Detecção de Ataques de Negação de Serviço Distribuídos com Algoritmos de Aprendizado de Máquina

Rodrigo R. Silva[1], Felipe da R. Henriques[1], Igor M. Moraes[2], Dalbert M. Mascarenhas[1]

[1]Centro Federal de Educação Tecnológica Celso Suckow da Fonseca - CEFET/RJ, Petrópolis - RJ - Brasil
[2]Laboratório MidiaCom – IC/TCC/PGC
Universidade Federal Fluminense (UFF), Niterói – RJ – Brasil

{felipe.henriques,dalbert.mascarenhas}@cefet-rj.br,
rodrigo.silva@aluno.cefet-rj.br, igor@ic.uff.br

18 de Setembro de 2024

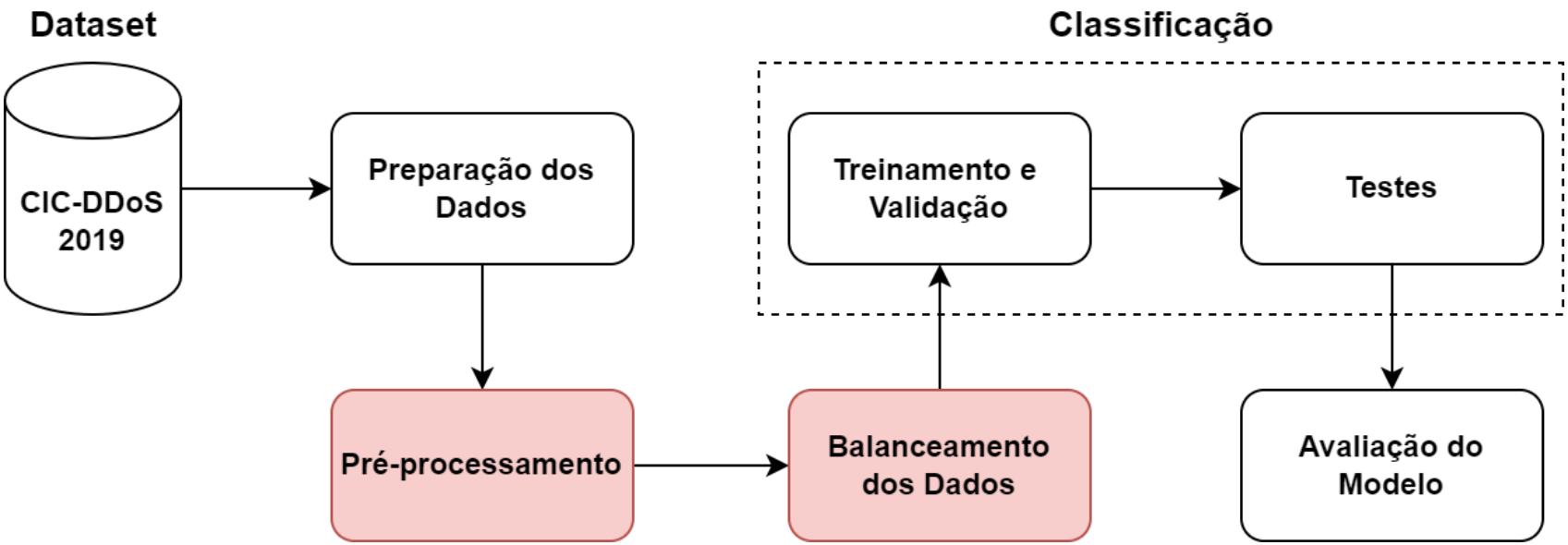
Sumário

Objetivos

- Apresentar uma metodologia para detectar e classificar ataques de negação de serviço distribuídos (DDoS):
  - Possibilitou a redução de atributos acima de 90% mantendo ótimos resultados.

