Secflow

Instituto de Computação

SBSeg | 23 JUIZ DE FORA

APRENDIZADO NÃO SUPERVISIONADO PARA ANÁLISE E DETECÇÃO DE ANOMALIAS EM REDES DE COMPUTADORES

Felipe Salles (1° autor)

Prof. Dr. Luiz Claudio Schara

Prof. Dra. Taiane Ramos

Índice

Contextualização sobre o Cenário de Detecção

Tipos de Sistemas de Detecção de Intrusão (IDS)

- Assinaturas
- Anomalias
- Híbridos

Dificuldades Atuais

- Acompanhar a **evolução** e mudança de **comportamento** dos ataques.
- Ter uma boa **base de dados** para estudo.

Proposta do Secflow

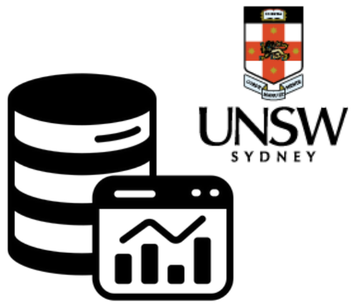
- Estudar, avaliar e **detectar** anomalias na rede da Universidade Federal Fluminense (UFF).

Clustering

USO DE APRENDIZADO NÃO SUPERVISIONADO

Pré-Processamento

UNSW-NB15



1

Extraímos 20% de cada *label*.

! Classes desbalanceadas → Refletem o cenário de uma rede real.

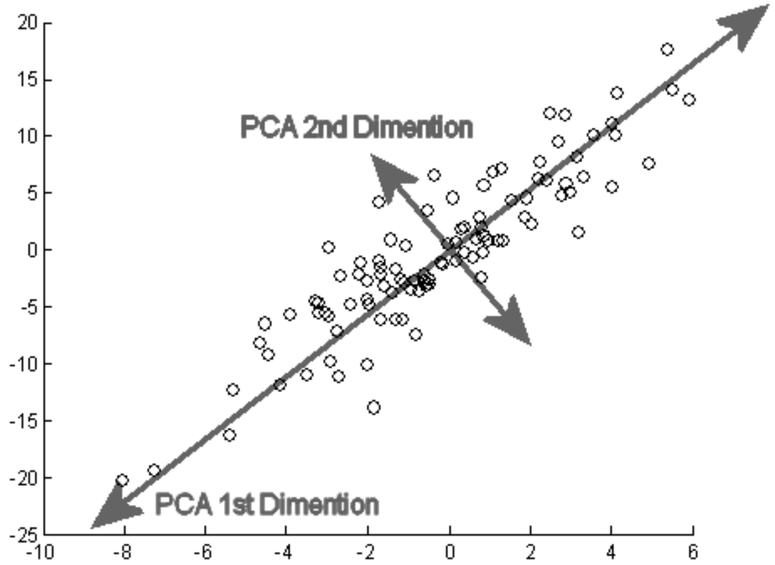
Fluxos Normais

Tipos de Fluxos de Ataque

Labels	Amostras	20% de Amostras
Normal	93.000	18.600
Analysis	2677	535
Backdoor	2329	465
DoS	16.353	3270
Exploits	44.525	8905
Fuzzers	24.246	4849
Generic	58.871	11.774
Reconnaissance	13.987	2797
Shellcode	1511	302
Worms	174	34

Pré-Processamento

2

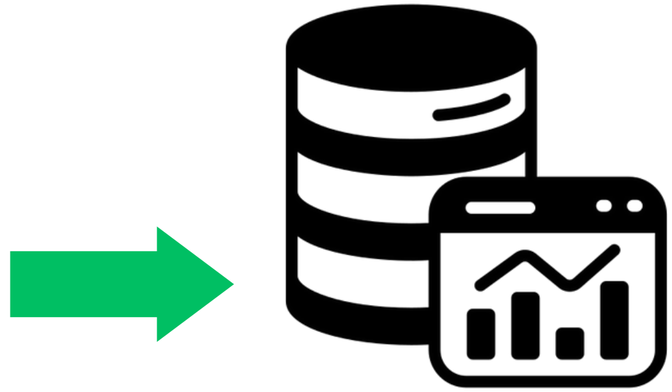


Aplicamos o PCA para seleção de features.

