Estudos para o desenvolvimento de um método de Classificação de Tráfego Cifrado em Redes Utilizando Machine Learning

Dyonatha Batista Kramer, Charles Neu

Universidade de Santa Cruz do Sul (UNISC)
Caixa Postal 96.815900 – Santa Cruz do Sul – RS – Brasil

dyonatha@mx2.unisc.br,charles1@unisc.br

Abstract. *The growing encrypted traffic reflects strong investments in information security. However, by offering more security on the Internet, network administrators create obstacles in management, monitoring control and security. For this reason, the characterization of encrypted traffic becomes an increasingly relevant topic of research, looking for solutions that avoid exposing the user's personal data, but that, at the same time, can offer network monitoring control. In order to help deal with these issues, this article proposes a system based on machine learning using statistical information to classify encrypted network traffic.*

Resumo. *O aumento do tráfego criptografado é reflexo de fortes investimentos em segurança da informação. Todavia, ao se oferecer mais segurança na Internet, cria-se, para os administradores de rede, obstáculos na gerência, controle de monitoramento e segurança da mesma. Por isso, a caracterização de tráfego cifrado se torna um tema cada vez mais pertinente de pesquisa, buscando soluções que evitem expor dados pessoais do usuário, mas que, ao mesmo tempo, possam oferecer um controle de monitoramento de rede. A fim de ajudar a lidar com essas questões, este artigo propõe um sistema baseado em machine learning utilizando informações estatísticas para classificar tráfego cifrado em rede.*

1. Introdução

A classificação de tráfego em redes é um assunto sempre novo nas áreas de redes e segurança. Algumas das principais razões de sua importância são, por exemplo, a criação de regras de controle e monitoramento que limitam a largura de banda consumida por algum aplicativo ou usuário e fornecem informações a respeito do que acontece em cada computador, a priorização de tráfego de acordo com as regras de negócio e a detecção de ataques maliciosos, assim como seu bloqueio.

A classificação de tráfego, usualmente, é feita a partir de algum algoritmo estatístico que combina protocolos conhecidos, números de portas e atributos específicos de aplicativos contidos no campo de dados (carga útil) dos pacotes. Entretanto, considerando que em uma rede transita um número cada vez maior de pacotes e que muitos deles são encriptados, torna-se praticamente impossível a identificação de seus atributos específicos a partir do campo de dados [Bar-Yanai et al. 2010].

Criptografia, em linhas gerais, pode ser definida como um conjunto de princípios e técnicas empregados para cifrar dados [Stallings 2015]. Ou seja, a

privacidade e segurança online são dela dependentes, pois são usadas como padrão para a troca de mensagens e dados de maneira segura na Internet. A tendência é que o percentual de tráfego global criptografado se mantenha alto devido à adoção da tecnologia como padrão de projeto [Cisco 2019]. Todavia, a criptografia também pode ser utilizada para comprometer a segurança de um sistema, como por exemplo, sendo aplicada por uma ferramenta de ataque malicioso, que irá cifrar o malware, impedindo que este seja facilmente detectado por um sistema convencional de segurança, além de consumir muitos recursos de hardware durante a análise, gerando um gargalo ou baixo desempenho no sistema [Neu 2019].

Portanto, uma abordagem interessante para a resolução deste problema é o uso de *machine learning* que, diferentemente de ferramentas usuais, hoje já consideradas antiquadas, pode identificar com maior precisão o tipo de tráfego que se tem em uma rede. Diversos algoritmos e métodos de *machine learning* tem sido estudados e aplicados na prática [Zhu et al. 2019], [Singhal, Mathur e Vyas 2013], diante disso, o tema de pesquisa propõe, com o uso desta tecnologia, fazer a classificação do tráfego cifrado de rede, informando a classe a qual pertence um fluxo de pacotes. Dessa maneira, será possível que o administrador ou responsável pela rede crie regras de qualidade de serviço, faça monitoramento, gere relatórios estatísticos e muito mais.

