Impacto da Anonimização do Tráfego em Redes na Identificação de Dispositivos e na Detecção de Anomalias

Ariel L. C. Portela[1], **Wanderson L. Costa**[1], **Rafael A. Menezes**[1], **Rafael L. Gomes**[1]

[1]Universidade Estadual do Ceará (UECE), Fortaleza, Ceará, Brasil.

ariel.portela@aluno.uece.br, wanderson.leonardo@ifpi.edu.br,

rafael.menezes@larces.uece.br, rafa.lopes@uece.br

Abstract. *Currently, a crucial aspect of networking management is monitoring it's traffic, where Machine Learning (ML) comes in. ML has been used to perform various tasks, such as loT device identification and detection of network anomalies. However, access to information about network traffic can affect users' privacy, thus violating existing privacy laws. Within this context, this article analyzes the impact of network traffic anonymization, ensuring privacy when identifying the device and detecting anomaly solutions by basing on feature selection techniques. The carried out experiments, used a real dataset, which results showed that when using the selection and ML techniques combined, the anonymization of traffic reduces the identification capacity. In addition, using those techniques also preserve user's privacy while maintaining the detection capacity of network anomalies.*

Resumo. *Atualmente, um aspecto crucial para o gerenciamento de redes é o monitoramento de tráfego de rede, onde técnicas de Aprendizagem de Máquina (ML) têm sido usadas sobre esses dados a fim de realizar diversas tarefas, como por exemplo identificação de dispositivos IoT e detecção de anomalias de rede. Contudo, o acesso a informações sobre o tráfego de rede pode afetar a privacidade dos usuários, ferindo assim as leis de privacidade existentes. Dentro deste contexto, este artigo analisa o impacto da anonimização de tráfego de rede, para garantir privacidade, sobre essas soluções de identificação de dispositivos e detecção de anomalias, a partir de técnicas de seleção de características. Os experimentos realizados utilizaram um conjunto de dados real, onde os resultados mostram que, quando utilizadas as técnicas de seleção e ML combinadas, a anonimização do tráfego reduz a capacidade de identificação, preservando assim a privacidade dos usuários, enquanto que mantêm a capacidade de detecção de anomalias de rede.*

1. Introdução

A sociedade vem evoluindo seu tradicional paradigma de comunicação baseado em chamadas de voz e mensagens de texto, agregando mais aplicações de interação multimídia e/ou compartilhando informações em redes sociais. Desta forma, a Internet emerge como o principal caminho para os serviços computacionais modernos (por exemplo, compartilhamento de conteúdo, sistemas inteligentes, automação de tarefas e outros), tornando-se crucial para a sociedade moderna.

Assim, um aspecto importante é entender o comportamento da rede e identificar o perfil de tráfego passante. A fim de realizar dinamicamente a definição de um perfil de comportamento, bem como a detecção de anomalias, são aplicadas técnicas de Inteligência Artificial (IA) sobre os dados de rede monitorados. Estas ferramentas possibilitam a construção, teste e implantação de soluções a partir de modelos gerados e aprimorados, permitindo a criação de estratégias eficientes e com alta precisão. Sendo assim, pode-se utilizar as características de tráfego dos dispositivos presentes nos ambientes a fim de gerar o perfil de comportamento da rede, usando diversas técnicas (e.g., clusterização, redes neurais e máquinas de vetor de suporte) a partir de diferentes abordagens: combinação de aprendizado supervisionado e não-supervisionado, cascata de técnicas supervisionadas ou técnicas específicas individualmente.

Contudo, há o acesso a informações sobre o tráfego de rede que pode afetar a privacidade dos usuários, ferindo assim as leis de privacidade, como por exemplo a Lei Geral de Proteção de Dados (LGPD - 13709/2018) no Brasil e General Data Protection Regulation (GDPR - 2016/679) na Europa. Estas leis têm como principal objetivo legitimar a garantia de proteção dos dados, os quais estão sendo fornecidos pelos usuários e utilizados pelos provedores de serviço, classificando então, o grau de sensibilidade da informação referente à cada indivíduo onde, quanto maior o impacto da exposição de um determinado dado, maior a sensibilidade deste. Consequentemente, tem-se um cenário de empresas que necessitam lidar com a integridade, confidencialidade, disponibilidade e autenticidade dos dados e serviços. Com isso, torna-se crucial validar os critérios de segurança da informação, dados estes que são denominados sensíveis, em consequência, várias medidas de segurança vêm sendo estudadas e aplicadas nos últimos anos em vários ambientes. A partir disso, faz-se necessário ter uma camada de proteção dos dados. Uma das funcionalidades necessárias para esta camada seria aplicar técnicas de anonimização de dados. Contudo, as técnicas de anonimização existentes atendem diferentes níveis de anonimização, que podem alterar o contexto dos dados impossibilitando a aplicação de técnicas de IA para identificação de perfil de tráfego.

