UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

SELEÇÃO DE VARIÁVEIS DE REDE PARA DETECÇÃO DE INTRUSÃO

DISSERTAÇÃO DE MESTRADO

Victor Machado Alves

Santa Maria, RS, Brasil

2012

SELEÇÃO DE VARIÁVEIS DE REDE PARA DETECÇÃO DE INTRUSÃO

por

Victor Machado Alves

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do grau de

Mestre em Computação

Orientador: Prof. Dr. Raul Ceretta Nunes (UFSM)

Santa Maria, RS, Brasil

2012

Alves, Victor Machado

Seleção de Variáveis de Rede para Detecção de Intrusão / por Victor Machado Alves. – 2012.
75 p.; 30 cm.

Orientador: Raul Ceretta Nunes (UFSM)
Dissertação (Mestrado) - Universidade Federal de Santa Maria, Centro de Tecnologia, Programa de Pós-Graduação em Informática, RS, 2012.

1. Informática 2. Computação 3. Segurança 4. Detecção de Intrusão I. Nunes, Raul Ceretta. II. Título.

Ficha catalográfica elaborada através do Programa de Geração Automática da Biblioteca Central da UFSM, com os dados fornecidos pelo autor.

E-mail: victoralves@inf.ufsm.br

Universidade Federal de Santa Maria
Centro de Tecnologia
Programa de Pós-Graduação em Informática

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

SELEÇÃO DE VARIÁVEIS DE REDE PARA DETECÇÃO DE INTRUSÃO

elaborada por
Victor Machado Alves

como requisito parcial para obtenção do grau de
Mestre em Computação

COMISSÃO EXAMINADORA:

Raul Ceretta Nunes (UFSM), Dr.
(Presidente/Orientador)

Roseclea Duarte Medina, Profª. Drª. (UFSM)

Rodrigo da Rosa Righi, Prof. Dr. (UNISINOS)

Santa Maria, 22 de Outubro de 2012.

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Informática
Universidade Federal de Santa Maria

SELEÇÃO DE VARIÁVEIS DE REDE PARA DETECÇÃO DE INTRUSÃO
AUTOR: VICTOR MACHADO ALVES
ORIENTADOR: RAUL CERETTA NUNES (UFSM)
Local da Defesa e Data: Santa Maria, 22 de Outubro de 2012.

Sistemas de Detecção de Intrusão são considerados mecanismos importantes para assegurar a proteção em redes de computadores. Entretanto as informações utilizadas por estes sistemas devem estar devidamente selecionadas, pois a precisão e desempenho são sensíveis à qualidade e dimensão dos dados analisados. A seleção de variáveis para Sistemas de Detecção de Intrusão (IDS - Intrusion Detection Systems) é assim um ponto chave no projeto de IDS. O processo de seleção de variáveis, ou de características, realiza a escolha das informações apropriadas através da remoção de dados irrelevantes que interferem no resultado da detecção. No entanto, abordagens existentes para auxiliar IDS selecionam as variáveis apenas uma vez, não se adaptando as mudanças comportamentais. As variações inerentes ao tráfego de rede não são assim acompanhadas dinamicamente por estes selecionadores. Uma estratégia para reduzir a taxa de falsos alarmes em IDS baseados em anomalias é avaliar se num mesmo intervalo de tempo ocorrem mudanças abruptas em mais de uma variável de rede. Porém, esta estratégia assume como hipótese que as variáveis analisadas são correlacionadas, exigindo um procedimento prévio de seleção de variáveis. Este trabalho propõe um método dinâmico de seleção de variáveis para IDS de rede, chamado SDCorr (Seleção Dinâmica por Correlação), que opera na modalidade de filtro e utiliza como avaliador o teste de correlação de Pearson. O método adapta-se dinamicamente as variações do tráfego de rede por meio da seleção de novas variáveis a cada iteração com o detector. Assim, possibilita acompanhar as mudanças nos dados e estabelecer relações entre variáveis. Como resultado, melhora-se a precisão e desempenho do IDS através da eliminação de variáveis desnecessárias e da redução da dimensão dos dados analisados.

Palavras-chave: Segurança, Sistemas de Detecção de Intrusão, Seleção de Variáveis.

ABSTRACT

Master's Dissertation
Graduate Program in Computer Science
Federal University of Santa Maria

NETWORK FEATURE SELECTION FOR INTRUSION DETECTION
AUTHOR: VICTOR MACHADO ALVES
ADVISOR: RAUL CERETTA NUNES (UFSM)
Defense Place and Date: Santa Maria, Month 22th, 2012.

Intrusion Detection Systems are considered important mechanisms to ensure protection for computer networks. However, the information used by these systems should be properly selected, because the accuracy and performance are sensitive to the quality and size of the analyzed data. The selection of variables for Intrusion Detection Systems (IDS) is a key point in the design of IDS. The process of selection of variables, or features, makes the choice of appropriate information by removing irrelevant data that affect the result of detection. However, existing approaches to assist IDS select the variables only once, not adapting behavioral changes. The variation of the network traffic is not so accompanied by these selectors. A strategy for reducing the false alarm rate based on abnormalities in IDS is evaluating whether a same time interval abrupt changes occur in more than one variable network. However, this strategy takes as hypothesis that the variables are related, requiring a prior procedure for variable selection. This paper proposes a dynamic method of selecting variables for network IDS, called SDCorr (Selection by Dynamic Correlation), which operates in the mode filter and as an evaluator uses the Pearson correlation test. The method dynamically adapts to changes in network traffic through the selection of new variables at each iteration with the detector. Therefore allow track changes in data and establish relationships between variables. As a result, it improves the accuracy and performance of the IDS by eliminating unnecessary variables and decreasing the size of the analyzed data.

Keywords: Security, Intrusion Detection Systems, Feature Selection.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

AR	*Auto-regressive*
ARIMA	*Auto-regressive Integrated Moving Average*
CIDS	*Correlation Intrusion Detection System*
CRF	*Conditional Random Fields*
DDoS	*Distributed Denial of Service*
DoS	*Denial of Service*
GLR	*Generalized Likelihood Ratio*
HIDS	*Host-based Intrusion Detection System*
IDS	*Intrusion Detection System*
IP	*Internet Protocol*
IPS	*Intrusion Prevention System*
MIB	*Management Information Base*
NIDS	*Network-based Intrusion Detection System*
PCA	*Principal Component Analysis*
PoD	*Ping of Death*
R2L	*Remote to Local*
SDCorr	*Seleção Dinâmica por Correlação*
SNMP	*Simple Network Management Protocol*
SSGBML	*Steady State Genetic-Based Machine Learning Algorithm*
TCP	*Transmission Control Protocol*
U2R	*User to Root*

SUMÁRIO

1 INTRODUÇÃO

A Internet trouxe facilidades na área das comunicações e trocas de informações entre os usuários devido sua característica de sistema aberto e escalável (PENG; LECKIE; RAMAMOHANARAO, 2007). De acordo com Tanenbaum (2002) , uma máquina está no domínio da Internet quando executa a pilha de protocolos TCP/IP (*Transmission Control Protocol/Internet Protocol*), tem um endereço de Protocolo de Internet (*IP*) e assim pode enviar pacotes para todos os dispositivos presentes na Internet. Desta forma, qualquer dispositivo que possua as características descritas, pode enviar e receber pacotes de qualquer dispositivo presente neste domínio. Assim, as comunicações são beneficiadas pela utilização da Internet, no passo que torna-se simples o processo de trocar informações através do envio de pacotes de redes entre dispositivos, não importa onde encontrem-se no mundo. Entretanto, esta facilitação na comunicação, onde qualquer dispositivo pode enviar e receber informações de qualquer outro dispositivo computacional, tem sido utilizada como meio para realização de ações de intrusão, interferindo no comportamento adequado da rede e prejudicando empresas e usuários comuns.

Uma intrusão no âmbito da segurança de redes é caracterizada quando um usuário não legítimo obtém acesso ilegal em um sistema computacional (WANG, 2009). Exemplos de ameaças são: ataques de controle de acesso, injeção de códigos maliciosos e ataques de negação de serviço. Estes comportamentos maliciosos tem sido cada vez mais evidentes já que a complexidade e a dimensão da Internet tem crescido rapidamente. Além disto, na própria Internet é possível encontrar ferramentas automáticas para iniciar a execução de ataques, sem que o agente responsável pelo ato possua conhecimento técnico avançado para realização de tal ação (KUMAR; SELVAKUMAR, 2009). Estas ferramentas, facilitam a execução de ataques já que não exigem que os autores tenham que tomar conhecimento em profundidade em relação as características do ambiente alvo.

Como parte da estratégia de defesa para combater os ataques, métodos tradicionais como antivírus e *firewalls* são empregados. Entretanto, este tipo de linha de defesa não é adequada para, sozinha, combater ataques desta natureza (BAI; KOBAYASHI, 2003), uma vez que abordagens utilizadas por este tipo de defesa realizam ações baseadas em assinaturas ou mesmo traços de assinaturas, dificultando o processo de identificação de variações de ataques. O fluxo de rede é dinâmico, havendo variações que podem ser difíceis de prever considerando o tipo de dados de serviços que são transportados através do canal de comunicação. Para atender a

esta característica variável do tráfego de rede em relação a defesa de ataques, IDS (*Intrusion Detection System*) tem sido amplamente utilizados. Os IDS de rede têm a função de, através de dados coletados no tráfego de rede, analisar e gerar alertas quando são identificadas ações de intrusão (ABOUABDALLA et al., 2009). Entretanto, as informações utilizadas por um IDS precisam estar devidamente selecionadas, pois a precisão e desempenho da detecção são sensíveis à qualidade e dimensão dos dados analisados. A seleção de variáveis para IDS é assim um ponto chave no projeto de IDS (NGUYEN; FRANKE; PETROVIC, 2010).