3

$$X^* = \frac{X - \mu}{\sigma}$$

Aplicamos o método *Z-Score*

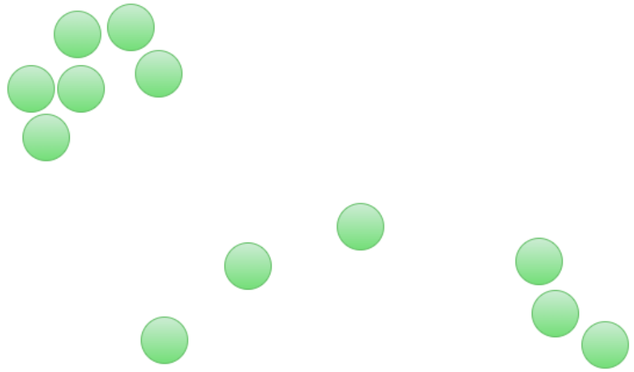


Conjunto de Dados Final

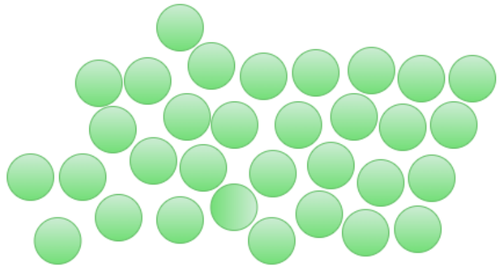
DBSCAN (ε, MinPts)

Foco em identificar anomalias

Parâmetros

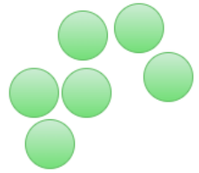


Explicando o funcionamento com um exemplo...



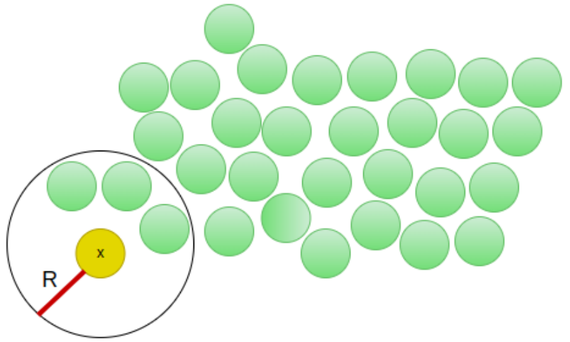
44 pontos

Exemplo DBSCAN (ε=R, MinPts=3)

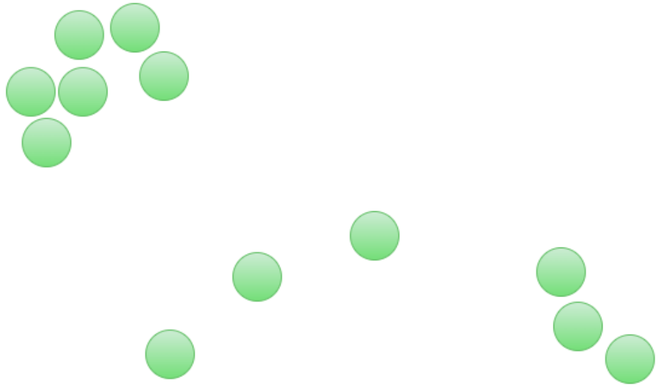


Exemplo: 44 pontos

1°Passo-Para cada ponto, verificar se é um ponto central.

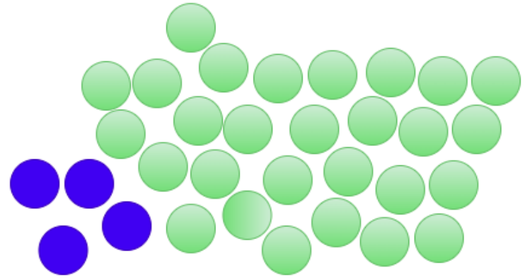


Exemplo DBSCAN (ε=R, MinPts=3)