Um dos pontos que existem em aberto é a identificação de dispositivos IoT, sendo que este ocorre a partir das informações do tráfego de rede usando técnicas de aprendizagem de máquina (*Machine Learning* - ML). No geral, as seguintes tarefas são executadas para realizar uma identificação: extrair as características do tráfego de rede, treinar o modelo de ML e realizar a identificação em si. Portanto, a identificação de dispositivos IoT pode afetar diretamente os aspectos de privacidade dos usuários. Apesar das preocupações relacionadas a privacidade, ainda faz-se necessário ter soluções de segurança em relação ao tráfego de rede, como por exemplo a detecção de anomalias, que é a base das ferramentas de Detecção de Intrusão (IDS).

Dentro deste contexto, este artigo analisa a capacidade de identificação de dispositivos IoT pelas técnicas de ML e o impacto da anonimização do tráfego de rede nesta identificação, preservando assim os aspectos de privacidade dos usuários. Adicionalmente, será analisada a capacidade de detecção de anomalias de redes com as características sobre o tráfego depois da anonimização. Assim, o trabalho realizado possibilitou identificar como que as técnicas de aprendizagem de máquina, usando características que podem ser extraídas do monitoramento da rede, podem afetar os aspectos de privacidade dos usuários de dispositivos IoT enquanto preservam o nível de segurança da rede ao

detectar anomalias.

Os experimentos realizados utilizaram um conjunto de dados real de dispositivos IoT, onde os resultados mostram que a anonimização do tráfego reduz a capacidade de identificação, preservando assim a privacidade dos usuários. Enquanto que pode manter um alto nível de detecção de anomalias de rede. No geral, os resultados mostraram que a técnica de seleção Lasso apresenta uma maior adequação a este contexto, visto que minimiza a identificação de dispositivos, enquanto que maximiza a detecção de anomalias.

O restante deste artigo está organizado da seguinte forma: A Seção 2 apresenta os trabalhos relacionados a análise de tráfego em redes IoT e técnicas de Anonimização. A Seção 3 descreve o trabalho realizado, enquanto que a Seção 4 descreve a configuração dos experimentos e analisa os resultados obtidos; Por fim, a Seção 5 apresenta a conclusão e os trabalhos futuros.

2. Trabalhos Relacionados

Sharafaldin at. al [Sharafaldin et al. 2019] apresentam um estudo sobre as características do tráfego de rede mais importantes para detecção de diferentes tipos de ataques DDoS em redes tradicionais, ou seja, redes TCP/IP. Nos experimentos realizados foram projetadas e implantadas duas redes com computadores tradicionais, ou seja, o comportamento extraído das amostras do conjunto de dados se torna diferente em comparação com o de redes projetadas com dispositivos IoT. O comportamento de redes IoT se comunica com um pequeno conjunto finito de pontos de extremidade e são propensos a ter padrões de tráfego de rede repetitivos (pacotes pequenos em intervalos de tempo fixos para fins de registro, por exemplo).

Clarke et al. [Clarke et al. 2017] propõem uma ferramenta de identificação de usuários que utiliza somente metadados com relação ao tráfego de rede monitorado, ou seja, não utiliza nenhuma informação de *payload*. O objetivo dos autores é reduzir o volume de dados a serem analisados a fim de identificar os usuários da rede. Similarmente, Li et. al [Li et al. 2018] apresentam um aprendizado profundo para IoTs no ambiente de computação de borda. Como os nós da borda existentes possuem capacidade de processamento limitada, também foi projetada uma nova estratégia de descarregamento para otimizar o desempenho de aplicativos de aprendizado profundo de IoT com computação de borda. Na avaliação de desempenho, foi testado o desempenho da execução de várias tarefas de aprendizado profundo em um ambiente de computação de ponta com nossa estratégia.