2. Fundamentação teórica

Na computação, a criptografia é implementada a partir de algoritmos, cada qual com sua finalidade e robustez, sendo alguns deles mais simples, como o *Data Encryption Standard* (DES), ou mais complexos como o *Advanced Encryption Standard* (AES). Além do algoritmo de encriptação, se faz necessária uma chave criptográfica que faz e desfaz a cifra, com a finalidade de esconder o texto ou decifrá-lo [Stallings 2015]. A partir deste contexto, torna-se complicada a classificação de um pacote, ou de um fluxo de pacotes, pois as propriedades que melhor o caracterizam estão cifradas, sendo necessário o uso de inteligência artificial para fazer este trabalho [Singhal, Mathur e Vyas 2013].

A inteligência artificial é composta por várias áreas, sendo o *machine learning* uma delas - que possui os resultados mais robustos em pesquisas relacionadas à classificação de tráfego em redes [Cossetti 2019] - e, portanto, a que utilizaremos neste trabalho. Geralmente descrevem-se cinco tipos de *machine learning*: i) Aprendizado Supervisionado; ii) Aprendizado Não-Supervisionado; iii) Aprendizado Semi Supervisionado; iv) Aprendizado por Reforço; v) Aprendizado Profundo.

O objetivo do aprendizado supervisionado é extrair conhecimento de exemplos rotulados, previamente, com classes, com a finalidade de predizer a qual classe pertencem novos exemplos expostos à máquina [Pila 2001]. No Aprendizado Não-Supervisionado não se tem informações antecedentes da classe, mas apenas dos atributos que compõem os dados de entrada (exemplos). Sendo assim, segundo Pila, seu o objetivo é firmar agrupamentos de padrões que são semelhantes e identificar potenciais classes nos exemplos de entrada [Pila 2001]. O Aprendizado Semi-Supervisionado é a combinação dos dois métodos apresentados anteriormente e possuindo alguns exemplos rotulados e outros não. É útil principalmente quando a rotulação dos exemplos é laborioso, exige muito tempo ou pode despender muito

dinheiro. O Aprendizado por Reforço não possui um conjunto de treinamento associado, nem sequer exemplos rotulados. O procedimento aqui se baseia na forma com que os algoritmos agem em um determinado ambiente, buscando maximizar alguma noção de recompensa cumulativa [Korbut 2017].

Como pode ser analisado na Seção 3, a abordagem escolhida para a investigação científica foi a classificação supervisionada. Dessa forma, é necessário mostrar ao modelo de *machine learning* um conjunto de conhecimentos que possa ensiná-lo a classificar tráfego encriptado. Este conjunto de conhecimentos pode ser obtido a partir de um *dataset*, que nada mais é do que uma coleção de dados geralmente tabulados que representam um determinado contexto, onde as linhas representam registros de acontecimentos, enquanto as colunas representam as características ou variáveis [Lashkari et al. 2017]. O *dataset* de treinamento, portanto, deve possuir exemplos suficientes de tráfegos de rede encriptados para que o modelo treinado seja preciso na classificação.

3. Trabalhos relacionados

Diversos trabalhos com diferentes metodologias foram propostos para fazer a classificação de tráfego de rede cifrado, utilizando métodos como Rede neural convolucional, *deep learning* [Liao et al. 2019] ou métodos mais tradicionais de *machine learning* [Lashkari et al. 2017]. A seguir serão discutidos os principais trabalhos na área de classificação de tráfego de rede e que servem de base para construção deste trabalho.

Zhang et al. [Zhang et al. 2017] tem como objetivo principal identificar no tráfego de rede aplicativos desconhecidos, isto é, que não foram rotulados durante o treinamento do *dataset* e que, por isso, não seriam identificados por um classificador convencional. Utiliza-se um classificador binário a partir do algoritmo *Random Forest* para decidir se o tráfego analisado é ou não conhecido e, a partir destes resultados, o modelo é treinado novamente, aumentando entre 5% e 15% a assertividade da próxima classificação binária.