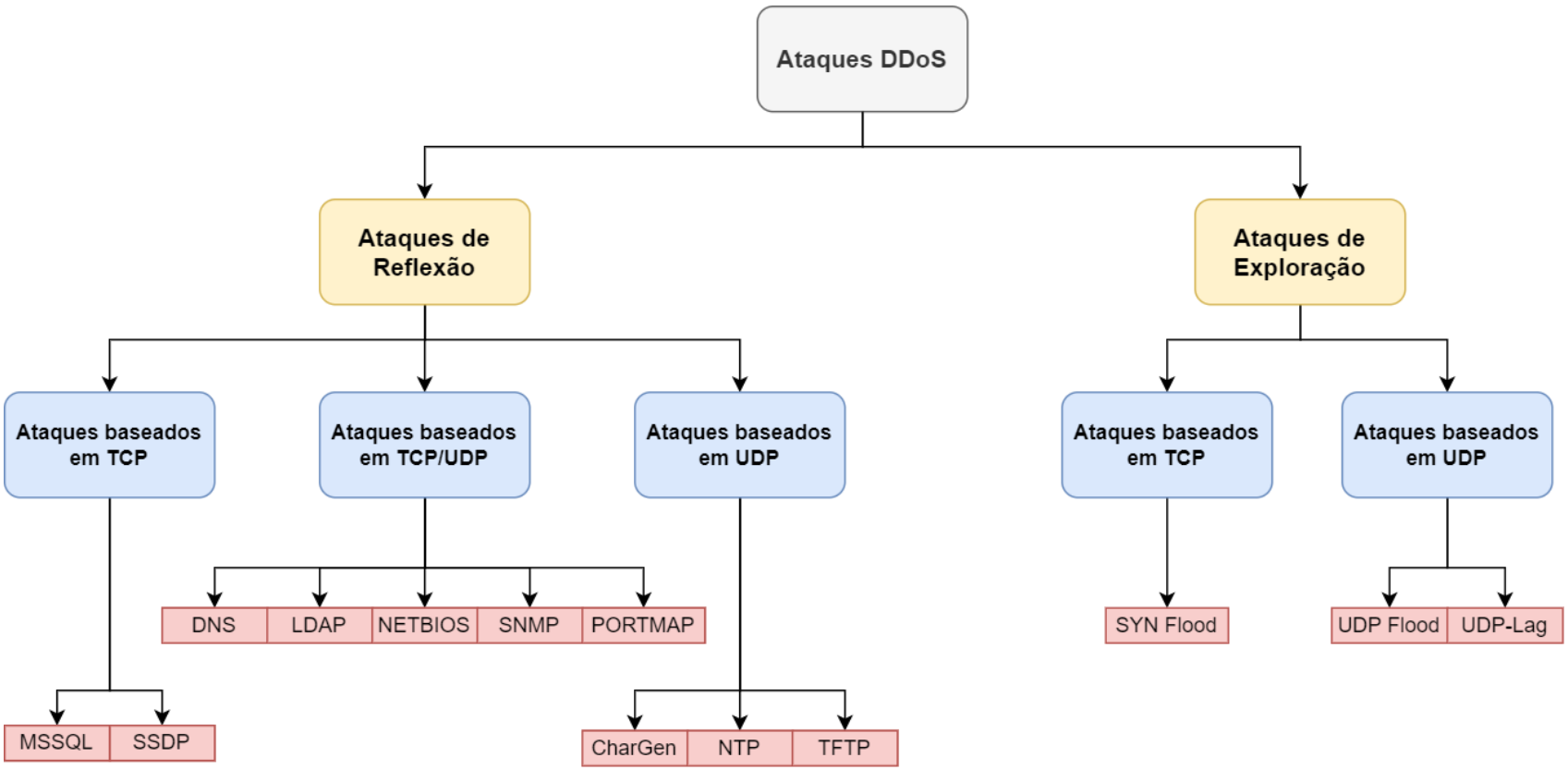
Arquitetura

Figura 1: Arquitetura da solução proposta



Dataset CIC-DDoS2019

Figura 2: Fluxograma dos ataques DDoS



Pré-processamento

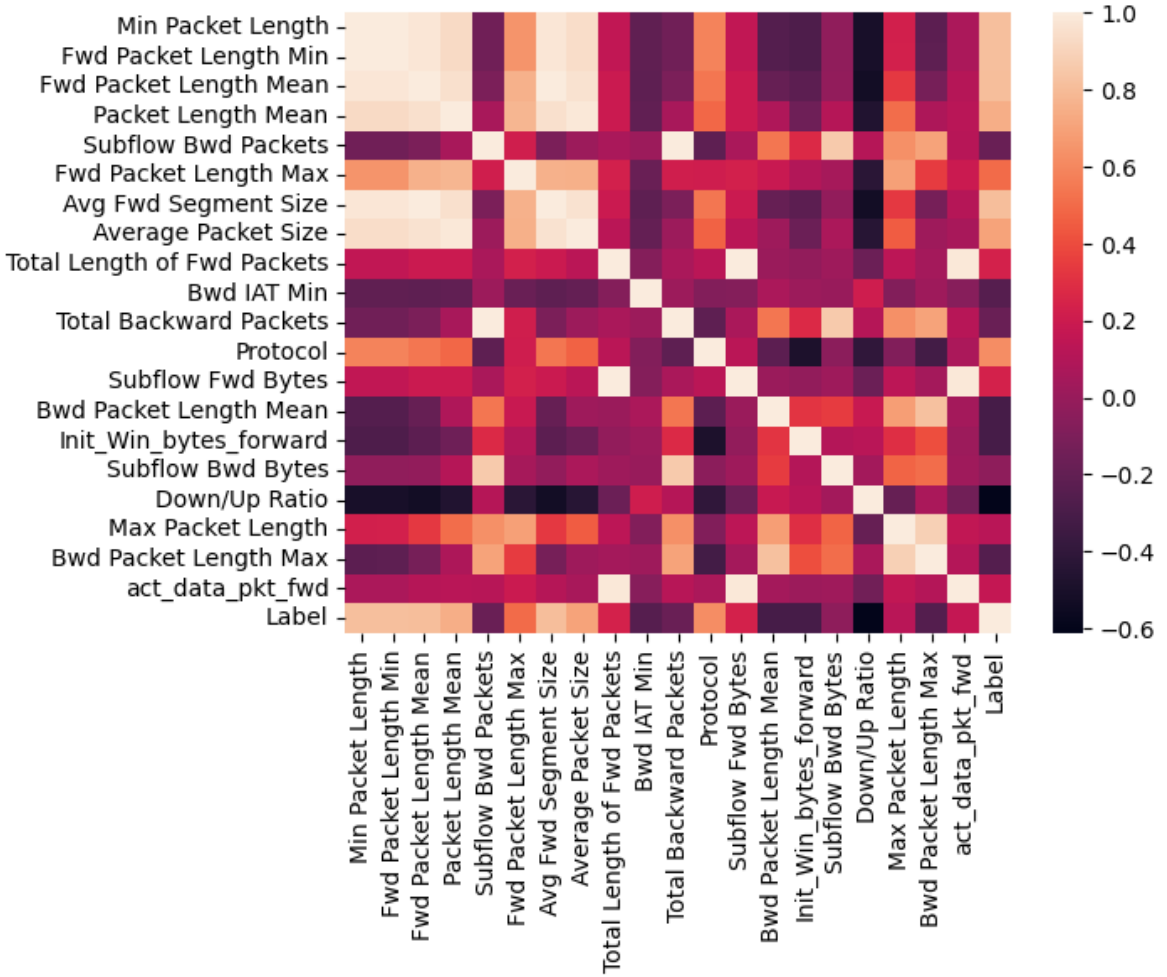
- Remoção de linhas com valores infinitos e Not a Number (NaN);
- Remoção de linhas duplicadas;
- Normalização dos dados para o intervalo [-1;1];
- Classificação binária:
  - Codificação do alvo em 0 (benign) e 1 (ataques);
- Classificação multiclasse:
  - Codificação do alvo em números inteiros (0, 1, 2 e 3);
- Codificação *one-hot-encode* para modelo MLP;

Seleção de Atributos (Etapa 1)