Ruoming et. al [Pang 2016] fazem uma medição de anonimização de rede incluindo os rastros de pacote, e propuseram uma ferramenta chamada tcpmkpub, para anonimizar rastros, e realizaram uma discussão sobre as políticas de anonimização onde descreveram o uso de metadados acompanhados por rastros de rede. Similarmente, a empresa CAIDA [CAIDA 2020] realizou um mapeamento das camadas de rede, e propôs os algoritmos de anonimização mais usados para determinados contextos, enfatizando a anonimização de endereço IP. Foi feito também uma segregação em camadas do cenário que compõe a internet apontando diferentes formas de como conseguir aplicar os conceitos de segurança da informação.

Embora vários estudos tenham sido realizados com a finalidade de analisar tráfego de redes e dispositivos IoT, nenhum desses trabalhos acima abordam um cenário realístico

de uma infraestrutura heterogênea em uma rede IoT avaliando aspectos de privacidade sobre os dispositivos IoT. Sendo assim, este trabalho avança o estado da arte no que se refere a soluções de segurança para redes IoT, focando na detecção desses dispositivos através da análise das características de tráfego e, dessa forma, preservando aspectos de privacidade, com isso respeitando a LGPD.

3. Proposta

Neste trabalho foram treinados dois modelos de ML distintos para realizar duas tarefas: Identificação de dispositivos IoT na Rede e Detecção de Anomalias de rede. O modelo para identificação visa determinar quais dispositivos IoT estão na rede, podendo assim ferir aspectos de privacidade. Por outro lado, o modelo de detecção atua como parte de uma solução de segurança de rede. Portanto, a avaliação consistiu nas seguintes etapas: (I) Monitoramento da Rede; Anonimização dos Dados Monitorados; (III) Extração de características de fluxo; (IV) Seleção de características; (V) Treinamento de Modelo;(VI) Detecção de Anomalias de rede; e, (VII) Identificação do dispositivo. A seguir, serão descritas as etapas mencionadas, detalhando suas particularidades, bem como o papel de cada uma.

3.1. Anonimização

O processo de anonimização consiste em alterar o conteúdo de algo, sem alterar a essência da informação. Com relação ao tráfego de rede, existem duas possibilidades de ferramentas: NFanon e CryptoPanWrapper. NFanon é uma ferramenta usada para anonimizar todos os endereços IPs (src, dst, próximo salto, IP do roteador etc) no registro de fluxo de rede usando o CryptoPAn (Cryptograpy-based prefix preserving Anonymization). A chave -K é usada para inicializar a cifra Rijndael. A chave é uma string de 32 caracteres ou uma string de 64 dígitos hexadecimais começando com 0x. Possui diversos modos de operação, porém, por estar desatualizada e sem manutenção há alguns anos, gerou inúmeros problemas de incompatibilidade. Similarmente, CryptoPanWrapper é uma biblioteca em Python que utiliza o CryptoPan (Cryptograpy-based prefix preserving Anonymization) e permite a anonimização de IPs, onde os endereços IPs são mapeados através de encriptação AES. Tecnicamente, o Crypto-PAn faz uma pseudonimização de endereços IP. Como cada endereço IP é mapeado exclusivamente (1-1) para outro endereço IP (via criptografia AES). A anonimização seria o processo de mapeamento de vários endereços IP para o mesmo endereço IP de destino, tornando efetivamente os IPs originais indistinguíveis (não mapeáveis de forma reversa).

Sendo assim, ambas as ferramentas atendem as exigências da LGPD no contexto de tráfego, porém, por conta dos diversos problemas com a NFanon, optamos por dar continuidade com a CryptoPanWrapper. A identificação de dispositivos IoT específicos, fere aspectos de privacidade. Então dentre as características selecionadas, verificamos quais delas são consideradas sensíveis. Chegamos a conclusão de que os IPs de origiem e destino, podem ser considerados sensíveis. Portanto, anonimizamos os IPs usando a ferramenta CryptoPanWrapper. Adicionalmente, foi retirada qualquer informação sobre o conteúdo dos payloads dos pacotes monitorados.

3.2. Extração e Seleção de Características

A partir da anonimização dos dados sobre monitoramento da rede é possível gerar um conjunto de dados em um formato PCAP, o qual permite a extração de até 80 (oi-

tenta) características de fluxos de rede (usando por exemplo a ferramenta CICFlowMeter [Sharafaldin et al. 2019]). Todavia, o uso de todas essas características pode gerar ruídos, ainda mais quando se trata de ambientes com alta variedade de tipos de dispositivos. Além disso, quando somente as características mais relevantes são consideradas, pode-se minimizar no tempo de treinamento do modelo e reduzir a demanda por recursos computacionais (processamento mais rápido, menor consumo de memória e menor espaço de armazenamento), devido a redução de dimensionalidade do problema. Portanto, o objetivo da seleção de características é habilitar a construção de modelos de ML que viabilizem entender os dados e maximizar a capacidade de detecção e identificação, ajudando a conhecer atributos irrelevantes e redundantes que podem ter impacto negativo no desempenho do modelo diminuindo a acurácia do modelo.