1.1 Definição do Problema de Pesquisa

O processo de seleção de variáveis, ou de características, realiza a escolha das informações apropriadas através da remoção de dados irrelevantes que interferem no resultado da detecção. A detecção pode ser prejudicada devido ao grande número de informações que não contribuem para a identificação de um ataque. Desta forma, uma fase de pré-processamento para selecionar adequadamente as informações relevantes em relação aos ataques e diminuir a dimensão total dos dados analisados é desejável. Esta fase adicional permite que a detecção seja mais precisa por considerar apenas os dados que possuem informações relevantes para identificação de ataques, descartando os demais que podem interferir na forma de ruídos, prejudicando a precisão do IDS. A seleção de variáveis de rede, alvo da fase de pré-processamento diminui a dimensão dos dados contribuindo para uma melhor precisão e um menor tempo de análise e detecção.

De acordo com Khor, Ting e Amnuaisuk (2009) os métodos de seleção podem ser classificados como *wrapper*, quando encapsulam o *engine* de seleção (máquina de aprendizagem), ou como filtros, quando utilizam etapas de pré-processamento independentes de selecionador, sendo, em ambos os casos, utilizado como um meio para redução da dimensão dos dados de entrada dos IDSs, potencializando uma detecção mais eficaz. No entanto, os selecionadores propostos para auxiliar IDSs (CABRERA et al., 2002) (SUEBSING; HIRANSAKOLWONG, 2009) (NGUYEN; FRANKE; PETROVIC, 2010) (MECHTRI; DJEMILI TOLBA; GHOUALMI, 2010) realizam a seleção de variáveis previamente, não se adaptando as mudanças comportamentais abruptas provocadas por uma intrusão. As variações do tráfego de rede não são assim acompanhadas por estes selecionadores, dado que selecionam as variáveis apenas uma vez.

Esta dissertação tem foco no problema de pesquisa de seleção de variáveis para sistemas de detecção de intrusão, abordando o problema através de uma metodologia dinâmica

de escolha dos dados através de correlação.

1.2 Objetivos e Abordagem Proposta

Uma estratégia para reduzir a taxa de falsos alarmes em IDSs baseados em anomalias é avaliar se num mesmo intervalo de tempo ocorrem mudanças abruptas em mais de uma variável de rede (THOTTAN; JI, 2003) (WU; SHAO, 2005). Porém, esta estratégia assume como hipótese que as variáveis analisadas são correlacionadas, exigindo um procedimento prévio de seleção de variáveis. No trabalho de Thottan e Ji (2003), as variáveis de tráfego disponíveis na MIB foram selecionadas via análise manual de contadores de filtro, que medem o nível de tráfego na entrada e na saída de cada camada de rede (THOTTAN; JI, 1998).

Diferentemente, este trabalho propõe um método dinâmico de seleção de variáveis para IDS de rede, chamado SDCorr (Seleção Dinâmica por Correlação), que opera na modalidade de filtro e utiliza como avaliador o teste de correlação de Pearson (NETO et al., 2011). De acordo com Suebsing e Hiransakolwong (2009), abordagens que utilizam métodos de similaridade são considerados de fácil implementação. Desta forma, a proposta enquadra-se nesta categoria e adapta-se dinamicamente as variações do tráfego de rede por meio da seleção de novas variáveis a cada iteração com o detector, o que possibilita acompanhar as mudanças nos dados e estabelecer correlações entre variáveis. Como resultado, o método permite melhorar a precisão e desempenho de IDSs através da eliminação de variáveis desnecessárias e da redução da dimensão dos dados analisados. Em outras palavras, o objetivo é fazer:

- seleção dinâmica de variáveis de acordo com variações da rede, baseada na correlação entre variáveis;
- diminuição do volume inicial de dados, possibilitando análise e detecção de maneira mais eficiente;
- melhoramento da taxa de detecção, dado o uso de variáveis de maior significância e remoção de ruídos que interferem na qualidade de detecção; e
- diminuição do tempo de processamento na detecção.

A abordagem apresentada neste trabalho utiliza como pivô, para seleção de variáveis, uma variável que apresenta variação abrupta decorrente de uma intrusão. As variáveis mais prováveis de conter informações relevantes são selecionadas, refinando o conjunto de informações

disponíveis e possibilitando que a detecção crie alarmes mais precisos de acordo com correlações encontradas nas características do tráfego de rede.

Como contribuição, este trabalho traz uma metodologia para seleção das variáveis que pode ser utilizada em estratégias de detecção que exigem que as variáveis analisadas sejam correlacionadas, tal como a proposta em Thottan e Ji (2003).

1.3 Estrutura do Texto

No Capítulo 2 são apresentados definições e classificações de IDSs, bem como realizada uma síntese em torno dos principais métodos e preocupações no campo de pesquisa dos IDSs.

O Capítulo 3 aborda os aspectos relacionados com a seleção de características para utilização dos dados em sistemas de detecção de intrusão. Pontos importantes são levantados neste capítulo através da análise de trabalhos realizados na área.

No Capítulo 4, a nova abordagem de seleção de características de rede para detecção de intrusões é apresentada, além de considerações em relação a implementação. O ambiente de detecção utilizado para aplicação da nova proposta também é apresentado neste capítulo, onde características e funcionamento são abordados.

No Capítulo 5, os experimentos para avaliação e validação são desenvolvidos. Aspectos da qualidade de detecção e volume de dados são discutidas.

Já no Capítulo 6 são apresentas as conclusões e contribuições do trabalho, assim como possíveis caminhos para trabalhos futuros.

2 SISTEMAS DE DETECÇÃO DE INTRUSÃO

Uma intrusão em um sistema computacional é definida como uma tentativa não autorizada de violação do sistema. Os sistemas de detecção de intrusão (IDS - *Intrusion Detection Systems*) são mecanismos utilizados para promover segurança frente a ações de intrusões. Para cumprir este papel, desempenham tarefas de monitoramento, análise e detecção. Este capítulo apresenta uma visão geral dos IDSs, no que diz respeito aos conceitos básicos (seção 2.1) e a sua classificação (seção 2.2), e realiza uma breve discussão a respeito das principais questões e desafios na área de projeto de IDSs (seção 2.3).

2.1 Conceitos Básicos

Intrusões em infraestruturas de rede são um dos principais problemas na segurança de redes de computadores (BAI; KOBAYASHI, 2003), pois decorre do aumento das redes e de ferramentas projetadas para executar intrusões. Sistemas de Detecção de Intrusão (IDS) são mecanismos de segurança que fazem uso dos dados provenientes dos ambientes que protegem, realizando a análise da informação para detecção de atividades não autorizadas (SABAHI; MOVAGHAR, 2008).

Como parte da linha de defesa para proteção contra intrusões, também existem métodos tradicionais como antivírus, criptografia e *firewalls*. Entretanto, estes métodos são estáticos, sendo limitados para defender sistemas contra intrusões. Diferentemente, os IDSs auxiliam para melhorar a defesa proporcionada por tais métodos (LIM; JONES, 2008) e, adicionalmente, são mecanismos apropriados a ambientes dinâmicos (EKTEFA et al., 2010) (BAI; KOBAYASHI, 2003). Os IDSs aumentam o nível de segurança porque atuam como uma ferramenta de monitoramento e análise dos dados diretamente relacionados com o ambiente ao qual defendem (BAI; KOBAYASHI, 2003).

O principal objetivo de sistemas de detecção de intrusão é identificar ou prevenir tentativas não autorizadas de acesso ou comprometimento do sistema, maximizando a taxa de acertos (verdadeiros positivos) e minimizando os alarmes falsos (falsos positivos) (BAI; KOBAYASHI, 2003) (MARHUSIN; CORNFORTH; LARKIN, 2008) (SPEROTTO et al., 2010) (KR; INDRA, 2010). Entretanto, além da precisão (sistema que detecta muitos ataques e resulta em poucos falsos alarmes) é desejável que ele trate de um grande volume de dados e seja rápido o bastante para tomar decisões num tempo aceitável (GUPTA; NATH; KOTAGIRI, 2010).

Os IDSs podem identificar intrusões, que já ocorreram ou que estão em curso, notificando, na forma de alarmes automatizados, os responsáveis para que as medidas de recuperação ou reação sejam tomadas. A metodologia de detecção básica desempenhada por um IDS é capturar informações em pacotes de rede ou em arquivos de log do sistema e desempenhar a detecção com base em métodos apropriados. De maneira geral, invasores frequentemente atuam de forma diferente de usuários legítimos pelos quais tentam se passar (WANG, 2009).

Esta diferença pode ser utilizada em análises quantitativas que executam a detecção com base em desvios do comportamento normal. Uma abordagem tradicional desempenhada por IDSs deste tipo é procurar por formas de identificar eventos considerados anormais, ou seja, por variações no comportamento do invasor em relação ao de um usuário legítimo. Em outras palavras, um IDS depende de métodos apropriados que podem basear-se na operação do sistema, protocolos de rede, abordagens estatísticas e mineração de dados.

De forma sintética, dentre as principais atividades desenvolvidas por um IDS´s tem-se (SABAHI; MOVAGHAR, 2008)(BAI; KOBAYASHI, 2003):

- monitoramento e análise das atividades de usuários;
- auditagem da estrutura do sistema de falhas;
- reconhecimento do modelo de atividades, identificando assinaturas de intrusões;
- análise estatística para identificação de anomalias;
- avaliação da integridade do sistema e/ou de informações críticas;
- análise de comportamento dos usuários de acordo com as políticas de segurança;

A avaliação dos IDSs é uma tarefa importante e os parâmetros para avaliar sua eficiência podem ser descritos por Debar (2000): precisão, desempenho, completude, tolerância a falhas e pontualidade.

A *precisão* indica a capacidade do IDS em detectar intrusões sem a presença de alarmes falsos. Neste sentido, uma detecção não é considerada precisa quando o IDS marca uma ação legítima como intrusão. Quando isto ocorre produz situações indesejadas, já que a emissão de um alarme foi realizada sem que fosse necessário, prejudicando a confiabilidade deste sistema de proteção.

O parâmetro *desempenho* avalia o IDS em relação a taxa de processamento das informações. Quando o desempenho é considerado bom, a velocidade da detecção permite que o IDS seja

projetado para operação em tempo real. Diferentemente, em caso de desempenho menor, é adequado para operação *off-line*, informando violações que já ocorreram.

A *completude* avalia a capacidade do IDS de detectar todos os ataques que ocorreram. Esta medida é considerada difícil de avaliar pois exige o conhecimento de todos os ataques que ocorreram.