- De um total de 89 atributos, removeu-se 11 atributos de metadados;
- Combinação de *variance threshold* de 0% até 30% e *feature_importances_* de 10% até 40%
- Foram feitos 16 testes em todos os 6 modelos para ambas classificações totalizando 192 testes;
- Classificação binária (20 atributos restantes):
  - *Variance threshold:* 0%
  - *Feature_importances_:* 30%
- Classificação Multiclasse (17 atributos restantes):
  - *Variance threshold:* 20%
  - *Feature_importances_:* 30%

Seleção de Atributos (Etapa 2) - Classificação Binária

- **Matriz de Correlação (antes da análise) - 20 atributos**

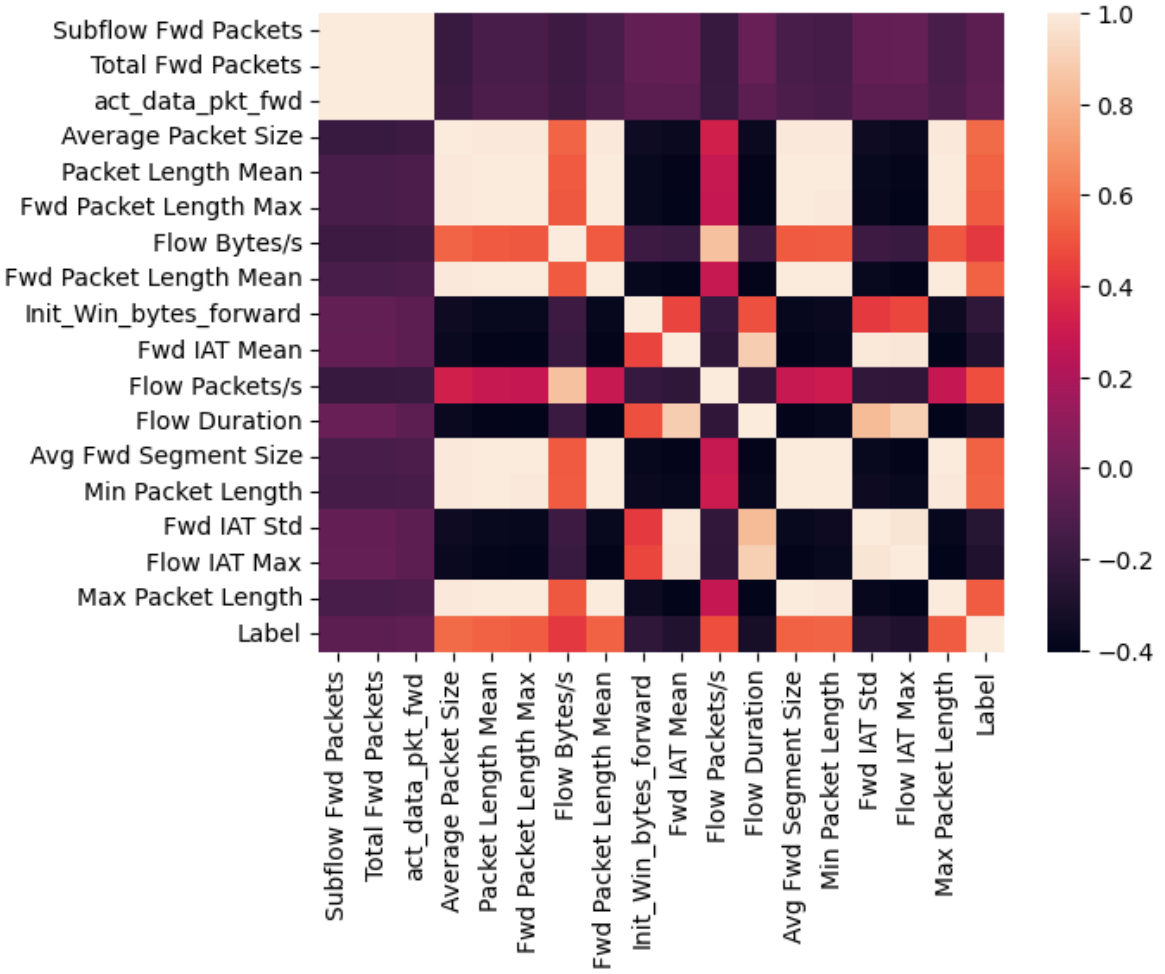


Seleção de Atributos (Etapa 2) - Classificação Binária

- **Matriz de Correlação (após a análise) - 8 atributos**

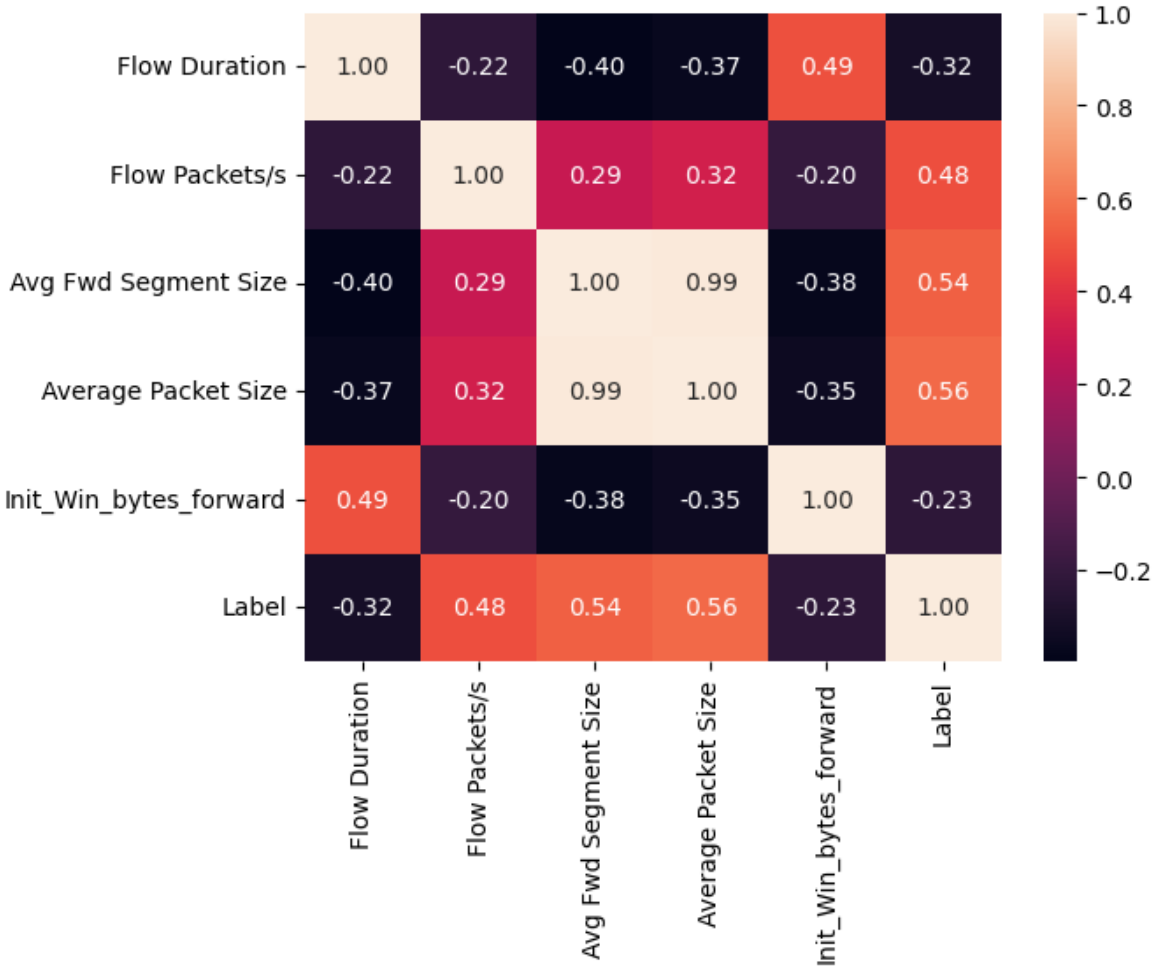
Seleção de Atributos (Etapa 2) - Classificação Multiclasse

- **Matriz de Correlação (antes da análise) - 17 atributos**



Seleção de Atributos (Etapa 2) - Classificação Multiclasse

- **Matriz de Correlação (após a análise) - 5 atributos**



Porque é importante reduzir os atributos ?

