# Detecção de Ataques *Port Scan* e *Slowloris* Através de Sistemas de Detecção de Intrusão com *Quadratic Discriminant Analysis*

**Vinícius M. Deolindo, Jeferson C. Nobre,**
**Luiz Ricardo Bertoldi de Oliveira, Allan de Barcelos Silva**

[1]Universidade do Vale do Rio dos Sinos - UNISINOS
São Leopoldo, RS, Brasil

vinicius.deolindo@outlook.com

{jcnobre,luizbertoldi,allanbs}@unisinos.br

***Abstract.*** *Identify and classify attacks through Intrusion Detection Systems (IDS) are one constant challenge for security professionals. Computer networks are one of the significant IT components that support classification operations. Machine Learning techniques can aid in this process by providing methods capable of making decisions based on the previously known information. Use of Quadratic Discriminant Analysis (QDA) as a classification method is not a widely used technique in the use of IDS classifiers. This study presents a way of applying a data sorting technique for the identification of Scan Port and Dos Slowloris attacks using QDA algorithm.*

***Resumo.*** *Identificar e classificar ataques através de Sistemas de Detecção de Intrusão (Intrusion Detection System - IDS), são um desafio constante para profissionais de segurança. As redes de computadores são um dos principais componentes de TI que suportam as operação de classificação. Técnicas de Machine Learning podem auxiliar neste processos, fornecendo métodos capazes de tomar decisões baseadas em informações previamente conhecidas. A utilização de Quadratic Discriminant Analysis (QDA) como método de classificação, não é uma técnica largamente explorada na utilização de classificadores para IDS. Este estudo apresenta uma maneira de aplicar uma técnica de classificação de dados para a identificação de ataques do tipo Port Scan e Dos Slowloris utilizando algoritmo QDA.*

## 1. Introdução

As redes de computadores fornecem um meio de comunicação entre os mais diversos dispositivos da atualidade. Estas comunicações tornam os dispositivos acessíveis para os usuários e sugestivos aos mais diversos tipos de ataques. Serviços como Sistemas de Detecção de Intrusão (*Intrusion Detection System* - IDS) fornecem ao ambiente um sistema capaz de identificar ocorrências que possam violar: a segurança; comprometer a integridade; confidencialidade; disponibilidade de um recurso ou serviço [Osareh and Shadgar 2008].

*Machine learning* é um área de conhecimento que tem como um de seus objetivos desenvolver técnicas que permitam que programas de computador possam adquirir informações e conhecimento de forma automatizada. Este processo para o aprendizado é

construído através de técnicas específicas que permitem que um sistema possa tomar decisões baseadas em experiências previamente avaliadas [Monard and Baranauskas 2003]. Tornar sistemas capazes de prever ocorrências das mais diversas é um dos desafio que os Cientistas de Dados enfrentam em *machine learning*. A tomada de decisão correta e automatizada de um sistema depende de uma série de fatores e análises preliminares para a definição de um modelo capaz de prever assertivamente uma informação.

Os métodos tradicionais utilizados por IDS, como Anomalia e Assinatura, podem apresentar problemas para identificar determinados tipos de ataques que possam ter sofrido alteração no método de atuação, isso faz com que o novo comportamento malicioso não seja percebido pelo Sistema. A detecção de anomalias de rede não é o suficiente para identificação de um ataque, uma questão importante no *Machine Learning* é que é preciso treinar algoritmos que possam detectar diversos ataques, e não apenas informar se o tráfego parece ou não legítimo [Sommer and Paxson 2010]. Assinaturas de ataque também podem sofrer variações sutis que podem impactar na detecção de um ataque, para isso é necessário estabelecer algoritmos que sejam capazes de entender perfis de ataques através de métricas específicas, direcionando assim a correta classificação e resposta pelo IDS.

A adoção de técnicas de *Machine Learning* para estabelecimento de padrões de ataque, que pode ser caracterizado como método de Detecção de Anomalia, pode auxiliar a determinar o tipo de ataque sofrido, facilitando assim a tomada de decisões para medidas de reação. Técnicas automatizadas são ideais para IDS, pois permitem monitorar e correlacionar um grande número de informações de padrões e assinaturas [Sinclair et al. 1999]. Foi adotado o QDA para a realização deste estudo, uma vez que, ele não é largamente utilizado para técnicas de ML em IDS. O QDA é uma das abordagens padrão para algoritmos de classificação de dados e tem bom desempenho quando existe bastante quantidade de dados para o estabelecimento de fatores [Srivastava et al. 2007].