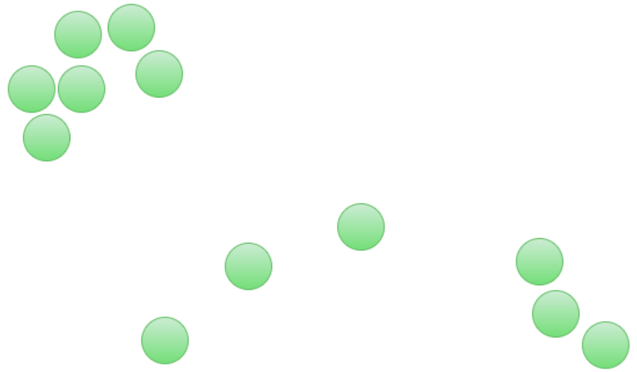


Exemplo: 44 pontos

2°Passo-Se for um ponto central, adiciona os pontos vizinhos no mesmo grupo.

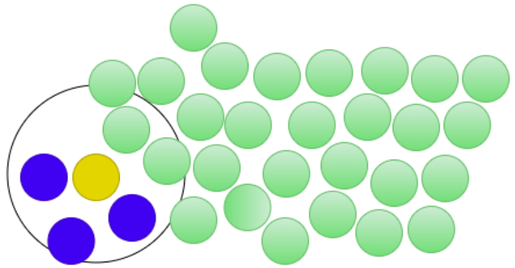


Exemplo DBSCAN (ε=R, MinPts=3)

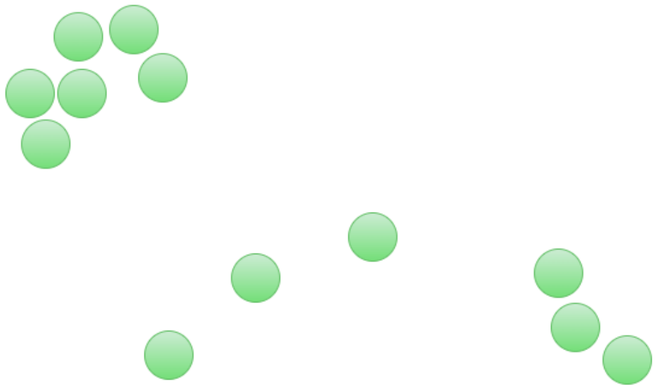


Exemplo: 44 pontos

3°Passo-Verifica se algum vizinho é um ponto central.

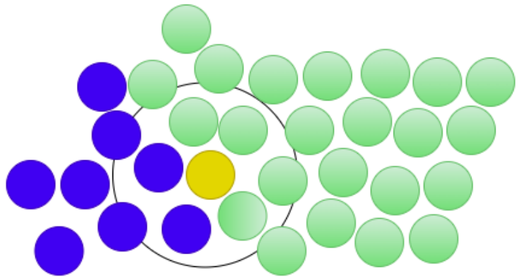


Exemplo DBSCAN (ε=R, MinPts=3)

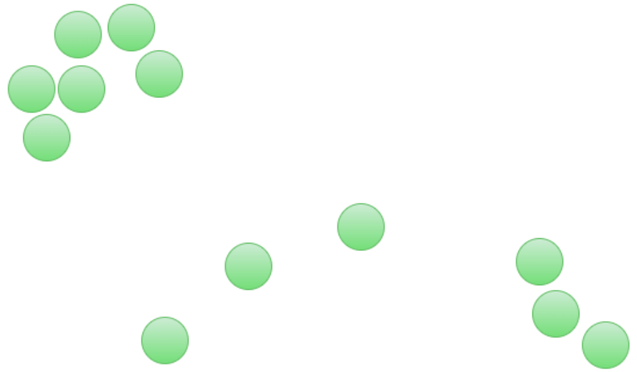


Exemplo: 44 pontos

4°Passo-Se for um ponto central, coloca os vizinhos no mesmo grupo.