O trabalho proposto por Zhao *et al*. [Zhao *et al*. 2020] apresenta um modelo baseado em CNN (Rede Neural Convolucional, em português) para identificar classes de aplicações conhecidas e possíveis *malwares*. A arquitetura geral de abordagem proposta contém dois módulos principais: pré-processamento e aprendizado de modelo. O módulo de pré-processamento converte fluxos brutos de rede de comprimentos variados em matrizes de tamanho fixo, que são então usadas como entradas do modelo CNN. Para enfrentar os desafios da diversidade dentro da classe e da semelhança entre classes, o módulo Aprendizado de Modelo aprende o modelo da CNN otimizando uma nova função objetivo, que adiciona um termo de regularização da aprendizagem métrica à perda tradicional de entropia cruzada. Alcançou-se resultados satisfatórios com precisão entre 98% e 99% durante os testes realizados.

Lashkari *et al*. [Lashkari et al. 2017], por sua vez, busca classificar o tráfego em uma rede Tor, assim como descobrir o tipo de aplicação que o gerou, para definir se os pacotes recebidos são provenientes de uma rede Tor ou não. Para fazer isso, foi utilizada uma técnica baseada na restrição de tempo, uma vez que diferentes tipos de

tráfego possuem divergências nesta unidade de medida devido às suas características. Para validar o modelo foi utilizado o software *WEKA* e algoritmos como C4.5, *Random Forest* e *K-Nearest Neighbors*. Os melhores resultados obtidos tiveram 85% de precisão com o algoritmo *Random Forest* durante a classificação do tráfego Tor.

Considerando os trabalhos apresentados, percebe-se que as abordagens de classificação de tráfego de rede são abrangentes, visto que o problema possibilita muitas alternativas de solução, entretanto, nenhum dos trabalhos se preocupa com o problema do crescente uso de criptografia, com exceção dos autores Lashkari *et al*. [Lashkari et al. 2017], que utilizam uma rede Tor nos seus estudos. Nas próximas seções é apresentado um sistema de classificação de tráfego de rede cifrado, baseado no estudo de Lashkari *et al*. [Lashkari et al. 2017], mas aplicado ao tráfego na rede Internet em vez de uma rede Tor.

4. Sistema para classificação de tráfego cifrado

Nesta seção é descrito, resumidamente, o sistema proposto para classificar tráfego cifrado de rede, sendo seu objetivo principal identificar o tipo de tráfego de um determinado fluxo de rede como *browsing*, email, *chat, streaming*, transferência de arquivo, *voIp, peer-to-peer*.

