POLVO-IIDS: Um Sistema de Detecção de Intrusão Inteligente Baseado em Anomalias

Paulo Manoel Mafra[1], **Joni da Silva Fraga**[1], **Vinícius Moll**[1],

Altair Olivo Santin[2]

Pós-Graduação em Engenharia de Automação e Sistemas (PGEAS)

Grupo de Computação Segura e Confiável (GCSEG)

Universidade Federal de Santa Catarina (UFSC)[1]

Pontifícia Universidade Católica do Paraná (PUC-PR)[2]

Roteiro

- Introdução
- Detecção de anomalias
- Modelo desenvolvido
- Testes
- Considerações

Introdução

Sistemas de detecção de intrusão são classificados como:

- Baseados em assinaturas
  - São muito bons para detectar intrusões conhecidas
  - Falham na detecção de novos ataques e em variantes de ataques já conhecidos
  - A geração de novas assinaturas é um processo custoso

Introdução

Sistemas de detecção de intrusão são classificados como:

- Baseados em anomalias
  - É gerado um modelo de comportamento normal do sistema
  - Detecta variações deste modelo em tempo real
  - Pode gerar muitos falsos positivos

Detecção de anomalias

Detectores baseados em anomalias podem ser:

- Baseados em sistemas especialistas
  - Possuem um conjunto de regras que descrevem o comportamento normal do sistema
- Baseados em aprendizagem
  - Aprendem o comportamento normal do sistema, geralmente de forma automática

Detecção de anomalias

Requisitos:

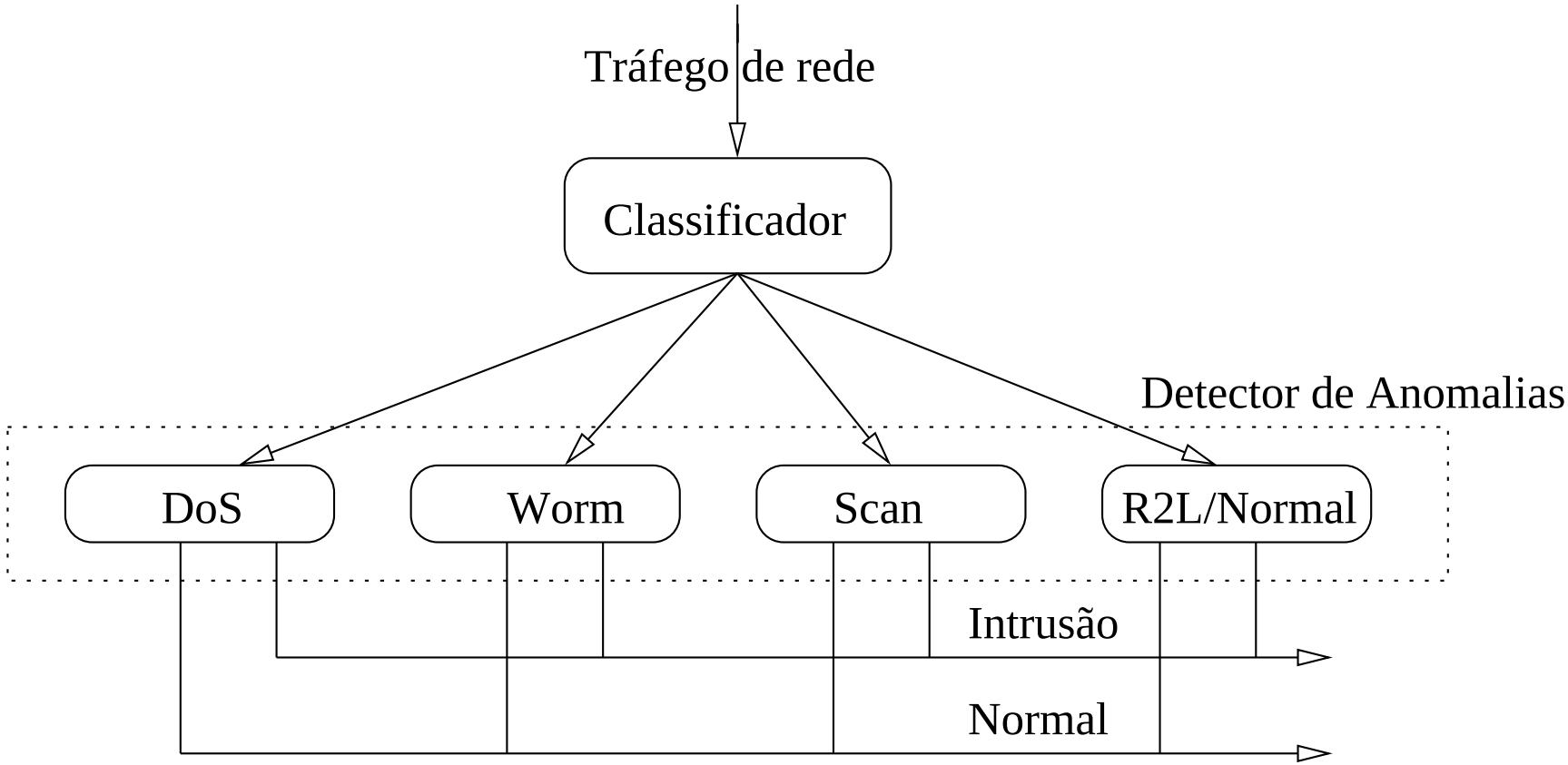
- Evolução e atualização do sistema de forma automática
- Ser genérico para qualquer serviço de rede
- Precisão na detecção de anomalias, com baixa taxa de falsos positivos
- Ser resistente a ataques similares (variações de ataques)
- Ser eficiente para operar em grandes larguras de banda, causando baixo impacto na latência da rede