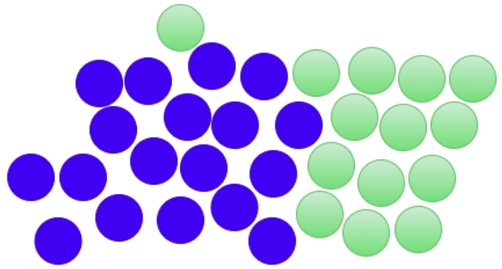


Exemplo DBSCAN (ε=R, MinPts=3)

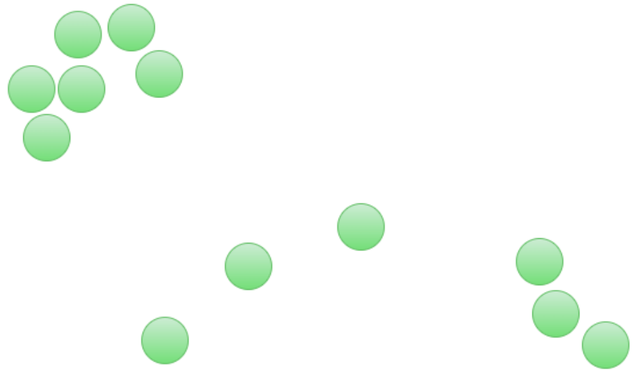


Exemplo: 44 pontos

5°Passo-Repetir até que nenhum vizinho a um ponto central seja um ponto central.

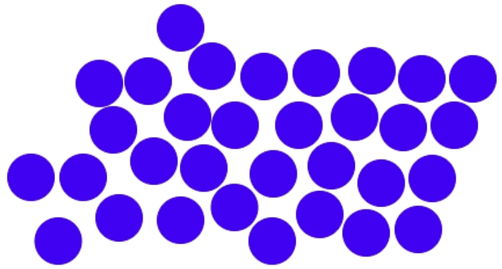


Exemplo DBSCAN (ε=R, MinPts=3)

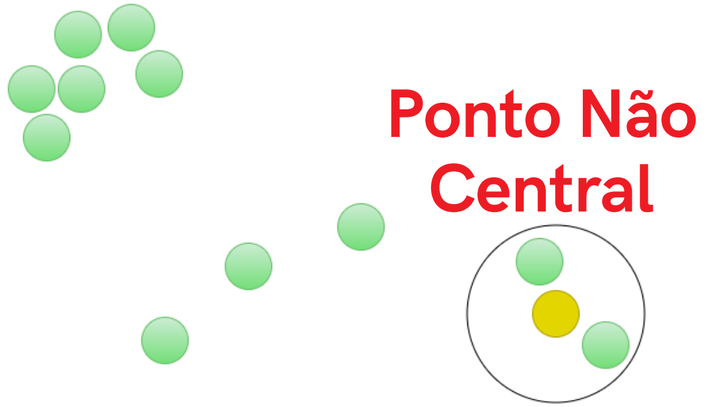


Exemplo: 44 pontos

5°Passo-Repetir até que nenhum vizinho a um ponto central seja um ponto central.

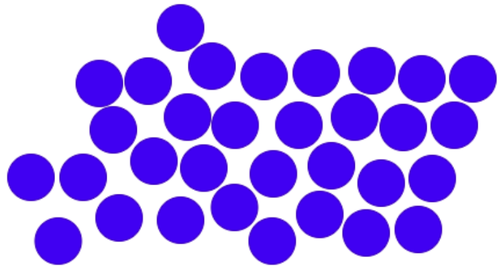


Exemplo DBSCAN (ε=R, MinPts=3)

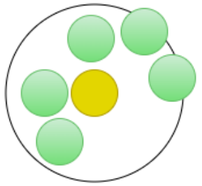


Exemplo: 44 pontos

6°Passo-Verifica outros pontos para avaliar se é um ponto central. Nenhum vizinho é central.

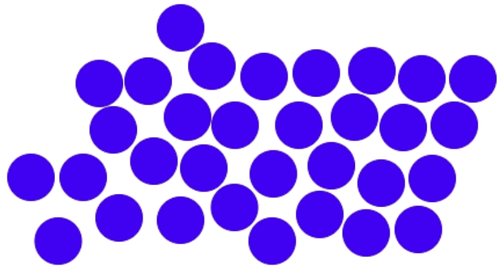


Exemplo DBSCAN (ε=R, MinPts=3)