A *tolerância a falhas* também é uma medida de avaliação de IDSs e refere-se a propriedade do IDS ser resistente, às investidas realizadas por invasores. Este parâmetro pode ser expressado também como *tolerância a intrusão*, se a falha for consequência de falhas do tipo maliciosa.

A *pontualidade* está relacionada com a medida desempenho. No entanto, o desempenho refere-se apenas a velocidade de processamento enquanto a pontualidade relaciona-se também com a capacidade de propagação da detecção, realizando a identificação de intrusões tão rápido quanto possível e propagando este conhecimento às partes responsáveis. Isto permite que a manutenção, correção ou mesmo reação sejam executadas antes que danos críticos sejam causados.

A arquitetura básica de um IDS típico é composta pelos seguintes componentes (MARHUSIN; CORNFORTH; LARKIN, 2008): avaliação, detecção e alarme. O componente *avaliação* é responsável por criar um perfil de segurança aceitável com base nas políticas e necessidades exigidas para proteção do sistema. O componente *detecção* desempenha a tarefa de coletar e analisar dados referentes ao sistema que protege, identificando desvios ou mesmo comportamentos inaceitáveis dentro do sistema. Exemplos de dados utilizados por este componente são logs de operação do sistema e pacotes de rede. O componente *alarme* sinaliza uma intrusão. O alarme pode corresponder a um aviso ao responsável pelo sistema ou a um outro sistema que seja responsável pelas ações de recuperação e reação.

2.2 Classificação

Os IDSs geralmente são classificados conforme os seguintes critérios: de acordo com a fonte de dados utilizada, pelo método de detecção e pelo tipo de reação diante uma ação de intrusão. De acordo com a fonte de dados pode ser denominado como baseado em *host* ou baseado em rede. Já de acordo com o método de detecção pode ser denominado baseado em assinatura, anomalia ou ainda um tipo híbrido. Relacionado com o tipo de reação diante uma intrusão é definido como passivo ou ativo. A Tabela 2.2 sintetiza a classificação mencionada e as próximas subseções explicam com maior profundidade a classificação por critério.

Tabela 2.1: Classificação dos IDS

Classificação dos IDS	Subcategorias de IDS
Fonte de Dados	Baseado em *host*
	Baseado em rede
	Híbrido
Métodos de Detecção	Assinatura
	Anomalia
	Híbrido
Reação a Intrusão	Passivo
	Ativo

2.2.1 Fonte de Dados

Os sistemas de detecção de intrusão podem ser classificados de acordo com a origem dos dados utilizados para detecção (SABAHI; MOVAGHAR, 2008). Estes dados são importantes na identificação da presença de uma intrusão, na investigação de incidentes ou mesmo na correlação de eventos entre IDS e demais fontes de dados. Um IDS pode ser enquadrado em três categorias básicas de acordo com a fonte dos dados utilizados: baseado em *host* (HIDS), baseado na rede (NIDS) ou híbrido. Cada uma das categorias possui uma abordagem diferente, com vantagens e desvantagens.

O modelo baseado em *host* monitora as características e atividades de um único *host*, na busca por atividades suspeitas que possam ser passíveis de um comportamento de intrusão. Isto é realizado através do monitoramento de eventos do sistema e comportamento dos usuários. A utilização de HIDS possui vantagens, por exemplo (BAI; KOBAYASHI, 2003): não exige a utilização de hardware adicional; os mecanismos de detecção e resposta são adequados a aplicações em tempo real; é ideal para ambientes criptografados. Em geral, HIDS permite que os dados coletados reflitam precisamente o que ocorre no *host* em questão. Entretanto, é limitado em relação ao que acontece na rede de intercomunicação, tendo informações apenas do dispositivo de rede no qual está instalado.

O IDS que trabalha com base na rede (NIDS), monitora o tráfego de informações na rede, dado um segmento particular o qual é responsável, analisando o comportamento da rede e aplicações que fazem uso do canal de comunicação. O NIDS busca por atividades suspeitas nor-

malmente por análise dos protocolos e pacotes. A utilização de IDS baseado em rede possui algumas vantagens e desvantagens (WANG, 2009). Vantagens: baixo custo, pois em um rede de larga escala, devido a instalação de sondas em um número pequeno de pontos selecionados, permite o monitoramento de toda a rede; não provoca interferência devido ao monitoramento ser desempenhado de forma passiva, encaminhando os pacotes para detecção e análise, sem interferir no tráfego normal da rede. Desvantagens: um NIDS pode não ser apropriado para processar um grande volume de dados, gerando alarmes falsos e prejudicando a taxa de acertos; possui dificuldade em detectar ataques que sejam segmentados.

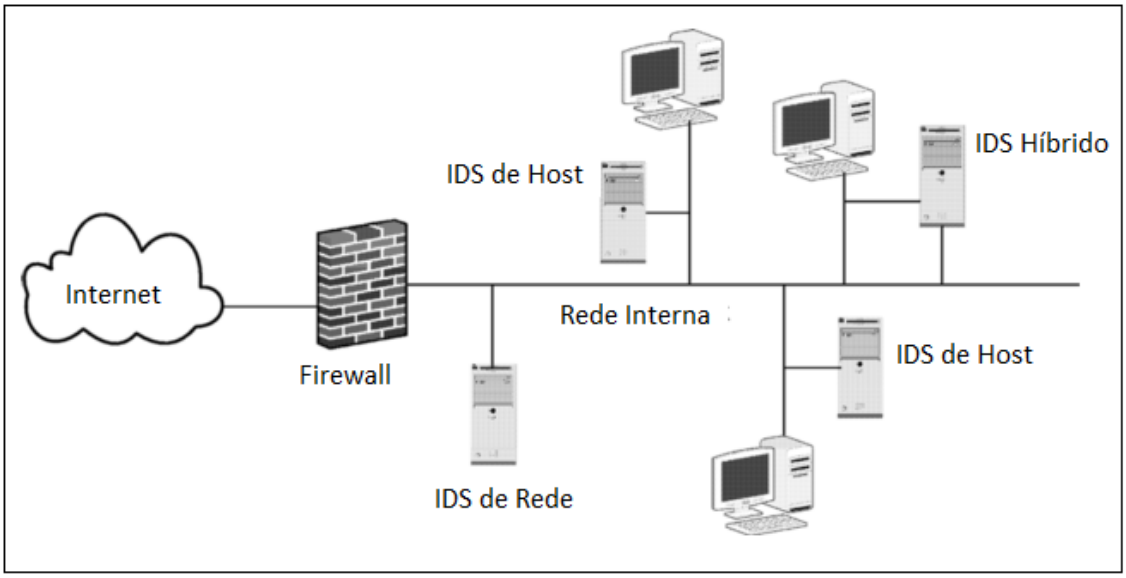

Figura 2.1: Posição de Diferentes Tipos de IDS, adaptado de Wang (2009)

Um IDS híbrido funciona desempenhando as atividades a partir das duas fontes de dados descritas (host e rede), de maneira simultânea, ou seja, é responsável por monitorar o comportamento de um *host* assim como da rede na qual está inserido. Este tipo é importante quando existe a necessidade de trabalhar com redes criptografadas e os dados destinados a um *host* em particular. Quando há criptografia, somente a origem e o destino podem ver qual é a informação que foi enviada. Portanto, o IDS híbrido trabalha com o uso de agentes não móveis, responsáveis por processar o tráfego de rede para que o *host* possa compreender as informações recebidas.

A Figura 2.1 mostra a posição de diferentes tipos de IDS de acordo com a fonte de dados utilizada.

2.2.2 Métodos de Detecção

Além da classificação de acordo com a origem dos dados, os IDSs também podem ser classificados de acordo com a abordagem utilizada no mecanismo de detecção para processamento da informação. Três são as abordagens: detecção de assinatura, detecção de anomalia e detecção híbrida (ABOUABDALLA et al., 2009) (MURALI, 2005) (GUPTA; NATH; KOTAGIRI, 2010).

A detecção de assinatura tem o objetivo de confrontar o comportamento analisado com comportamentos disponíveis em uma base de comportamentos relacionados aos ataques (base de assinaturas de ataques). Dessa forma, o método procura por padrões de ataques conhecidos para identificar violações de segurança. Um alarme é gerado caso a assinatura do ataque seja a mesma do comportamento que foi comparado. Por mais de uma década a detecção por assinatura tem se mostrado uma estratégia dominante para IDSs (LIM; JONES, 2008), principalmente pela confiabilidade da detecção. O principal desafio neste método é encontrar intrusões quando há mutação nas assinaturas. O método não é eficiente na detecção de intrusões novas ou desconhecidas (BAI; KOBAYASHI, 2003).

Já a detecção por anomalia funciona baseada no conceito de que variações no comportamento usual podem estar relacionadas com intrusões. Métodos de detecção por anomalia propõem algoritmos que analisam variações no conjunto de informações coletadas, onde métodos são empregados para detectar desvios do comportamento normal conhecido. Assim, uma das preocupações no projeto de um detector baseado em anomalias é distinguir entre comportamentos legítimos e comportamentos anômalos relacionados com intrusões (SABAHI; MOVAGHAR, 2008). A detecção por anomalia determina a anormalidade medindo a distância entre variações de comportamento entre atividades suspeitas e atividades normais através de um limite (*threshold*). Dependendo como é gerenciado este *threshold* pode-se reduzir ou exceder o número de acertos e falsos positivos. Adicionalmente, não é necessário um conhecimento específico das intrusões e assim a detecção é capaz de identificar novos ataques. Um problema latente neste método é quando as atividades intrusivas confundem-se com atividades normais anômalas, podendo gerar uma das seguintes situações: atividades anômalas que não são intrusivas são marcadas como comportamentos intrusivos, resultando em falsos positivos; atividades intrusivas que não são anômalas resultam em falsos negativos, quando a intrusão não é detectada (BAI; KOBAYASHI, 2003).

O método de detecção híbrido, une os dois tipos apresentados anteriormente. Considera

padrões normais e anormais para treinar o sistema e assim realizar a classificação dos dados de teste. Os IDSs híbridos podem ser eficientes, sujeitos ao método de classificação dos dados usados e pode também ser usado para rotular novas variações de classes de ataques conhecidos. Isto é possível devido ao aprendizado durante o treinamento que permite ao sistema aprender características de todos os casos.