A Figura 1 ilustra as principais etapas dos processos implementados

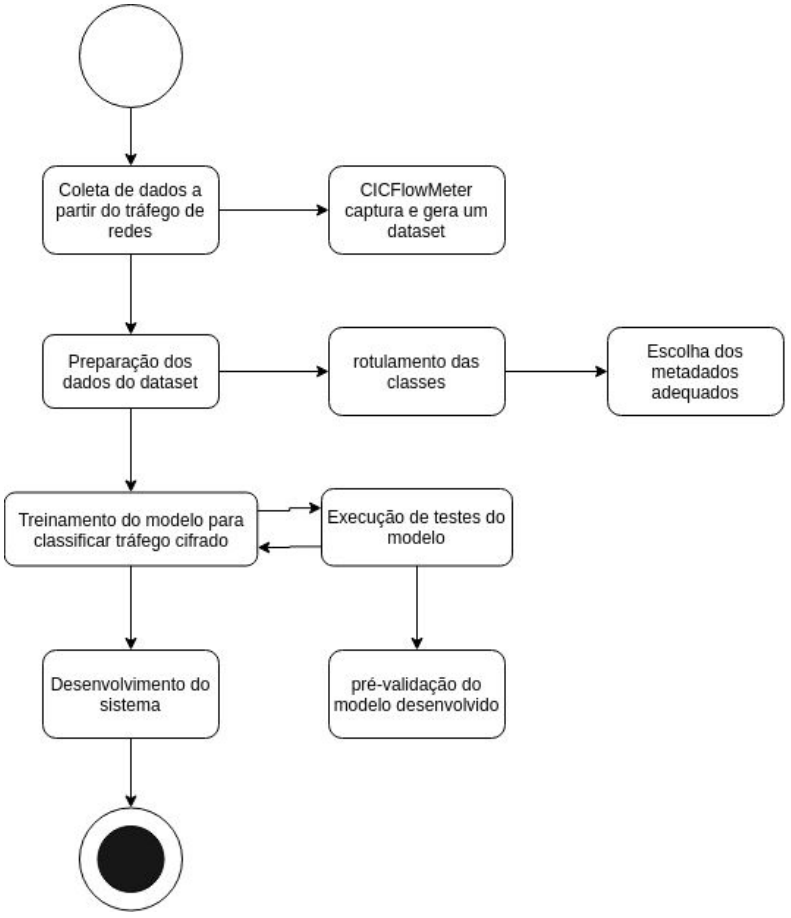

Figura 1. Diagrama do fluxo de atividades

Coleta de dados: Nessa fase, o sistema utilizará o *CICFlowMeter* [Lashkari et al. 2016] para coletar tráfego em uma rede controlada. Segundo Sukeyosi et al. [Sukeyosi et al. 2013], os ambientes controlados de rede (no inglês, network testbeds) são soluções com a finalidade de simular ambientes reais. Para a fase de desenvolvimento, a adoção dessa tecnologia facilita a obtenção de conjunto de dados para a realização da parte central do trabalho, isto é, as fases descritas a seguir.

Tratamento do *dataset*: Com o *dataset* adquirido a partir da etapa anterior, agora os dados deverão ser tratados. Ainda utilizando a ferramenta *CICFlowMeter* [Lashkari et al. 2016], que gera fluxo de tráfegos, a partir do *dataset* gerado na etapa anterior e, a partir desse fluxo, cria novas variáveis no *dataset*. Estas variáveis são características baseada em tempo, uma vez que as informações importantes de um pacote cifrado não são legíveis, entretanto, informações temporais de tráfego em uma rede podem possuir padrões úteis para uma classificação mais precisa e, a partir deles, será treinado o modelo.

Treinamento do modelo: Os algoritmos escolhidos para essa etapa, *Random Forest*, *Naive Bayes* e árvore de decisão, passarão por um processo de *Voting Classifier*, onde é escolhido um número ímpar de algoritmos e, após cada um deles resultar em uma saída, será escolhida a classe resultante pela maioria [Scikit-learn 2020]. O treinamento do modelo consiste em torná-lo apto a reconhecer a qual classe pertence o fluxo de rede, e para que isto seja feito é importante repetir o processo muitas vezes. Para isso, uma porção do *dataset* será utilizada, cerca de 75%. Essa amostra estará rotulada, diferentemente dos outros 25% que servirá para testes, portanto o algoritmo deverá classificá-los, pois são exemplos desconhecidos para ele. Durante o treinamento, variáveis do modelo podem ser removidas ou adicionadas, por exemplo, de maneira a aperfeiçoar e aumentar a classificação do sistema. As variáveis escolhidas para iniciar o treinamento foram as baseadas tempo, as mesmas utilizadas por Lashkari et al [Lashkari et al. 2017]. Dessa forma, espera-se que o sistema seja capaz de reconhecer um tráfego encriptado de maneira precisa, assim como reconhecer um tráfego não encriptado.