A partir desta realidade, este artigo analisa as seguintes técnicas de seleção de características:

- Extra-Árvore (EA) [Geurts et al. 2006]: O algoritmo Extra-Árvore constrói um conjunto de árvores de decisão ou regressão não podadas de acordo com o procedimento clássico de cima para baixo. Suas duas principais diferenças com outros métodos de conjuntos baseados em árvores são que ele divide os nós escolhendo pontos de corte completamente aleatoriamente e usa toda a amostra de aprendizado para cultivar as árvores. Essa classe implementa um meta-estimador que se encaixa em várias árvores de decisão aleatórias em várias subamostras do conjunto de dados e usa a média para melhorar a precisão preditiva e controlar o ajuste excessivo.
- Lasso [Friedman et al. 2010]: O Lasso é um modelo linear que estima coeficientes esparsos. Este é comumente aplicado em alguns contextos devido à sua tendência de preferir soluções com menos coeficientes diferentes de zero, reduzindo efetivamente o número de características dos quais a solução fornecida depende. Portanto, ele consiste em um modelo linear com um termo de regularização adicionado.

Para cada uma dessas técnicas utilizadas nos experimentos, um conjunto de características são escolhidas como sendo as mais relevantes, logo, cada técnica, de acordo com seu algoritmo de escolha de característica relevantes, extrai os atributos que julga serem melhores. Dessa forma, selecionamos as características presentes na interseção indicada pelas técnicas.

3.3. Treinamento do modelo

Após a seleção das características mais adequadas, é iniciada a fase de treinamento do modelo de ML. O treinamento do modelo de ML engloba a recepção dos dados sobre o tráfego de rede e execução da técnica de ML. Cada técnica de ML aplica uma abordagem distinta para compreender os dados. Neste artigo foram avaliadas as seguintes técnicas que possuem estratégias distintas:

- Nearest Neighbor (KNN) [Hwang and Wen 1998]: O KNN é um dos algoritmos não-paramétricos mais importantes no campo de reconhecimento de padrões, sendo um algoritmo de classificação de aprendizado supervisionado. As regras de classificação do KNN são geradas pelas próprias amostras de treinamento sem nenhum dado adicional. O algoritmo de classificação KNN prevê a categoria da

amostra de teste de acordo com as amostras de treinamento que são os vizinhos mais próximos da amostra de teste, e a julga para aquela categoria que possui a maior probabilidade de categoria. O vizinho mais próximo refere-se ao vetor de característica multidimensional que é usado para descrever a amostra mais próxima, e o critério mais próximo pode ser a distância euclidiana do vetor de característica.

- Regressão logística (RL) [Meurer and Tolles 2017]: Regressão logística é um modelo de ML usado para prever a probabilidade de ocorrência de um evento em face de um conjunto de variáveis explanatórias. A função logística, também conhecida como função sigmoide, é usada para calcular o modelo logístico no qual cada valor do infinito negativo ao infinito positivo é fornecido como entrada e saída limitadas no intervalo de 0 e 1. Este algoritmo pode entender variáveis vetoriais e avaliar os coeficientes ou pesos para cada variável de entrada e, em seguida, prever que a classe expressou o valor do vetor de palavras.
- Multi-layer Perceptron (MLP) [Alanis et al. 2019]: um algoritmo de aprendizado supervisionado que aprende uma função $f(.) : R^m \rightarrow R^o$ treinando em um conjunto de dados, onde m é o número de dimensões para entrada e o é o número de dimensões para saída. Diferentemente da regressão logística, entre a camada de entrada e a de saída pode haver uma ou mais camadas não lineares, chamadas de camadas ocultas. As vantagens do MLP são sua capacidade de aprender modelos não lineares, bem como a capacidade de aprender modelos em tempo real (aprendizado on-line) usando partial_fit.