2.2.3 Reação a Intrusões

Quando um IDS identifica uma intrusão deve executar alguma função para que a detecção alcance o objetivo de proteger o sistema. Dessa forma, IDSs são divididos em duas categorias em relação a reação na presença de intrusões: reação passiva e reação ativa (SPEROTTO et al., 2010). Quando a reação determina ações corretivas ou proativas, o sistema é considerado ativo. Caso contrário, se apenas emite alarmes quando ocorre uma intrusão, é chamado passivo.

Geralmente, IDSs reagem diante de ataques de forma passiva. O modo passivo apenas informa o administrador sobre eventos maliciosos, sem qualquer outra medida. Neste caso, o mais importante é a velocidade de notificação quando ocorre uma intrusão. Diferentemente, alguns IDSs podem desempenhar reação ativa quando intrusões ocorrem, respondendo a eventos críticos. Entretanto, reações ativas geralmente não executam medidas perfeitas diante de intrusões, isto porque é necessária a utilização de mais recursos e concentração para definição das tarefas que devem ser desempenhadas, o que não é recomendado por tornar-se mais caro computacionalmente (SABAHI; MOVAGHAR, 2008).

Os sistemas que operam de forma ativa são frequentemente referenciados como sistemas de prevenção a intrusão (IPS). Algumas das tarefas que um IPS pode realizar quando detecta um ataque são (WANG, 2009): modificar regras de perímetro de rede, isolar segmentos afetados, ou mesmo encerrar serviços. É importante salientar que um IDS típico que desempenha função passiva não possui tarefas de prevenção a intrusão, tendo como responsabilidades apenas reportar os eventos ao administrador (SABAHI; MOVAGHAR, 2008).

2.3 Questões e Desafios

Quando um IDS identifica uma atividade legítima como suspeita de intrusão, diz-se que houve um falso positivo, ou seja, o IDS comete um erro. Deste modo, a indicação de falsos positivos (ABOUABDALLA et al., 2009) é um problema para os IDSs. Um falso negativo ocorre quando uma intrusão não é identificada pelo IDS, o que também não é desejável. Os

falsos positivos e falsos negativos normalmente são medidas inversamente proporcionais, pois para redução da detecção de falsos positivos as políticas de segurança podem ser refeitas com o objetivo de aceitar maior variabilidade nas atividades como normais, o que significa que será detectado menos atividades anormais, resultando em um aumento de detecção de falsos negativos. Por outro lado, para redução de falsos negativos a revisão das políticas pode ser ajustada para aceitar menor variabilidade nas atividades normais, ou seja, existe maior chance de atividades normais serem detectadas como anormais, resultando em um aumento de detecções de falsos positivos. Assim, a formulação e calibragem das políticas para balanceamento entre taxas de falsos positivos e falsos negativos é um dos desafios na área de IDSs (WANG, 2009).

Geralmente o método de detecção de anomalia encontra dificuldades, especialmente em ambientes de larga escala. A detecção por anomalias depende do fluxo dos dados, o qual é frequentemente reportado em intervalos grandes de tempo. Esta latência pode propagar-se às máquinas do ambiente, sendo comprometidas antes do ataque ser detectado. Assim, é necessário um método que trate de grandes volumes de dados com menos perdas de informação (LIM; JONES, 2008). A eficiência da detecção depende do comportamento normal da rede que é modelado. Diversas técnicas, tem sido pesquisadas e propostas para melhorar a eficiência do perfil gerado. Entretanto, o comportamento normal da rede pode ser subjetivo e as anomalias podem não ser bem definidas devido a dificuldade de modelagem do perfil de comportamento normal que é subjetivo. A detecção por anomalia encontra desafios no campo de redução do número de falsos positivos. Isto provoca nos usuários uma falsa percepção, por conta do grande número de falsos positivos.

Dentre as técnicas utilizadas na detecção de intrusão, a mineração de dados tem sido usada como passo benéfico para solução de vários problemas em diferentes questões como, por exemplo, no processamento de grande quantidade de dados e na seleção apropriada de variáveis para o processo de detecção (EKTEFA et al., 2010). Além do mais, sistemas que utilizam seleção dos dados possibilitam executar a sumarização das informações e visualização o que ajuda na análise e segurança de sistemas computacionais.

A utilização de correlação para encontrar comportamentos implícitos no conjunto de dados é uma estratégia para identificar anomalias. O trabalho de Azevedo et al. (2011) explora a correlação entre descritores de tráfego de rede através da transformada *Wavelet 2D*.

Em sistemas de segurança é comum a emissão de alarmes de acordo com as atividades de usuários, protocolos e demais variáveis individualmente. Os alarmes, independentes, podem

ser confrontados para identificar a presença de um ataque com mais precisão, diminuindo o número de falsos positivos. Esta abordagem foi utilizada nos trabalhos de Thottan e Ji (2003) e Wu e Shao (2005). Entretanto, mesmo com a utilização desta abordagem, este procedimento pode resultar num número indesejado de falsos positivos, pois a qualidade dos dados de entrada interfere no resultado (ABOUABDALLA et al., 2009).

Quando os dados de entrada não são apropriados, não possuem conteúdo relacionado a ações de intrusão ou possuem um volume proibitivo para tratamento, o IDS terá dificuldades para identificar verdadeiros positivos e evitar falsos positivos. Os dados devem ser assim preparados para apresentar conhecimento útil ao IDS, potencializando a detecção de intrusões.

A utilização de grandes volumes de dados e/ou de dados inadequados provocam ruídos na saída do IDS, prejudicando o funcionamento desejado. Adicionalmente, o processamento dos dados na sua forma original demanda tempo e carga computacional, o que interfere no tempo de resposta dos alertas emitidos pelos IDS.

Neste contexto, a seleção de variáveis é um ponto importante no projeto de IDS, pois permite reduzir o número de variáveis de entrada, melhorar a precisão da detecção e diminuir o tempo de processamento.

2.4 Conclusões Parciais

Os sistemas de detecção de intrusão são ferramentas que auxiliam na linha de defesa contra comportamentos abusivos. Dentre os tipos de IDSs abordados, destacam-se os que funcionam através da análise de anomalias na rede. Este tipo é caracterizado pela capacidade de detectar tipos novos de ataques, sem a necessidade de especificação de assinaturas. Seu funcionamento depende do conhecimento do perfil normal de comportamento da rede. Comportamentos que diferem deste perfil de normalidade são considerados anomalias (alertas).

Uma questão importante que interfere na qualidade de detecção de um IDS é o volume e a qualidade dos dados utilizados. O grande volume de dados provoca atrasos na detecção e pode interferir na qualidade dos resultados devido a ruídos que atrapalham a detecção. A alimentação do IDS com dados irrelevantes provoca maior dificuldade na identificação de uma intrusão. Dados relevantes são aqueles que estão alinhados com a necessidade IDS, ou seja, são filtrados para sua utilização, removendo os que não são importantes. Em IDS que necessitam de informações correlacionadas, esta propriedade é utilizada para definir a relevância das informações através de um processo de seleção de variáveis. A redução da quantidade de dados,

selecionando os mais adequados é um ponto chave no projeto de IDSs.

3 SELEÇÃO DE CARACTERÍSTICAS PARA DETECÇÃO DE INTRUSÃO

Este capítulo traz um estudo relacionado a seleção de características (variáveis) para detecção de intrusão. A seleção de características é definida como uma fase de pré-processamento dos dados para utilização deste conjunto de informação em IDS. Este processo tem o objetivo de selecionar as características mais apropriadas para o processo de detecção, diminuindo o volume de dados e retirando ruídos que prejudicam a qualidade da taxa de acertos e o tempo de processamento. A Seção 3.1 aborda conceitos importantes da área de seleção de características; a Seção 3.2 apresenta as classificações utilizadas; a Seção 3.3 traz abordagens para resolução de problemas na área através de um estudo de trabalhos encontrados na literatura.

3.1 Conceitos Básicos

A seleção de características, também conhecida como seleção de subconjuntos ou seleção de variáveis, é um passo importante de pré-processamento de dados (NGUYEN; FRANKE; PETROVIC, 2010), onde o subconjunto de características disponíveis é selecionado para posterior aplicação. A seleção mostra-se necessária quando há informações escondidas no volume de dados e mesmo pelo desnecessário e dispendioso processo de análise do conjunto original, que comporta um grande volume de informações. Esta seleção consiste na identificação das características relevantes presentes no conjunto de informações, considerando as mais importantes e descartando aquelas identificadas como irrelevantes. As vantagens estão relacionadas com o desempenho dos algoritmos de seleção, redução do volume de dados e simplicidade (BOLÓN-CANEDO; SANCHEZ-MAROO; ALONSO-BETANZOS, 2009). O desempenho é alcançado por meio da interpretação dos dados, reconhecendo traços e comportamentos que estão escondidos no grande volume de informações. A redução da quantidade de dados auxilia na economia dos recursos de processamento, diminuindo custos. Além disto, a simplicidade inerente aos métodos permite o uso de modelos simples e eficientes, ganhando velocidade no processo de seleção das características.

A seleção é considerada um passo de pré-processamento que implementa o aprendizado dos dados disponíveis para seleção do subconjunto mais apropriado de características de acordo com o método de detecção de intrusão que será utilizado. Assim, o espaço de características é reduzido através de um critério de avaliação (SHEEN; RAJESH, 2008). Os dados originais

normalmente tem grande volume. Assim, é desejável que seja selecionado um subconjunto que represente de forma adequada o volume original de dados. O procedimento é executado considerando o mínimo de atributos necessários para representar com precisão o conjunto integral de informações. É importante também pela remoção de ruídos provocados por variáveis que não são utilizadas (EL-KHATIB, 2010).