Este trabalho apresenta os passos para criação de um modelo de ML para identificação de ataques de *Port Scan* e *DoS Slowloris* com base nos seus comportamentos. A seção 2 apresenta a fundamentação teórica, em seguida a seção 3 descreve os resultados obtidos e a seção 4 as conclusões do estudo.

## 2. Fundamentação Teórica

O presente trabalho utilizou tecnologias e métodos de *Machine Learning* (ML), *CRISP-DM* [Wirth and Hipp 2000], IDS e *Quadratic Discriminant Analysis* (QDA). No que tange o primeiro termo, é uma área da Inteligência Artificial que consiste em um conjunto de algoritmos matemáticos e estatísticos que podem ser aplicados a tarefas utilizando um conjunto de dados [Horst 1999]. Em sua maioria, os métodos possuem o objetivo de construir sistemas que possam adquirir conhecimento de forma automática, tomando decisões baseadas em informações repassadas anteriormente [Monard and Baranauskas 2003]. Em uma das etapas, denominada classificação dos dados, é adotado um método de aprendizado para o algoritmo, podendo ser Supervisionado e Não Supervisionado. No método Supervisionado o conjunto de dados utilizado para identificação já possui algum tipo de classe, e o desafio é treinar o algoritmo para que estas classes possam ser novamente classificadas[Horst 1999]. O Não-Supervisionado, dado o conjunto de dados, este método busca identificar classes ou padrões, não tendo ainda estes estabelecidos [Horst 1999].

Para a aplicação das tecnologias de inteligência artificial na área de segurança, o presente trabalho propõe uma alternativa para identificação de ataques do tipo *Port San* e *Slowloris* através da coleta de dados de IDS. Através do estudo *Toward generating a new intrusion detection dataset and intrusion traffic characterization*[Sharafaldin et al. 2018], que criou um *dataset* completo com diversos tipos de ataques, foi possível realizar a extração dos padrões de ataques para a realização deste estudo. Com o mapeamento de informações relevantes para cada tipo de ataque, é possível estabelecer padrões que tornaram possíveis a utilização de algoritmos de ML para classificação e consequente identificação de determinados ataques. Pra este estudo foi utilizado o *framework* CRISP-DM, o qual é composto de seis etapas (*Business Understanding*, *Data Understanding*, *Data Preparation*, *Modeling*, *Evaluation* e *Deployment*) e fornece padrões para a realização das atividades de um projeto de ML [Nadali et al. 2011]. De modo a promover a melhor compreensão, serão apresentadas as etapas de preparação dos dados e criação do modelo para a classificação de ataques.

### 2.1. Etapa - *Business Understanding*

A fase de entendimento do negócio, neste projeto, pode ser resumida no problema que o projeto busca resolver. Neste caso o problema principal do estudo, é a identificação de padrões de ataque de *Port Scan* e *Slowloris* através de IDS com o uso de QDA.

Neste etapa, foram avaliados os conjuntos de dados de IDS e a suas características, para que fosse possível determinar que dados seriam utilizados no projeto. Destacou-se o projeto *Toward generating a new intrusion detection dataset and intrusion traffic characterization*, que desenvolveu um conjunto de dados com informações completas de diversos tipos de ataques coletados por IDS. Também foi possível verificar que os dados foram construídos utilizando técnicas recentes de ataque, e o ambiente utilizado para a realização do estudo foi documentado. Informações como, origem dos ataques, sistemas utilizados, dados da rede interna (LAN) e externa (WAN), períodos em que ocorreram os ataques estavam disponíveis neste conjunto de dados. Além de todo o detalhamento disponível sobre o estudo, o principal fator que destacou o *dataset* foram os *lebels* de classificação de cada ataque, o que facilitou a criação futura de um modelo de classificação.

### 2.2. Etapa - *Data Understanding*

Para iniciar a avaliação dos dados foi necessária a solicitação, através de e-mail, dos *datasets* para a *Canadian Institute for Cybersecurity* que é a organização que conduziu o estudo de criação dos conjuntos de dados de IDs contendo os ataques. Após a disponibilização dos conjuntos de dados, iniciou-se a avaliação.