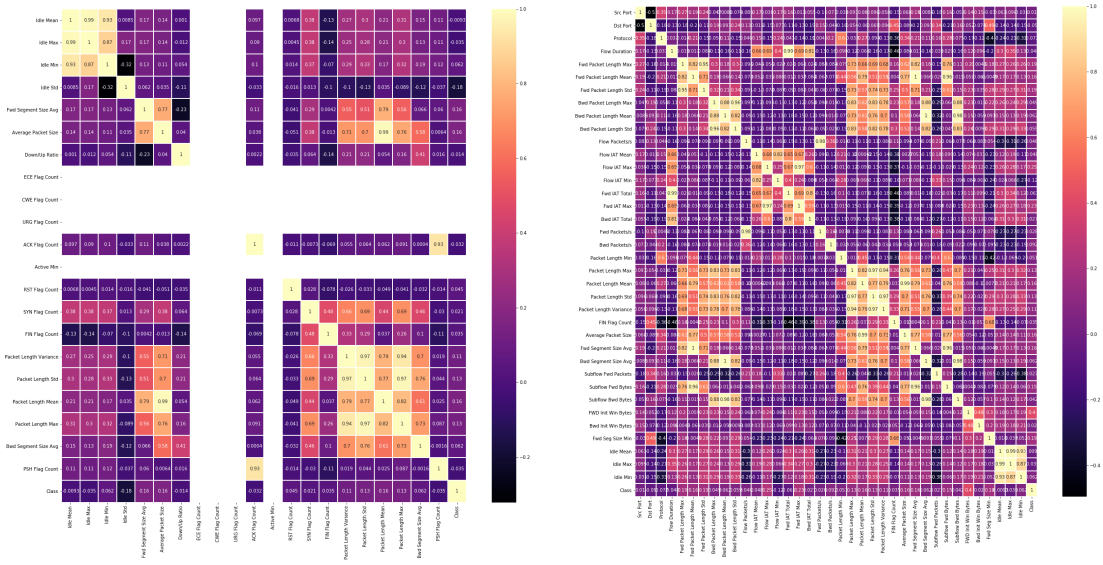
Como dito anteriormente, foram treinados dois modelos de ML distintos para identificar dispositivos IoT e detectar de anomalias no tráfego de rede. Assim, podem ser utilizados qualquer uma das técnicas citadas, em conjunto com as técnicas de seleção descritas, para realizar as atividades previstas.

4. Resultados

Os experimentos foram realizados utilizando um conjunto de dados de rede dentre os disponibilizados por diversas universidades e instituições ao redor do mundo, no formato PCAP. Os arquivos PCAPs contém vários dados de rede, como: tamanho do cabeçalho, tipo de serviços, tamanho total do pacote, endereço ip de origem, destino e sua respectiva versão (Ipv4/Ipv6), porta de origem e destino, protocolo de pacotes, FLAGs, endereço MAC, payload e outros. Os experimentos foram baseados no conjunto de dados "UNSW-IoT", criado por [Sivanathan et al. 2018], que tem tráfego normal (benigno) de IoT e dispositivos pessoais.

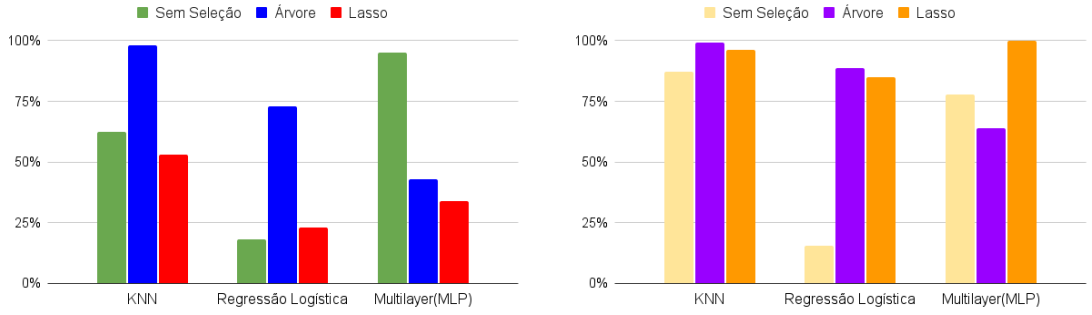
A Figura 1 tem-se a correlação entre as características de tráfego selecionadas por cada técnica de seleção (Lasso e EA), onde destaca-se a importância das características selecionadas (cores mais claras). Similarmente, avaliou-se, tanto para a detecção de anomalias, quanto para a identificação de dispositivos, a Acurácia de cada um dos modelos de ML descritos em conjunto com as técnicas de seleção, proporcionando assim uma ampla avaliação da capacidade de cada combinação possível. É válido ressaltar que devido a limitação de espaço, outras métricas não foram avaliadas.