Para a construção de modelos de IDS, a seleção das características torna-se um passo crucial (EL-KHATIB, 2010). Durante o processo, o conjunto de atributos mais importantes deve ser considerados para a construção de algoritmos de detecção ajustáveis. O problema chave neste campo de pesquisa é escolher o conjunto ótimo de atributos, já que nem todas as características são relevantes para o algoritmo de detecção e em alguns casos representam partes irrelevantes ou mesmo redundantes que podem introduzir ruídos que prejudicam a capacidade de detecção dos algoritmos (HOON; SHIN; CHUNG, 2006). Avanços nos algoritmos de seleção de características tem ajudado IDSs a desempenhar papéis mais robustos, tornando a detecção mais precisa. O desafio é converter este volume de informações em conhecimento útil para aplicação.

Características redundantes podem reduzir o desempenho do sistema de reconhecimento de padrões (NZIGA, 2011). Entretanto, a redundância de características em detectores que baseiam-se em correlação, funciona auxiliando o processo de filtragem de alarmes, onde a correlação de características indica propagação de anomalias no fluxo de dados, ajudando na diferenciação de comportamentos legítimos e os provocados por intrusões. Desta forma, o método de avaliação de subconjuntos de características deve ser adequado ao método de detecção de intrusão empregado.

O desafio está em selecionar características em um IDS, escolhendo medidas apropriadas que consigam precisamente determinar a relevância e o relacionamento entre as características de um conjunto de dados (NGUYEN; FRANKE; PETROVIC, 2010). A Figura 3.1 apresenta o processo básico de seleção de características, onde o conjunto original passa pelo método de seleção e a saída corresponde às características selecionadas pelo processo.

3.2 Classificação

Os métodos utilizados para realização da seleção de características são classificados de acordo com sua atitude em relação ao desempenho da classificação (KHOR; TING; AMNUAISUK, 2009). Quando há envolvimento com o desempenho da classificação, o método é denominado *wrapper*. Diferentemente, quando não existe relação com o desempenho é chamado de

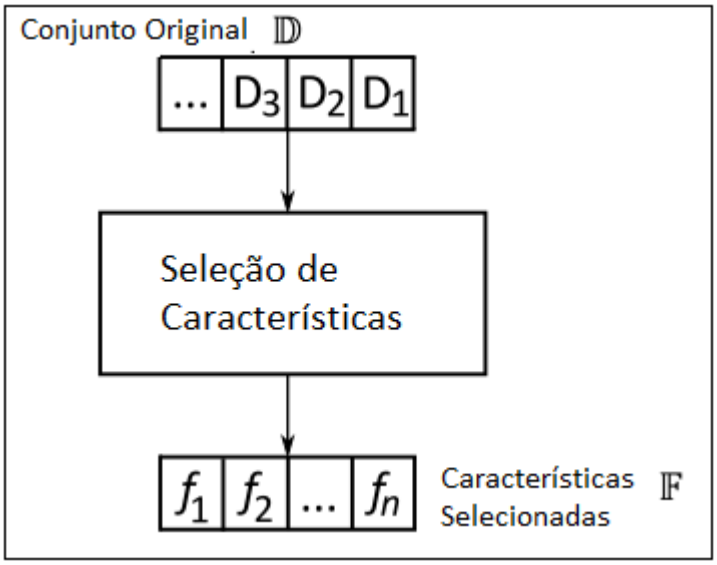

Figura 3.1: Processo de Seleção de Características, adaptado de Nguyen, Franke, Petrovic (2010)

filtro. A seguir estes dois tipos de métodos de seleção de características serão abordados.

3.2.1 *Wrappers*

Os métodos do tipo *wrapper* envolvem-se na otimização do resultado da seleção como parte do processo de identificação das informações mais relevantes. Métodos baseados neste tipo de seleção, avaliam quanto cada característica contribui para o processo, utilizando algoritmos de aprendizagem. A seleção por um método *wrapper* termina quando um conjunto ótimo é gerado. Desta forma, o resultado é sempre monitorado com o objetivo de alcançar o melhor subconjunto. Uma preocupação latente na utilização de métodos do tipo *wrapper* é o grande consumo de recursos computacionais (NZIGA, 2011). Isto ocorre devido a sobrecarga que normalmente pode ocorrer em algoritmos de aprendizagem, provocando perda de desempenho na detecção de intrusão.

Assim, *wrappers* levam em consideração a natureza do modelo utilizado para seleção das características, sendo diretamente ligado ao efeito que a seleção deste subconjunto terá no desempenho da aplicação dos dados escolhidos. Portanto, uma vez disponível o subconjunto de treinamento, este método de seleção procura sempre por um resultado que seja capaz de maximizar o desempenho através de uma seleção ótima dos dados. Entretanto, apesar de resultar em uma seleção de dados considerada a melhor possível, estes métodos são caros e provocam perdas de desempenho do detector, devido à alterações na velocidade da detecção provocadas pelo tempo necessário para que o selecionador consiga alcançar o melhor conjunto de variáveis

(NGUYEN; FRANKE; PETROVIC, 2010).

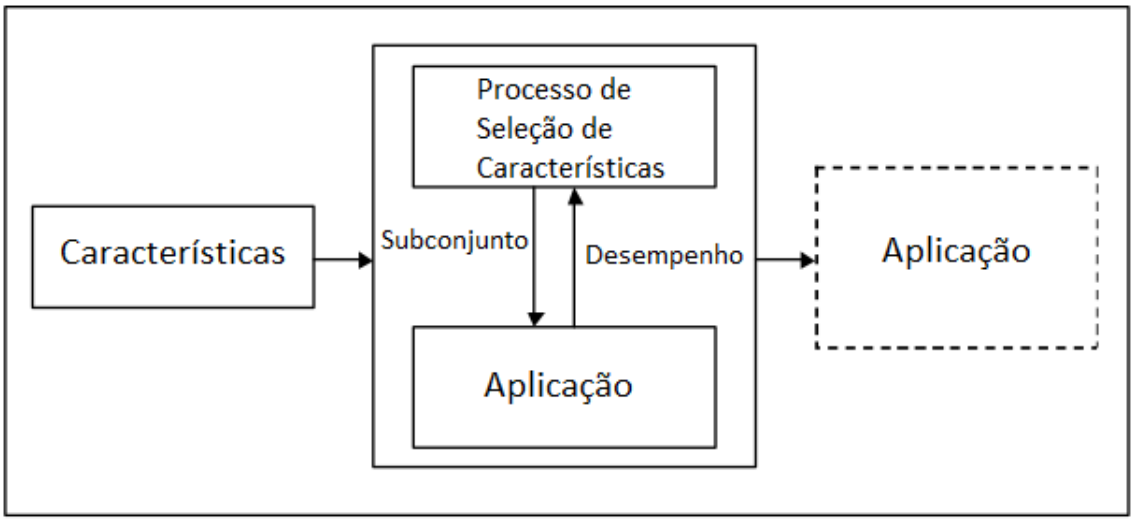

Figura 3.2: Método *Wrapper* de Seleção, adaptado de Puma-Villanueva, Santos, Von Zuben (2006)

Na Figura 3.2 é possível acompanhar como funciona um método *wrapper*. As características passam pelo processo de seleção, condicionada pela avaliação de desempenho em relação a aplicação que será utilizada. Em caso de desempenho insatisfatório, o subconjunto é ajustado até que seja considerado o melhor possível para a aplicação em questão.

3.2.2 Filtros

Métodos baseados na abordagem de filtros funcionam com características gerais de treinamento dos dados para seleção das características, independentemente do classificador e do desempenho gerado por esta seleção. Este tipo de abordagem realiza a seleção dos dados e redução do volume de informações por meio de técnicas mais baratas em relação aos *wrappers* (NZIGA, 2011). Desempenha a seleção por meio da avaliação de contribuição de cada característica individual para compor o subconjunto desejado. As características que não alcançam o valor de avaliação esperado são desclassificadas no processo. Portanto, apenas o valor de cada característica importa para a seleção, o qual depende do tipo de seleção de filtro adotado, não havendo preocupação com o desempenho gerado por este subconjunto.

Abordagens de filtro utilizam-se da linearidade do classificador como uma hipótese prévia, onde índices de correlação podem ser referenciados como principais representantes desta classe (PUMA-VILLANUEVA; SANTOS; VON ZUBEN, 2006). De acordo com Puma-Villanueva, Santos e Von Zuben (2006), em relação a *wrappers* e filtros, é possível avaliar que filtros são computacionalmente menos custosos que *wrappers* mas requisitam uma disponibilidade maior

de amostras para o conjunto de treinamento. Filtros consideram características estatísticas do conjunto de dados de forma direta, sem envolvimento com o desempenho do classificador como métrica. Devido a eficiência computacional, o método de filtro é geralmente utilizado para selecionar características de conjunto de dados com grandes dimensões, como em sistemas de detecção de intrusão.

Assim, a seleção de características para detecção de intrusão busca diminuir o volume de dados para execução da detecção mais rápida e precisa, sem a necessidade de analisar o conjunto integral e adicionalmente remover dados irrelevantes que podem provocar ruídos, interferindo na qualidade de detecção. A minimização do volume por meio de métodos *wrapper* é demorada e dispendiosa, diferentemente, métodos baseados em filtros são mais práticos e simples, consumindo menos recursos e selecionando as características apropriadas através de uma hipótese prévia bem definida (PUMA-VILLANUEVA; SANTOS; VON ZUBEN, 2006).

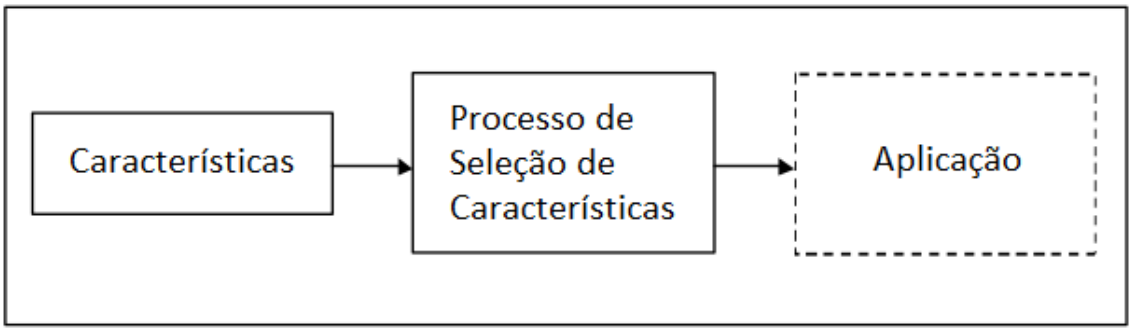

Figura 3.3: Método Filtro de Seleção, adaptado de Puma-Villanueva, Santos e Von Zuben (2006)

A Figura 3.3 mostra como métodos de filtros desempenham sua função. A seleção de características é executada de forma mais simples em relação aos *wrappers*, onde cada característica é selecionada de acordo com sua contribuição para o conjunto, sem preocupar-se com o desempenho do subconjunto selecionado. O processo de seleção neste caso, mostra-se semelhante a Figura 3.1 que apresenta o processo de seleção de forma abstrata.