Problemas presentes em IIDS

- Precisão na detecção
- Treinamento do sistema
  - Dados usados no treinamento $\neq$ dados Internet
  - Modelo de sistema adotado
- Grande quantidade de dados para analisar
- Dificuldade de atuação em ambientes reais (Internet)
- Dificuldade de resposta do IIDS

Modelo desenvolvido

Modelo geral do sistema:



Modelo desenvolvido

Classificador:

- Faz a seleção prévia do tráfego
- Usa uma rede neural de Kohonen
- Aprende padrões de forma automática
- Os padrões são genéricos
- Categorias geradas: **DoS**, **Worm**, **Scan** e **R2L/Normal**
- Rede neural com 41 entradas e 4 saídas
- Distância entre neurônios $d = \sqrt{\sum_i (v_i - w_i)^2}$
- Aprendizado por reforço
- Ajuste nos valores de aptidão de 0,9 para 0,8

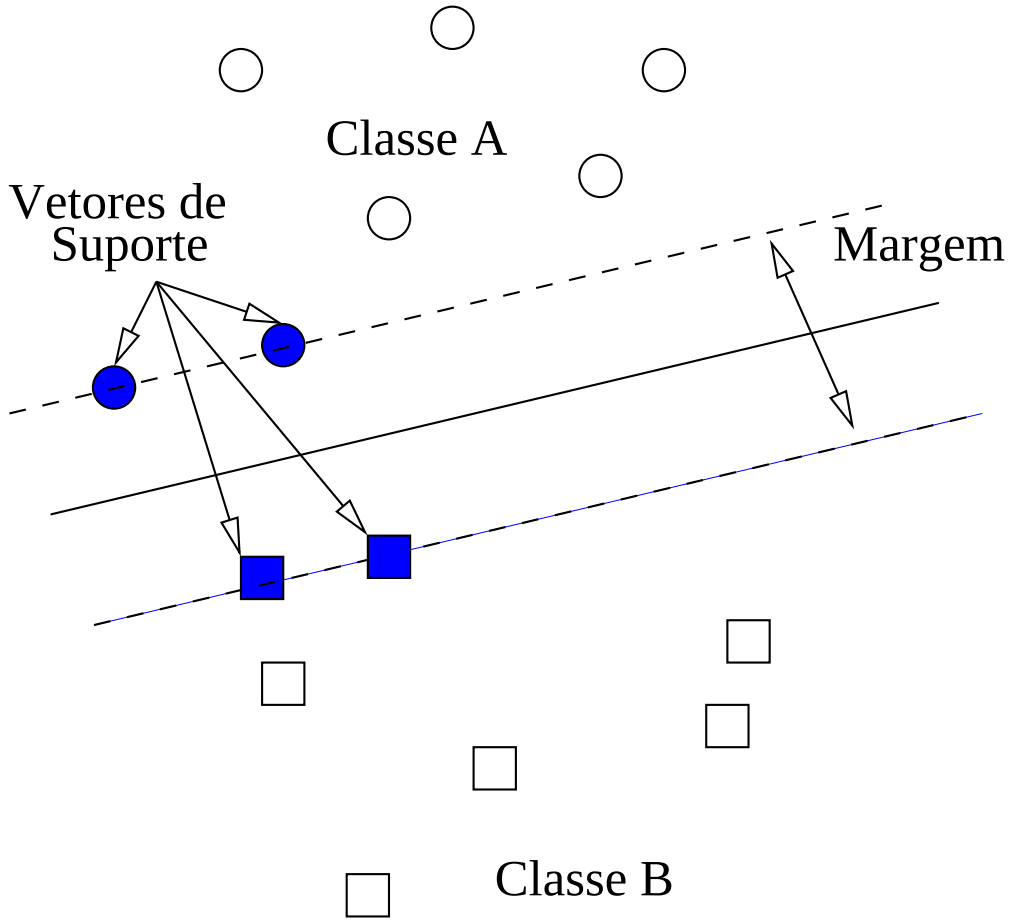
Modelo desenvolvido

Detector de Anomalias:

- Usa quatro *Support Vector Machines* (SVM)
- Cada SVM é responsável por um tipo de detecção
- SVMs suportam ruído na entrada
- Testes indicam que SVMs funcionam melhor ao separar apenas duas classes
- Simplicidade na configuração