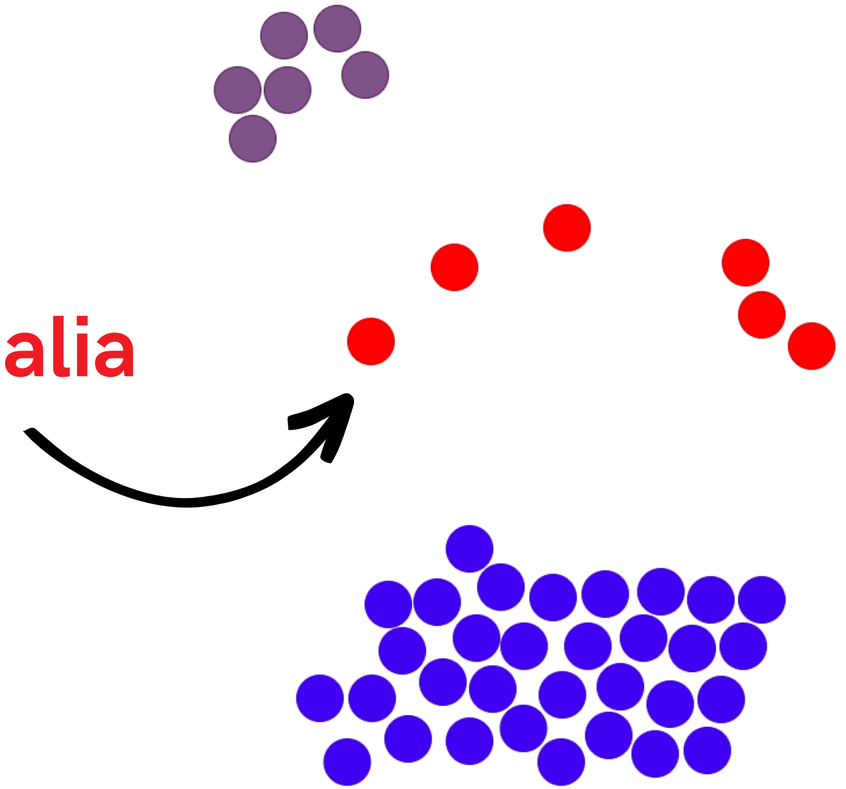
Exemplo: 44 pontos

7°Passo-Encontrou um ponto central. Inicia a formação de um novo cluster.



Exemplo DBSCAN (ε=R, MinPts=3)

Exemplo: 44 pontos



8°Passo-Pontos que não foram agrupados em nenhum cluster são anomalias.

Abordagem aplicada...