3.3 Abordagens Para Seleção de Características

Nesta seção são apresentados trabalhos encontrados na literatura, que apresentam soluções para o problema de seleção de características para detecção de intrusão. As abordagens, assim como suas características e principais questões, são descritas com o objetivo de entender como funcionam os métodos envolvidos na seleção de características.

Existem diferentes abordagens para seleção de variáveis para IDS, conforme segue. Em

Cabrera et al. (2002) foi apresentada uma abordagem para detecção de ataques do tipo *Distributed Denial of Service* (DDoS). Esta abordagem baseia-se na definição manual de uma variável que serve como pivô para seleção das demais. Esta variável pivô é comparada com as outras disponíveis na forma de pares utilizando o método de causalidade de *Granger*. O método de causalidade Granger foi utilizado para avaliar a relação de causalidade entre a variável pivô e as demais, propondo uma detecção proativa de intrusões, ou seja, o objetivo é detectar o ataque antes que ele danifique o sistema alvo. A classificação do selecionador neste caso é de filtro.

Em Suebsing e Hiransakolwong (2009) é proposto um método de seleção de variáveis, baseado no conceito de filtros. O método utiliza cálculos de similaridade como critério de seleção. Os cálculos utilizados são o distância euclidiana e a medida de similaridade cosseno. A seleção com o cálculo da distância euclidiana é utilizada para detecção por assinatura e a medida de similaridade cosseno é utilizada para detecção de ataques desconhecidos.

Em Nguyen, Franke e Petrovic (2010) é proposto a utilização de um dos métodos mais conhecidos de seleção, o *Correlation Feature Selection* (CFS). Este método, do tipo filtro, considera a correlação entre as características e as classes de ataques, assim como a inter-correlação entre variáveis. A inter-correlação representa a correlação entre um conjunto de variáveis, sem a utilização de uma variável externa ou dependente. O algoritmo trabalha com a hipótese de que boas variáveis possuem correlação com as classes de ataques, e não possuem correlação com outras variáveis. Quando o número de informações para selecionar é pequeno, o método utiliza força bruta para escolha. Caso contrário, estratégias como heurística e busca aleatória são empregadas.

A proposta apresentada em Mechtri, Djemili Tolba e Ghoualmi (2010) realiza a seleção de variáveis para IDS com base na análise dos componentes principais (*Principal Component Analysis* - PCA). O método PCA é um dos mais utilizados para redução da dimensão de grandes volumes de dados.

O estudo apresentado por Shankarapani et al. (2010), aborda um método para classificação de código maliciosos que pode levar a ataques automatizados e intrusões por meio de *kernel* de uma máquina. Para identificação das características de ataques, utiliza métricas de similaridade (SHANKARAPANI et al., 2010) como similaridade cosseno, medida estendida de Jaccard e correlação de Pearson. As características são analisadas por meio destas métricas, uma vez que são medidas populares para medir a similaridade de sequências. Os autores utilizam a média destas medidas para calcular o grau de similaridades entre diferentes ameaças. Sobretudo, o mé-

todo de Correlação de Pearson chama a atenção por ser capaz de analisar a correlação cruzada entre séries temporais. Desta forma, é uma medida interessante para utilização de um avaliador para seleção de características.

Tabela 3.1: Características de Trabalhos Relacionados

Trabalho	Operação	Tipo de Seleção	Foco em Redução	Medidas de Similaridade	Seleção Dinâmica
Cabrera et al.	Semiautomática	Filtro	Não	Sim	Não
Suebsing e Hirasokolwong	Automática	Filtro	Não	Sim	Não
Nguyen, Franke e Petrovic	Automática	Filtro	Sim	Sim	Não
Mechtri, Djemili Tolba e Ghoualmi	Automática	Filtro	Sim	Não	Não
Shankarapani et al.	Semiautomática	Filtro	Não	Sim	Não

A Tabela 3.1 sumariza os trabalhos relacionados comentados, classificando-os de acordo com o tipo de operação, tipo de seleção, foco em redução de conjuntos, medidas de similaridade e seleção dinâmica. O tipo de operação refere-se à capacidade do selecionador trabalhar de forma automática (sem intervenção humana), semi-automática (com intervenção humana) e manual (seleção totalmente humana). O tipo de seleção está alinhado com a classificação apresentada na Seção 3.2, podendo ser do tipo filtro ou *wrapper*. O foco em redução aponta se o trabalho está concentrado apenas em selecionar para encontrar um subconjunto menor, mas com todas as características do conjunto original. O parâmetro medidas de similaridade aponta a utilização deste tipo de métrica de avaliação ao selecionar dados. A seleção dinâmica classifica de acordo com a capacidade do selecionador escolher diferentes tipos de conjuntos de variáveis de acordo com variações do conjunto analisado, adaptando-se à possíveis mudanças de comportamento.

Ao analisar a Tabela 3.1 é possível avaliar que em termos de tipo de operação, os trabalhos apresentam-se como automáticos e semiautomáticos. A abordagem automática é importante por não ter intervenção humana, entretanto o tipo semiautomático permite que a escolha dos dados seja assistido, auxiliando o processo de seleção. Já quanto ao tipo de seleção, filtros são predominantes na seleção de dados que serão usados em IDS, por sua característica rápida como discutido na Seção 3.2.2. O foco em redução é importante quando o objetivo é dimi-

nuir o conjunto em um subconjunto equivalente, não sendo adequado à situações onde busca-se encontrar comportamentos implícitos no conjunto original. Em relação às medidas de similaridade, pode-se avaliar que este tipo de métrica para seleção torna-se um grande aliado para encontrar comportamento entre variáveis. Quanto a seleção dinâmica, os trabalhos apontados não apresentam este comportamento, de escolha de novos subconjuntos na medida que os dados são analisados, não adaptando-se à mudanças de comportamento entre variáveis.

No Capítulo 4 é apresentado o SDCorr, um método de seleção de variáveis de rede através de comportamentos de correlação entre os dados, selecionando de forma dinâmica novos subconjuntos de acordo com o processo de análise dos dados disponíveis. Esta característica permite encontrar novas variáveis a cada análise, auxiliando o processo de detecção.

3.4 Conclusões Parciais

A seleção de características é definida como uma fase de pré-processamento dos dados de um IDS. Este processo tem o objetivo de selecionar as características mais apropriadas para o processo de detecção, diminuindo o volume de dados e removendo ruídos que prejudicam a qualidade da taxa de acertos e o tempo de processamento.

Dentre os tipos de abordagens existentes para seleção, os filtros destacam-se pela capacidade de selecionar informações rapidamente, alinhando-se ao projeto de IDS, onde o tempo de resposta é importante. Técnicas de seleção baseadas em filtros tem como principais representantes índices de correlação. Para implementação, medidas de similaridade são estratégias de fácil implementação onde é possível classificar as informações de acordo com seu grau de correlação.

4 SELEÇÃO DINÂMICA POR CORRELAÇÃO

Este capítulo apresenta o SDCorr (Seleção Dinâmica por Correlação), uma nova abordagem para seleção de características, bem como questões de implementação da abordagem. A Seção 4.1 apresenta a abordagem proposta, seus conceitos e definições; a Seção 4.2 mostra os aspectos relacionados com a implementação do selecionador proposto; e a Seção 4.3 apresenta o método de detecção que utiliza a abordagem e que permite avaliar o selecionador.

4.1 Seleção de Variáveis pelo Método de Correlação de Pearson

Técnicas de seleção de variáveis devem empregar critérios de escolha que desempenhem papel de decisão na seleção das variáveis adequadas. Deste modo, no caso de seleção de variáveis para sistemas de detecção de intrusão, o subconjunto selecionado deve estar alinhado aos interesses do método de detecção empregado. De acordo com Thottan e Ji (2003), em cenários onde intrusões geram variações abruptas nas variáveis de rede, tal como em ataques de negação de serviço, para identificar variáveis envolvidas numa intrusão é importante priorizar a correlação entre variáveis do tráfego de rede.

Duas abordagens podem ser utilizadas para análise da correlação: analisá-la no próprio detector de intrusões, junto ao algoritmo de análise de mudança abrupta; ou analisá-la numa fase de seleção de variáveis a serem direcionadas ao detector. Além disto, a análise da correlação pode ser realizada *off-line*, de maneira manual ou semi-automática (com intervenção humana) e antes da execução do algoritmo de detecção de intrusão, ou *on-line*, de maneira automática (sem intervenção humana) e dinâmica (reavaliando de tempos em tempos o conjunto de variáveis selecionadas).

A abordagem de seleção proposta neste trabalho explora a seleção semi-automática e dinâmica de variáveis, buscando correlacionar variáveis que foram atingidas por intrusões e que podem auxiliar na detecção de violações de segurança. É fato que, se bem executada, a seleção de variáveis permite melhorar a taxa de detecção e diminuir o tempo de processamento, pois variáveis desnecessárias à detecção geram ruídos que interferem na qualidade e no desempenho do algoritmo de detecção.