Os resultados dos experimentos destacam a importância da seleção de características para a acurácia, tanto na detecção de anomalias, quanto na identificação de



(a) Lasso. (b) Extra-Árvore.

Figura 1. Correlação entre as características selecionadas.



(a) Acurácia da Identificação de Dispositivos. (b) Acurácia da Detecção de Anomalias.

Figura 2. Resultados.

dispositivos IoT. Por exemplo, usando a técnica de seleção mais apropriada, o desempenho dos classificadores KNN e RL para a identificação podem variar de 35% a 45% e 50% a 55%, respectivamente, como mostra a Figura 2(a). Além disso, a técnica RL usando os 80 atributos extraídos (sem seleção) tem uma acurácia ineficaz, enquanto usando a técnica Extra-Árvore, ela atinge mais de 72% de acurácia. Um comportamento similar é percebido com relação a detecção de anomalias, onde o desempenho de cada combinação varia devido as diferentes abordagens usadas nas técnicas de seleção e nos modelos de ML. Como mostrado na Figura 2(b), utilizando Árvore com KNN, por exemplo, temos uma acurácia de 99%, o que a torna uma ótima opção. Porém, comparada a seleção com Lasso, os resultados mostram que há um melhor aproveitamento, considerando diferentes técnicas de ML, pois obtivemos uma acurácia superior a 80% em todas as técnicas, diferente da Árvore, que mostrou uma acurácia um pouco reduzida com MLP. Nota-se também, que novamente o processo de RL, utilizando os atributos sem seleção, possui uma acurácia ineficiente. Portanto, percebe-se que há técnicas de seleção que são mais compatíveis com certos modelos de ML.

A partir dos resultados apresentados, pode-se verificar que a precisão das técnicas de ML varia de acordo com a técnica de seleção aplicada, principalmente quando essas técnicas de ML são baseadas em abordagens que focam na dimensionalidade, como os classificadores KNN e RL. Sendo assim, a técnica de seleção Lasso apresenta uma maior adequação a este contexto, visto que minimiza a identificação de dispositivos, enquanto que maximiza a detecção de anomalias de rede. Além disso, percebe-se que apesar da anonimização minimizar, a identificação dos perfis de tráfego ainda foi possível.

5. Conclusão e Trabalhos Futuros

Este artigo apresentou uma análise em relação ao impacto da anonimização de dados sobre o tráfego de rede monitorado sobre as tarefas de identificação de dispositivos IoT que pode ferir aspectos de privacidade, bem como a detecção de anomalias de rede usada para soluções de segurança de redes. Os resultados mostraram que a técnica de seleção Lasso apresenta uma maior adequação a este contexto, visto que minimiza a identificação de dispositivos, enquanto que maximiza a detecção de anomalias. Como trabalhos futuros, pretendemos analisar experimentalmente outras métricas de avaliação, bem como outras técnicas de seleção e ML.

Referências

Alanis, A. Y., Arana-Daniel, N., and Lopez-Franco, C. (2019). *Artificial neural networks for engineering applications*. Academic Press.

CAIDA (2020). Summary of anonymization best practice techniques. `https://www.caida.org/`.

Clarke, N., Li, F., and Furnell, S. (2017). A novel privacy preserving user identification approach for network traffic. *Computers Security*, 70:335–350.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software*, 33(1):1.

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Machine learning*, 63(1):3–42.

Hwang, W.-J. and Wen, K.-W. (1998). Fast knn classification algorithm based on partial distance search. *Electronics letters*, 34(21):2062–2063.

Li, H., Ota, K., and Dong, M. (2018). Learning iot in edge: deep learning for the internet of things with edge computing. *IEEE Network*, 32(1):96–101.

Meurer, W. J. and Tolles, J. (2017). Logistic regression diagnostics: understanding how well a model predicts outcomes. *Jama*, 317(10):1068–1069.

Pang, R. (2016). The devil and packet trace anonymization. *Computer Communication Review*, 36(1):29–38.

Sharafaldin, I., Lashkari, A. H., Hakak, S., and Ghorbani, A. A. (2019). Developing realistic distributed denial of service (ddos) attack dataset and taxonomy. In *2019 International Carnahan Conference on Security Technology (ICCST)*, pages 1–8. IEEE.

Sivanathan, A., Gharakheili, H. H., Loi, F., Radford, A., Wijenayake, C., Vishwanath, A., and Sivaraman, V. (2018). Classifying iot devices in smart environments using network traffic characteristics. *IEEE Transactions on Mobile Computing*, 18(8):1745–1759.