Janela Deslizante

Análise Individual dos Fluxos na Janela

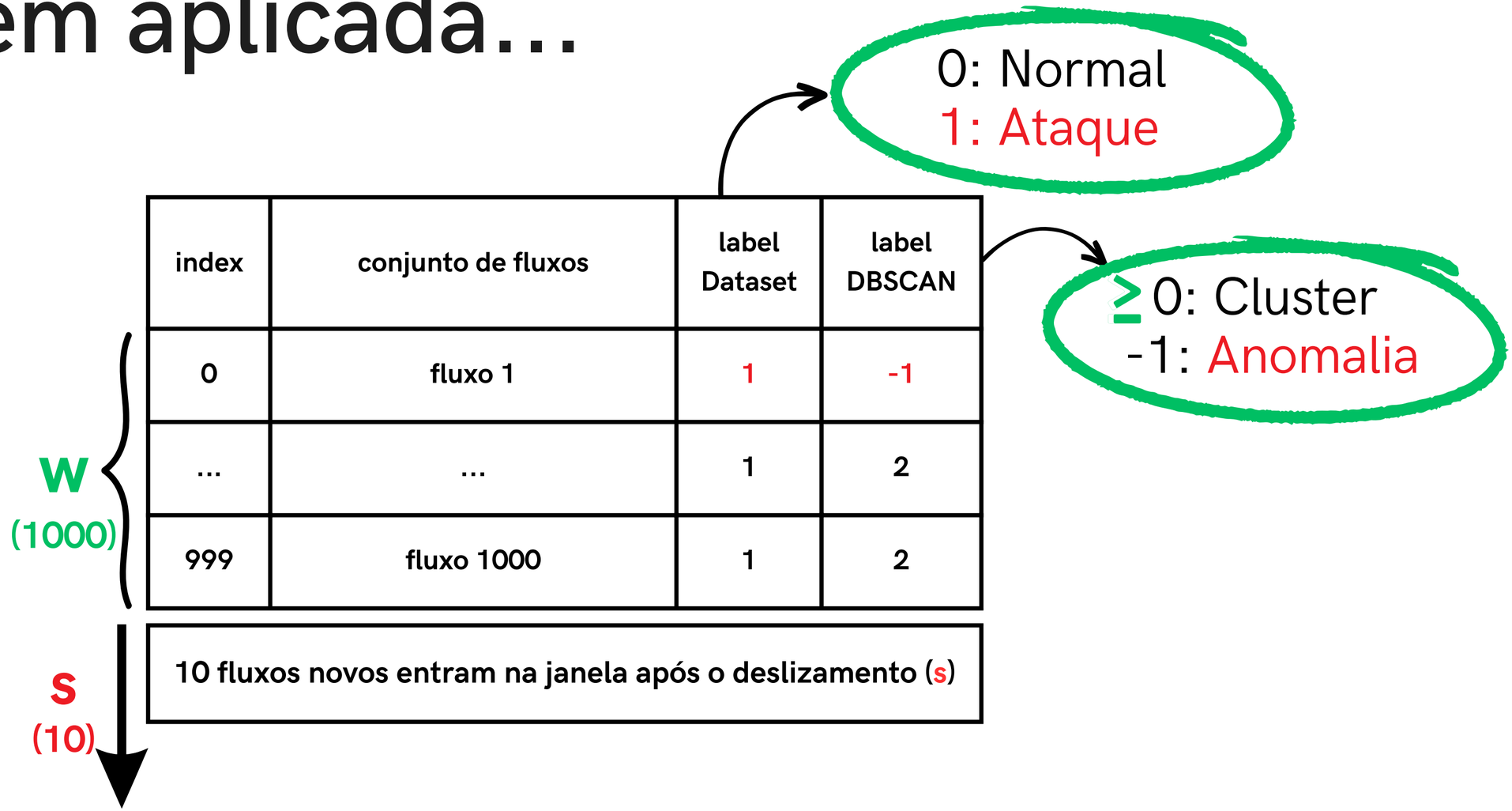


Tabela 1

Matriz de Confusão Média das Janelas

Avaliação 1

Classe Real	Previsão	
	Ataque	**Normal**
Ataque	28.41 (TP)	98.10 (FN)
Normal	68.87 (FP)	804.60 (TN)

ε = 3
MinPts = 10
Acurácia Média = 0.83

Avaliação 2

Classe Real	Previsão	
	Ataque	**Normal**
Ataque	28.55 (TP)	97.96 (FN)
Normal	178.84 (FP)	694.63 (TN)

ε = 1.5
MinPts = 2
Acurácia Média = 0.72

Tabela 1

Matriz de Confusão Média das Janelas

Avaliação 3

Classe Real	Previsão	
	Ataque	**Normal**
Ataque	51.27 (TP)	75.24 (FN)
Normal	351.87 (FP)	521.60 (TN)

ε = 1.5
MinPts = 5
Acurácia Média = 0.57

Avaliação 4

Classe Real	Previsão	
	Ataque	**Normal**
Ataque	53.81 (TP)	72.70 (FN)
Normal	492.17 (FP)	381.30 (TN)

ε = 0.95
MinPts = 2
Acurácia Média = 0.43

Anomalia Coletiva numa Janela Deslizante...

Abordagem Inicial

Avaliação Individual

- Avaliar a influência de cada fluxo novo individualmente.

Abordagem Futura

index	conjunto de fluxos	label Dataset	label DBSCAN
990	fluxo 991	1	-1
...	...	0	2
999	fluxo 1000	1	3

(10)

Avaliação Coletiva

- Avaliar a influência de um conjunto de fluxos novos quando clusterizados.

Anomalia Coletiva numa Janela Deslizante...

N+P (1000)

N (990)

index	conjunto de fluxos	label Dataset	label DBSCAN
0	fluxo 1	0	1
1	fluxo 2	0	2
...	...	0	3
989	fluxo 990	0	1

P (10)

index	conjunto de fluxos	label Dataset	label DBSCAN
990	fluxo 991	1	-1
...	...	0	2
999	fluxo 1000	1	3

Dúvidas?