Neste sentido, é um mecanismo proposto para seleção de características de rede é baseado na metodologia utilizada por Cabrera et al. (2002), a qual utiliza uma variável pivô para escolha das demais variáveis de interesse, e no método de correlação de Pearson (NETO et al.,

2011), um método rápido e eficiente para avaliar a correlação entre variáveis. Embora similar, o mecanismo difere da abordagem de Cabrera et al. (2002) em dois aspectos. Em Cabrera et al. (2002), a variável pivô, que serve para escolha das demais, é definida manualmente de acordo com o ataque que espera-se detectar e as demais variáveis são selecionadas a partir de causalidades encontradas entre a variável pivô e as demais disponíveis, sendo esta última parte automática. Porém, mesmo sendo classificado como semi-automático, o método depende da intervenção humana e realiza a seleção da variável pivô de maneira manual. Adicionalmente, o método de avaliação é diferente, pois em Cabrera et al. (2002) o objetivo foi encontrar relações de causalidade através do teste de Granger (GRANGER, 1969) e não uma avaliação de similaridade por correlação estatística, obtida através do teste de Pearson.

O método de correlação de Pearson é utilizado como avaliador no processo de seleção, identificando variáveis que possuem correlação na busca por subconjuntos que possuam variações relacionadas com intrusões. O método de seleção é classificado como filtro de acordo com o que foi discutido na Seção 3.2. Este tipo de selecionador é barato do ponto de vista computacional e promove uma seleção rápida e apropriada para ser utilizada em conjunto com sistemas de detecção de intrusão. A seção 4.1.1 apresenta em detalhes este método.

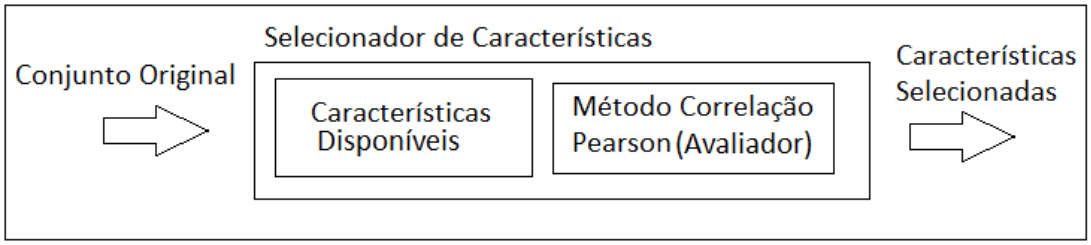

Figura 4.1: Arquitetura do Mecanismo de Seleção SDCorr

A Figura 4.1 apresenta a arquitetura modular do SDCorr. O conjunto original de dados é inserido no selecionador de características, que avalia o conjunto de características disponíveis e realiza a seleção semi-automática da variável pivô. O método de Correlação de Pearson, que faz papel de Avaliador, é então aplicado e resulta em um subconjunto de dados correlacionados. Salienta-se que o processo de seleção de características é composto pela metodologia utilizada para seleção (uso de variável pivo + método de correlação) e o avaliador utilizado neste processo (método de correlação de Pearson). A seguir, estes assuntos serão abordados detalhadamente.

4.1.1 Correlação de Pearson

A correlação de Pearson (NETO et al., 2011) (SHANKARAPANI et al., 2010) (KALOUSIS; PRADOS; HILARIO, 2005) é amplamente utilizada em análise estatística, reconhecimento de padrões e processamento de imagens (XIONG et al., 2004) (NAGARAJAN; UPRETI, 2006) e serve para definir se há existência de correlação linear entre dois vetores de informações. Logo, se aplicado a descritores de tráfego, o método adotado permite identificar a força de relacionamentos lineares entre variáveis descritivas do tráfego de rede. Do ponto de vista estatístico, a medida de correlação linear é um indicador de dependência entre duas variáveis aleatórias. Deste modo, esta medida pode ser usada para definir dependências (medidas por variabilidade similar) entre variáveis de um conjunto de entrada dos IDS.

O cálculo de correlação de Pearson é dado pela Equação 4.1, onde ω e ω' são os vetores de variáveis, ω_i e ω'_i são os valores dos vetores na posição i, e μ_ω e $\mu_{\omega'}$ representam médias aritméticas. O valor resultante, C_ω, pode ter valores entre -1 e 1, onde o valor 1 representa a correlação máxima positiva e o valor 0 que não há correlação. O valor -1 representa que são negativamente correlacionadas.

$$C_w(\omega, \omega') = \frac{\sum_i (\omega_i - \mu_\omega)(\omega'_i - \mu_{\omega'})}{\sqrt{\sum_i (\omega_i - \mu_\omega)^2 \sum_i (\omega_i{}' - \mu_{\omega'})^2}} \tag{4.1}$$

A partir das duas séries submetidas ao teste de correlação, um intervalo de confiança é criado e o resultado entre -1 e 1 é mapeado, com nível de 95%. O coeficiente de Correlação de Pearson é então utilizado como medida de avaliação para seleção das características que possuam correlação com a característica informada via variável pivô. O objetivo é identificar aquelas variáveis que são afetadas por variabilidades na variável pivô, conforme detalhado na próxima seção.

4.1.2 Metodologia de Seleção SDCorr

Quando aplicada na seleção de variáveis, a análise de correlação necessita de uma variável de referência, aqui chamada de *pivô*. A variável pivô deve ser informada com base no conhecimento do usuário, pois deve ser representativa de ataques que tendem a apresentar variações abruptas. Deste modo, a metodologia de seleção proposta, chamada SDCorr, realiza a identificação da variável pivô de forma semi-automática. O método de correlação de *Pearson* é então utilizado como avaliador no processo de seleção, verificando a existência de correlações entre a variável pivô e as demais.

Com este método as variáveis potencialmente afetadas por uma intrusão serão selecionadas, pois permite que a partir da variável pivô, as demais sejam escolhidas. Salienta-se que o processo de seleção visa descobrir as variáveis que possuem correlação com a variável pivô informada (variável potencialmente envolvida nos ataques).

A metodologia de seleção do SDCorr pode ser definida através dos seguintes passos:

1. Através de um algoritmo rápido de análise de variação abrupta, indica-se ao usuário as variáveis que mais apresentam variabilidade para um dado conjunto de dados de entrada;
2. Usuário seleciona variável pivô dentre as indicadas;
3. Avaliação da correlação das variáveis de entrada (características) com a variável pivô, em pares, de acordo com o método de Pearson;
4. Se houver correlação entre as duas variáveis, seleciona a característica;
5. Se não houver correlação entre as duas variáveis, descarta a característica.

É importante salientar que apenas uma variável é utilizada como pivô para ser testada com as demais variáveis candidatas a classificação pelo processo de seleção. A utilização de apenas uma variável é adotada, pois se mais variáveis fossem consideradas como pivôs, a complexidade da metodologia cresceria (EL-KHATIB, 2010) sendo necessário um número maior de avaliações, o que pode degradar o desempenho do método de seleção de características.

A existência de constante mutação no tráfego de rede e nos padrões de assinaturas (HENKE et al., 2011) indicam que é importante evitar métodos estáticos e é necessário dar atenção a interpretação dos padrões de rede. Métodos dinâmicos podem melhorar a taxa de detecção e diminuir o número de falsos positivos (LAKHINA; CROVELLA; DIOT, 2005). Porém, as abordagens de seleção de características costumam realizar a seleção apenas uma vez, não acompanhando as mudanças nos comportamentos entre variáveis. Com a seleção realizada apenas uma vez, aplicação de método estático, a análise de correlação pode não capturar correlacionamentos temporários. Além disto, a detecção pode ser prejudicada pela seleção equivocada de dados que interferem na forma de ruídos, alterando a precisão do detector (HOON; SHIN; CHUNG, 2006).

Neste trabalho a redução do conjunto de dados é desempenhada através da análise de janelas de dados (conceito de janela deslizante) que permite que o teste de correlação fique alinhado ao

comportamento do tráfego, dado que a troca de janela possibilita a adaptação a novos comportamentos do tráfego. A reavaliação automática sob uma nova janela resulta na avaliação dinâmica dos dados.

As mudanças comportamentais do tráfego de rede são acompanhadas a cada janela analisada e apenas as variáveis que possuírem correlação com a variável pivô serão escolhidas, resultando conjuntos diferentes a cada análise. A seleção dinâmica é assim um critério importante para acompanhar a mudança de comportamento entre variáveis do conjunto de entrada. Em outras palavras, considera que duas variáveis podem ser correlacionadas em um determinado período e em outro não.

A Figura 4.2 exemplifica a questão de correlação em diferentes períodos de tempo para duas variáveis (1 e 2). Ao considerar uma janela com todas as amostras de cada variável (amostras do período t e do período $t+1$), ao aplicar o teste de correlação de Pearson tem-se como resultado o valor 0.6123724, que indica uma correlação positiva. Por outro lado, ao considerar as mesmas variáveis 1 e 2, mas em duas janelas de dados (uma para o período t e outra para o período $t+1$), a correlação no período t é 1, ou seja, são 100% correlacionadas positivamente, e no período $t+1$ é zero, indicando que não existe correlação entre estas variáveis.