5. Considerações finais e trabalhos futuros

Este trabalho apresenta uma solução para a classificação de tráfego cifrado em rede de computadores, considerando que a maioria dos pacotes contam, de alguma forma, com técnicas de criptografia, possibilitando, dessa forma, um controle e monitoramento mais assertivo para a rede.

No estudo realizado no presente trabalho, nota-se que existem diversas possibilidades de classificar um tráfego de rede, tanto pela quantidade de classificadores em *machine learning* quanto pela quantidade ainda desconhecida de métodos de entrada para estes modelos de aprendizagem. Entretanto, ainda são poucos os trabalhos que consideram a criptografia como algo relevante durante a classificação do tráfego e uma abordagem de classificação mais simplificada que não cause impacto no desempenho da rede. Este será o objetivo a ser alcançado no decorrer do presente trabalho.

Para a próxima etapa deste trabalho, terá continuidade o desenvolvimento de um modelo de classificação de tráfego cifrado em rede, fazendo o uso de *machine*

learning. Além disso, pretende-se fazer com que o sistema capture, trate e classifique o tráfego da maneira mais autônoma possível. Também serão executados testes e validações para avaliar o desempenho do sistema desenvolvido.

6. Referências

Bar - Yanai R., Langberg M., Peleg D., Roditty L. (2010) "Realtime Classification for Encrypted Traffic". In: Festa P. (eds) Experimental Algorithms. SEA 2010. Lecture Notes in Computer Science, vol 6049. Springer, Berlin, Heidelberg.

Cisco. "encrypted traffic analytics (white paper)". (2019). https://www.cisco.com/c/dam/en/us/solutions/collateral/enterprise-networks/enterprise-network-security. Junho de 2020.

Cossetti, Melissa Cruz. (2019). "O que é inteligência artificial?". https://tecnoblog.net/263808/o-que-e-inteligencia-artificial. Maio de 2020.

Korbut, Daniil. (2017). "Machine Learning Algorithms: which one to choose for your problem". https://statsbot.co/blog/machine-learning-algorithms/. Julho de 2020.

Lashkari, A. H., Draper-Gil, G., Mamun, M. S. I., & Ghorbani, A. A. (2017). "Characterization of tor traffic using time based features". In ICISSp, pages 253-262.

Neu, C. V. (2019). "Detecting encrypted attacks in software-defined networking".

Pila, Adriano Donizete. (2001). "Seleção de atributos relevantes para aprendizado de máquina utilizando a abordagem de Rough Sets".

Scikit-learn. "User Guide". (2020). https://scikit-learn.org/stable/user_guide.html. setembro de 2020.

Singhal, P., Mathur, R., & Vyas, H. (2013). "State of the Art Review of Network Traffic Classification based on Machine Learning Approach". International Journal of Computer Applications, 975, 8887.

Stallings, William. (2015), Criptografia e segurança de redes: princípios e práticas. São Paulo:Pearson. 6ª edição.

Sukeyosi, W. A. P., de Mattos, E. P., Zarpelão, B. B., & de Souza Mendes, L. (2013). "Ambientes Controlados de Geração de Anomalias: uma Reprodução de Ataques de Negação de Serviço". Anais do Computer on the Beach, pages 88-97.

Zhao, L., Cai, L., Yu, A., Xu, Z., & Meng, D. (2020). "A novel network traffic classification approach via discriminative feature learning". In Proceedings of the 35th Annual ACM Symposium on Applied Computing, pages 1026-1033.

Zhang, Z., Kang, C., Fu, P., Cao, Z., Li, Z., & Xiong, G. (2017). "Metric learning with statistical features for network traffic classification". In 2017 IEEE 36th International Performance Computing and Communications Conference (IPCCC), pages 1-7. IEEE.

Zhu, Q., Li, D., Xin, Y., Yu, X., & Mu, G. (2019). "A Survey on Network Traffic Identification". In International Conference on Artificial Intelligence and Security, pages 91-100. Springer, Cham.