- Menos atributos = menor tempo de detecção;
- No contexto de ataque DDoS:
  - Redes de médio a grande porte que têm tráfego massivo de dados se beneficiam desta característica.

Balanceamento & Conjunto de Dados

- **Classificação Binária:**
  - Utilizou-se geração de dados sintéticos com SMOTE (*Synthetic Minority Over-sampling Technique*).

Tabela 1: Quantidade de amostras de tráfego de ataque e benigno em cada cenário antes e depois da reamostragem de dados

		Qntd. Tráfego		
Reamostragem	**Cenário**	**Ataque**	**Benigno**	**Proporção**
Antes	1 e 2	639.175	2.870	0,45%
Depois	1	320.000 30.000	320.000 30.000	
	2	2.870	2.870	

Balanceamento & Conjunto de Dados

- **Classificação Multiclasse:**

Tabela 2: Ataques escolhidos e quantidade de amostras lidas.

Categoria	Baseado em	Tipo do Ataque	Quantidade
Ataque de Reflexão	TCP	MSSQL	61.926
	TCP/UDP	DNS	15.552
		LDAP	21.704
		NETBIOS	13.919
	UDP	NTP	56.767
Ataque de Exploração	TCP	SYN Flood	38.721
	UDP	UDP Flood	155.022

Tabela 3: Quantidade de amostras por classe após reamostragem.

Classe	Quantidade
UDP	13.919
SYN Flood	13.919
NTP	13.919
Outros	55.676

Métricas de Avaliação

- $$Acurácia = \frac{VP}{VP + VN + FP + FN} \quad (1)$$
- $$Precisão = \frac{VP}{VP + FP} \quad (2)$$
- $$Sensibilidade = \frac{VP}{VP + FN} \quad (3)$$
- $$F1\text{-}Score = 2 \times \frac{Precisão\ .\ Sensibilidade}{(Precisão + Sensibilidade)} \quad (4)$$
- $$ROC-AUC = [0, 1] \quad (5)$$