Para avaliação inicial do conjunto de dados, foram observadas informações como estrutura das tabelas de dados. Também foi avaliado a diversidade de *labels* de classificação de ataque contidos nos arquivos, pois para a criação do modelo utilizando QDA, é importante que existam um número relevante de dados para geração do modelo.

Foi possível perceber que existem 127537 dados classificados como *BENIGN* e 158930 *PortScan*, estas informações representam mais da metade dos dados classificados com o ataque (*PortScan*), isso significa que existe uma grande quantidade de dados para criação do modelo de identificação de ataque. Devido a estrutura dos arquivos ser a mesma, o procedimento de análise para os dados relacionados ao ataque de *Slowloris* foram os mesmos.

### 2.3. Etapa - *Data Preparation*

Qualidade dos dados avaliados é fundamental para o sucesso do projeto, na etapa de preparação são removidas as informações que não serão utilizadas no futuro, pois, a maioria dos algoritmos de aprendizado utiliza o conhecimento extraído dos dados sem a utilização e outras fontes externas. [Batista et al. 2003] Para a construção dos conjuntos de dados para cada tipo de ataque, foram utilizadas as informações do estudo *Toward generating a new intrusion detection dataset and intrusion traffic characterization*, que indicou quais as informações mais relevantes para a avaliação de cada tipo de ataque. A Tabela 1 representa as colunas com maior relevância para cada tipo de ataque de acordo com autores anteriores [Sharafaldin et al. 2018]. De posse das informações mais relevantes é possível montar os dados que serão utilizados no modelo.

**Tabela 1. Informações para ataques *Slowloris* e *Port Scan***

| *Label* | *Feature* | *Weigh* |
|---|---|---|
| *PortScan* | *Init Win F.Bytes* | 0.0083 |
| | *B.Packets/S* | 0.0032 |
| | *PSH Flag Count* | 0.0009 |
| *DoS slowloris* | *Flow Duration* | 0.0431 |
| | *F.IAT Min* | 0.0378 |
| | *B.IAT Mean* | 0.0300 |
| | *F.IAT Mean* | 0.0265 |

### 2.4. Etapa - *Modeling*

Conforme já descrito anteriormente, este estudo busca criar modelos para identificação de ataques *Port Scan* e *Slowloris* utilizando o QDA, e a fase de Modelagem é a etapa em que os testes de criação dos modelos ocorre. Inicialmente, para que fosse possível a criação do modelo em QDA, foi necessário a realização de algumas etapas de preparação dos dados. Foi estabelecido um percentual de 60% dos dados para treinamento do modelo e os restantes 40% para avaliação dos resultados. Para a criação dos conjuntos de dados de treinamento e teste, foi utilizado a função SAMPLE(), que tem a função de extrair quantidades aleatórias de um conjunto, ou seja, quando criarmos um conjunto com 60% dos dados, estes serão extraídos de maneira aleatória e não os primeiros 60%. O código a seguir demostra como foram criados os conjuntos, estes passos são idênticos para a criação de conjuntos de ambos os ataque.

### 2.5. Etapa - *Evaluation*

Nesta etapa, foi realizada a fase de avaliação do modelo criado na etapa *Modeling*. A diversidade dos dados para a viabilização de um modelo é a chave para estabelecimento de um padrão, outro fator que também determinou os ataques que seriam utilizados neste estudo foi a precisão alcançada pelo Modelo criado.

### 2.6. Etapa - *Deployment*

Nesta etapa ocorre a validação com as partes interessadas e também a instalação em produção não foi realizada neste estudo. Uma vez que, este é um estudo de caso, realizado em um ambiente controlado para avaliar a viabilidade de utilização de métricas de

rede para a identificação e classificação de ataques específicos, a implementação do modelo em ambiente de produção pode ser relacionado ao trabalho futuro, posterior e este estudo.

