2015 10th International Conference for Internet Technology and Secured Transactions (ICITST)*, pages 336–341, 2015.

[16] Pascal Maniriho, Ephrem Niyigaba, Zephanie Bizimana, Valens Twiringiyimana, Leki Jovial Mahoro, and Tohari Ahmad. Anomaly-based intrusion detection approach for iot networks using machine learning. In *2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, pages 303–308, 2020.

[17] Fernando H. Y. Nakagawa, Sylvio Barbon Junior, and Bruno Bogaz Zarpelão. Attack detection in smart home iot networks using clustream and page-hinkley test. In *2021 IEEE Latin-American Conference on Communications (LATINCOM)*, pages 1–6, 2021.

[18] Bambang Susilo and Riri Fitri Sari. Intrusion detection in software defined network using deep learning approach. In *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, pages 0807–0812, 2021.

[19] Bruno Bogaz Zarpelão, Rodrigo Sanches Miani, Cláudio Toshio Kawakani, and Sean Carlisto de Alvarenga. A survey of intrusion detection in internet of things. *Journal of Network and Computer Applications*, 84:25–37, 2017.

[20] Zhi-Kai Zhang, Michael Cheng Yi Cho, Chia-Wei Wang, Chia-Wei Hsu, Chong-Kuan Chen, and Shiuhpyng Shieh. Iot security: Ongoing challenges and research opportunities. In *2014 IEEE 7th International Conference on Service-Oriented Computing and Applications*, pages 230–234, 2014.