Modelo desenvolvido

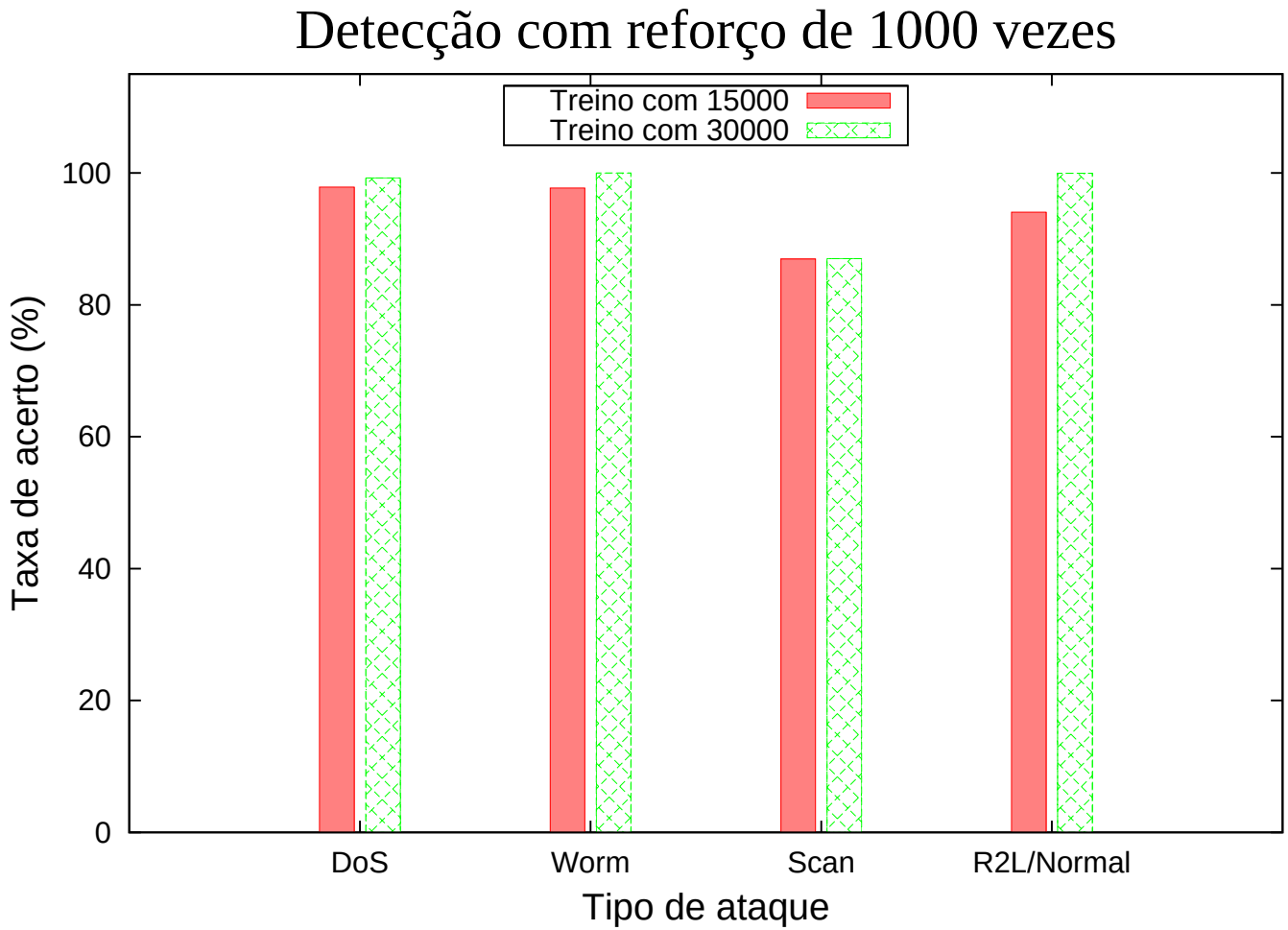
Detector de Anomalias:



Testes

- Protótipo foi desenvolvido usando o *framework* JOONE (Java Object Oriented Neural Engine)
- Hardware: um computador com 4GB de memória e dois processadores Intel Xeon de 1.6GHz
- Dados do KDD Cup 1999 para detecção de intrusão
- Treinamento: 15.000 e 30.000 entradas com reforço de 100 vezes e 1000 vezes
- Testes: 97.143 entradas (Tráfego normal, DoS, Worm, Scan ou R2L)

Testes


Detecção com reforço de 1000 vezes
Treino com 15000
Treino com 30000
Taxa de acerto (%)
0
20
40
60
80
100
DoS
Worm
Scan
R2L/Normal
Tipo de ataque

Testes

Comparação de resultados entre IIDSs:

IIDS	Acerto Médio	Desvio Máximo
Anomalous Payload-based IDS	58,80 %	41,20 %
HPCANN	77,49 %	22,53%
MADAM ID	77,97 %	17,97%
Multi-level Hybrid Classifier	89,19 %	22,52%
POLVO-IIDS	96,55 %	9,53 %

Considerações

- O emprego de redes neurais artificiais e SVM permitiu uma taxa de detecção maior
- O modelo possibilita o emprego de tráfego real
- O treinamento pode acontecer de maneira constante, mesmo com algum tráfego malicioso
- Tempo de treinamento é elevado
- Ainda não possui um mecanismo de resposta

Perguntas