## 3. Resultados

Este estudo utilizou um conjunto de dados construído especificamente para a realização de experimentos. Desenvolvido pela *Canadian Institute for Cybersecurity*, os conjuntos de dados apresentam uma série de informações referentes aos principais tipos de ataques detectados por IDS. Para a avaliação do modelo, foi utilizado o pacote CARET, que contém diversas funções para treinar e apresentar modelos de classificação e regressão como o QDA [Kuhn et al. 2008]. Este pacote permite que sejam exibidos resultados gerais do modelo de classificação, resultados estes que utilizam como parâmetros os dados fornecidos pela matriz confusão do modelo. A Tabela 2 apresenta o resultado da matriz confusão e das métricas extraídas dos modelos utilizados para classificação dos ataques de *Port Scan* e *DoS Slowloris*:

**Tabela 2. Resultados das métricas obtidas.**

| Métrica | *PortScan* | *DoS Slowloris* |
|---|---|---|
| TP | 7568 | 27060 |
| FP | 1 | 168 |
| FN | 27 | 348 |
| TN | 63426 | 2023 |
| Acurácia | 0,9996 | 0,9826 |
| Precisão | 0,9999 | 0,9938 |
| *Recall* | 0,9964 | 0,9873 |
| *F-measure* | 0,9982 | 0,9906 |
| Especificidade | 1,0000 | 0,9233 |

Com base nos resultados apresentados na Tabela 2, é possível dizer que o modelo criado para classificação dos ataques de *DoS Slowloris*, utilizando as variáveis *Flow Duration, F.IAT Min, B.IAT Mean, F.IAT Mean* obteve um percentual de 99,4% de f-measure na classificação das amostras de testes. Os resultados para a classificação e identificação do ataque de *Port Scan* foram melhores se comparados aos resultados do *DoS Slowloris*.

## 4. *Conclusão*

O presente trabalho busca identificar, através de métricas de IDS e reconhecimentos de padrões, ataques do tipo *Port Scan* e *Dos Slowloris*. Para a detecção dos ataques foi adotada a utilização de algoritmo de classificação QDA com o método de aprendizagem Supervisionado. Como guia para a realização do estudo, foi utilizado o *framework* CRISP-DM, que forneceu uma guia com etapas claras para a realização do estudo.

Os resultados demostraram que a identificação dos ataques através da utilização de QDA e utilizando as métricas mapeadas obtiveram um alto percentual de precisão. A diversidade e qualidade de dados utilizados para a construção do modelo, foram de grande importância na obtenção dos resultados. Como trabalhos futuros, cabe a analise de performance e comparativo de precisão com outros algoritmos de classificação.

## Referências


Batista, G. E. d. A. P. et al. (2003). *Pré-processamento de dados em aprendizado de máquina supervisionado*. PhD thesis, Universidade de São Paulo.

Horst, P. S. (1999). *Avaliação do conhecimento adquirido por algoritmos de aprendizado de máquina utilizando exemplos*. PhD thesis, Universidade de São Paulo.

Kuhn, M. et al. (2008). Caret package. *Journal of statistical software*, 28(5):1–26.

Monard, M. C. and Baranauskas, J. A. (2003). Conceitos sobre aprendizado de máquina. *Sistemas inteligentes-Fundamentos e aplicações*, 1(1):32.

Nadali, A., Kakhky, E. N., and Nosratabadi, H. E. (2011). Evaluating the success level of data mining projects based on crisp-dm methodology by a fuzzy expert system. In *Electronics Computer Technology (ICECT), 2011 3rd International Conference on*, volume 6, pages 161–165. IEEE.

Osareh, A. and Shadgar, B. (2008). Intrusion detection in computer networks based on machine learning algorithms. *International Journal of Computer Science and Network Security*, 8(11):15–23.

Sharafaldin, I., Lashkari, A. H., and Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. In *Proceedings of fourth international conference on information systems security and privacy, ICISSP*.

Sinclair, C., Pierce, L., and Matzner, S. (1999). An application of machine learning to network intrusion detection. In *Computer Security Applications Conference, 1999.(ACSAC'99) Proceedings. 15th Annual*, pages 371–377. IEEE.

Sommer, R. and Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *Security and Privacy (SP), 2010 IEEE Symposium on*, pages 305–316. IEEE.

Srivastava, S., Gupta, M. R., and Frigyik, B. A. (2007). Bayesian quadratic discriminant analysis. *Journal of Machine Learning Research*, 8(Jun):1277–1305.

Wirth, R. and Hipp, J. (2000). Crisp-dm: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, pages 29–39. Citeseer.