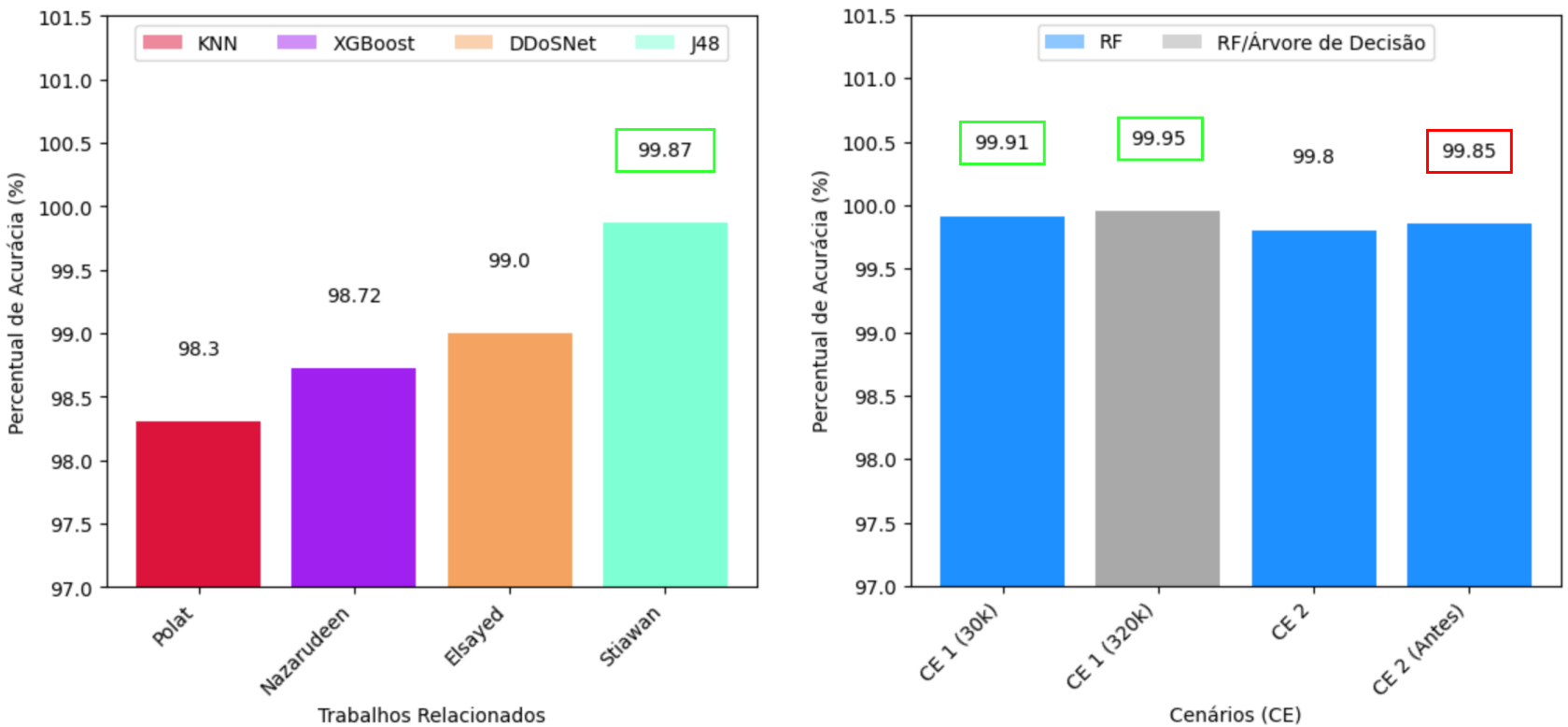
Modelos Avaliados

Todos os modelos avaliados foram utilizados usando a biblioteca *scikit-learn* do *Python*, exceto MLP:

- *Naive Bayes* sem PCA
- *Naive Bayes* com PCA
- MLP – Biblioteca Keras (*Tensor Flow*)
- Árvore de Decisão
- *Random Forest*
- SVM (*Support Vector Machine*)

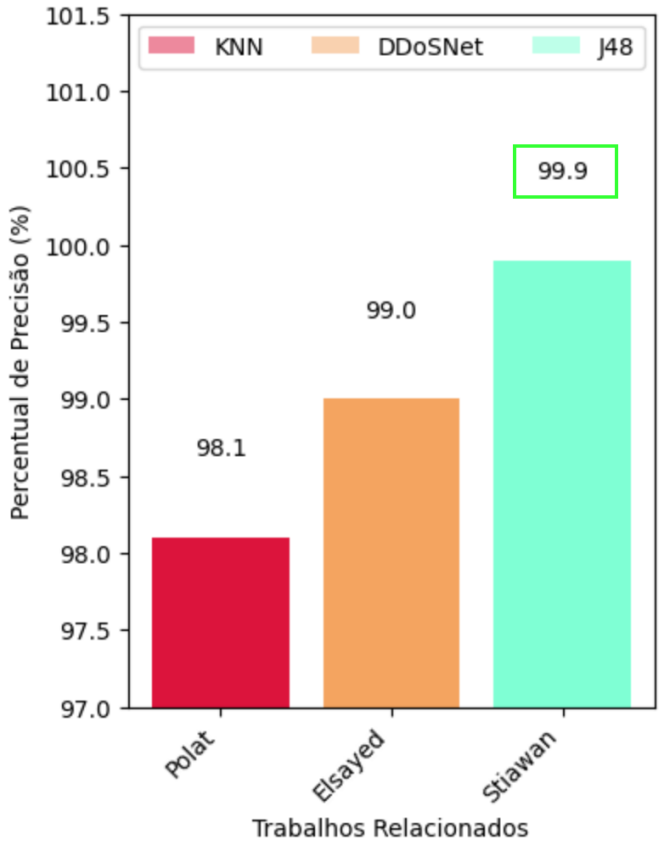
Classificação Binária - Acurácia

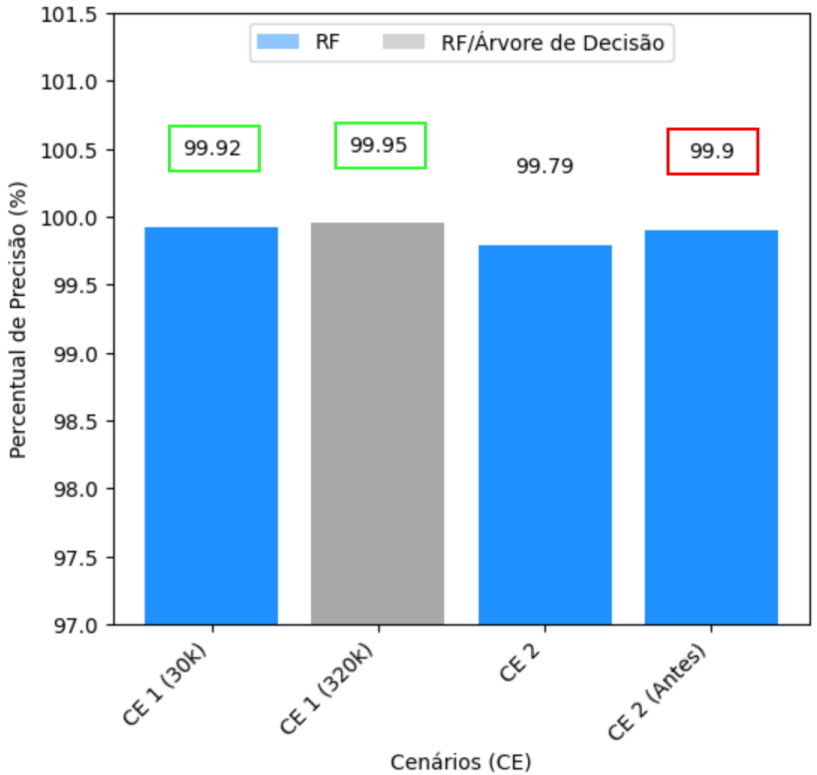
Figura 3: Resultados da acurácia comparados ao melhor modelo dos trabalhos relacionados



Classificação Binária - Precisão

Figura 4: Resultados da precisão comparados ao melhor modelo dos trabalhos relacionados





Classificação Multiclasse

Figura 5: Relatório de classificação *random forest* após análise da matriz de correlação (etapa 2 da seleção de atributos)

	precision	recall	f1-score	support
NTP	0.99	0.99	0.99	4772
OUTROS	0.99	0.99	0.99	18919
SYN	1.00	1.00	1.00	4753
UDP	0.99	0.98	0.99	4684
accuracy			0.99	33128
macro avg	0.99	0.99	0.99	33128
weighted avg	0.99	0.99	0.99	33128

Conclusão

- Os modelos utilizados obtiveram excelentes resultados comparados aos outros trabalhos.
- Classificação Binária (8 atributos - 91,01% de redução total):
  - Resultados do CE1 foram melhores do que qualquer outro trabalho.
  - Resultados do CE2 são superiores a todos os trabalhos comparados, exceto (Stiawan) que obteve 99,87% de acurácia, usando 52 atributos, contra 99,85% no CE2 (antes análise da matriz de correlação) com 20 atributos.
- Classificação Multiclasse (5 atributos - 94,38% de redução total):
  - Resultados equivalentes ao trabalho (Nazarudeen) (11 ataques)
  - 100% de detecção do ataque SYN-Flood.
  - Detecção $>= 98\%$ para as 3 demais categorias.

Agradecimento

Este trabalho foi realizado com recursos da RNP, CNPq, CEFET/RJ, CAPES, FAPERJ e PGC/UFF.