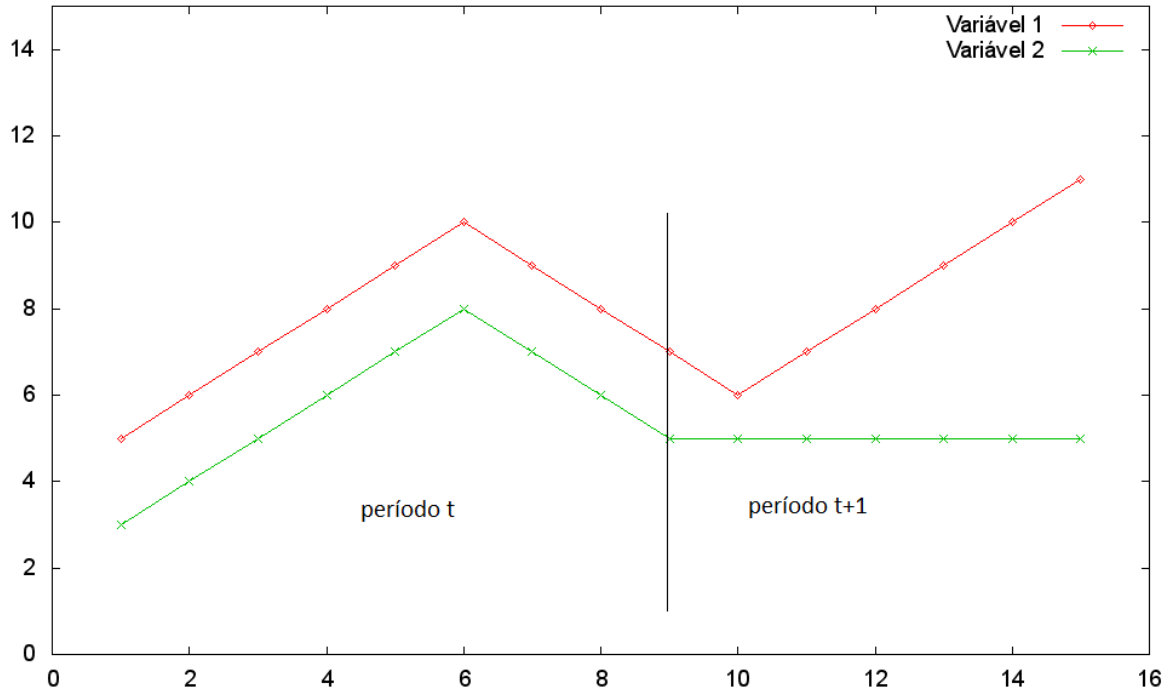

Figura 4.2: Correlações entre Variáveis em Períodos Diferentes

O Algoritmo 1 especifica o funcionamento do SDCorr, o método de seleção de variáveis proposto. Como parâmetro de entrada devem ser informadas as variáveis descritoras de rede.

```
   Entrada: Variável Pivô, Variáveis descritivas de rede
   Saída: Subconjunto selecionado
1  enquanto Existir Novas Amostras faça
2  |  Adiciona Nova Amostra na Janela Deslizante;
3  |  Cada elemento da Janela forma par com a variável pivô;
4  |  Os pares são submetidos ao teste de correlação de Pearson;
5  |  Resultado do teste de correlação é mapeado para valores booleanos;
6  |  se Se avaliação do Par é Verdadeiro então
7  |  |  Seleciona Variável;
8  |  senão
9  |  |  Descarta Variável;
10 |  fim
11 fim
```

Algoritmo 1: Algoritmo do Método de Seleção Proposto, SDCorr

A saída é o subconjunto de variáveis selecionadas. Para cada janela de interesse (Linha 2) é realizado a avaliação pelo método de correlação de *Pearson* (Linha 4), sendo as variáveis selecionadas ou descartadas (Linhas 6 a 10). O resultado é o subconjunto de variáveis que foram selecionadas. Como a seleção é feita em cada janela de interesse, que corresponde a uma janela deslizante no tempo, a seleção permite a identificação das variáveis mais importantes (inclusive a pivô) a cada intervalo de tempo, adaptando-se dinamicamente as variações de interesse no tráfego de rede.

4.2 Aspectos de Implementação do SDCorr

Esta seção apresenta os aspectos referentes a implementação do SDCorr. A implementação foi realizada através da linguagem de programação Java que permite a portabilidade de plataforma. Para manipulação estatística foi utilizado o R (R Development Core Team, 2012) que é uma biblioteca de funções estatísticas que permite a manipulação de séries de dados através de cálculos estatísticos. O R foi escolhido por ser uma biblioteca de software livre e por possuir aplicações auxiliares que permitem seu uso a partir de programas escritos na linguagem Java.

A Figura 4.3 mostra o diagrama de classes projetado para o selecionador de características SDCorr. As classes envolvidas no processo de seleção são: **Pearsons, Rcommand, Rconnection, SelectionDetection, DataSet** e **Config**. As classes **GraphicsD** e **Date** são classes de apoio à experimentação apenas. As classes do processo de seleção são explicadas detalhadamente a seguir.

Rcommand e Rconnection

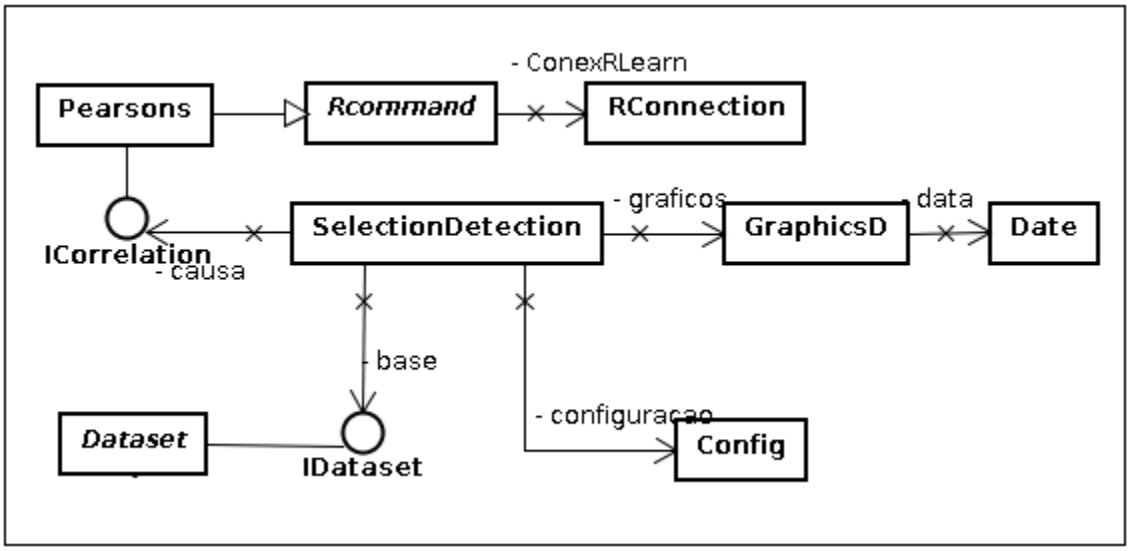

Figura 4.3: Diagrama de Classes do Seletor de Características Proposto

Estas classes são responsáveis pela API (*Application Programming Interface*) referente a biblioteca de manipulação estatística R. O acesso a biblioteca é realizado por uma aplicação auxiliar, o *Rserve* (URBANEK, 2003), que envia, por meio de mensagens TCP/IP, chamadas de rotina de uma aplicação Java ou C/C++ para R. A classe **Rcommand** possui as primitivas necessárias para realização de chamadas para execução de tarefas pelo R e a classe **Rconnection**, pertencente a biblioteca, é responsável pela conexão entre o Java e o R.

Pearsons

Esta classe possui a implementação do método de correlação de *Pearson* através de primitivas do R, por isto estende a classe **Rcommand**. A classe **Pearsons** implementa a interface **ICorrelation**, que especifica como deve ser a assinatura dos métodos que implementam cálculos de correlação. Esta estruturação do código possibilita a implementação de diferentes métodos ou variantes de maneira facilitada. Como parâmetros de entrada, esta classe recebe duas séries temporais a serem testadas, o que neste trabalho corresponde a duas variáveis de interesse, a pivô e uma outra a ser avaliada em relação a pivô. A classe visa confinar o algoritmos que encontram variáveis correlacionadas positivamente. Variáveis que possuem correlações negativas são descartadas. As variáveis que possuem correlação positiva e são mapeadas dentro do intervalo de confiança são selecionadas, ou seja, indica-se que o par de variáveis analisadas possui o tipo de correlação buscado para seleção.

SelectionDetection

Esta classe é a responsável pela coordenação do método. Direciona as informações da base de dados, chama procedimentos para criação de gráficos e manipula a instância do método de

correlação de *Pearson* especificado pela interface **ICorrelation**.

DataSet

Esta classe contém os métodos necessários para utilização da base de dados desejada. Qualquer coletor ou base que seja utilizado deve estender esta classe para que possa ser utilizada pelo processo de seleção.

Config

Esta classe corresponde as informações de configuração para o funcionamento do algoritmo de seleção. Os parâmetros que devem ser informados são: tamanho da janela deslizante, diretório onde se encontram os arquivos referentes a base, identificação da base e identificação do método de correlação. Os arquivos de configuração devem possuir formato texto (txt) e possuir os contadores de pacotes referentes às variáveis. Foi implementado um tipo de base e um método de correlação, sendo identificados pelos seguintes parâmetros Darpa e Pearson, respectivamente. Entretanto é possível implementar diferentes parâmetros para utilização no projeto, desde que estenda a classe **DataSet**, no caso da base, e implemente a interface **ICorrelation**, no caso de métodos de correlação.

4.3 Detector de Intrusões Baseado em Variações de Sinais Abruptos

O método de seleção proposto na seção 4.1, o SDCorr, foi projetado para selecionar dinamicamente variáveis do tráfego de rede, adaptando-se ao comportamento do tráfego. Porém, para auxiliar IDSs, ele necessita estar associado a um IDS que realize detecção de intrusão por meio da análise de anomalias no tráfego de rede. Esta seção apresenta um algoritmo que considera o correlacionamento de variáveis como estratégia para uma detecção mais efetiva, logo, que pode se beneficiar do mecanismo de seleção SDCorr.

Algumas abordagens de detecção de intrusões, tal como em (THOTTAN; JI, 2003) (WU; SHAO, 2005), fazem uso de matrizes de correlação junto ao algoritmo de detecção para poderem reduzir a taxa falsos positivos e melhorar a qualidade da detecção. A hipótese básica destas abordagens é a de que durante um ataque há propagação de variações abruptas em diferentes variáveis de rede correlacionadas. A ideia é que através da análise de correlação via matriz (ordem representada pelo número de variáveis utilizadas) de correlação seja possível diferenciar a variação no tráfego provocada por um ataque e a variação gerada por um comportamento legítimo. Entretanto, em (THOTTAN; JI, 2003) (WU; SHAO, 2005) a seleção de variáveis de interesse é realizada de forma manual e estática.

Para fins de avaliação do SDCorr, considera-se como referência o método para identificação de anomalias proposto em (THOTTAN; JI, 1998) e uma adaptação deste algoritmo proposto por Vogt (2011) , no qual o modelo de séries temporais autoregressivo (AR) foi substituído pelo modelo autoregressivo integrado e de médias móveis (ARIMA) a fim de permitir o tratamento de eventuais comportamentos não estacionários. A estruturação do modelo de detecção de anomalias considerado é apresentado na Figura 4.3.

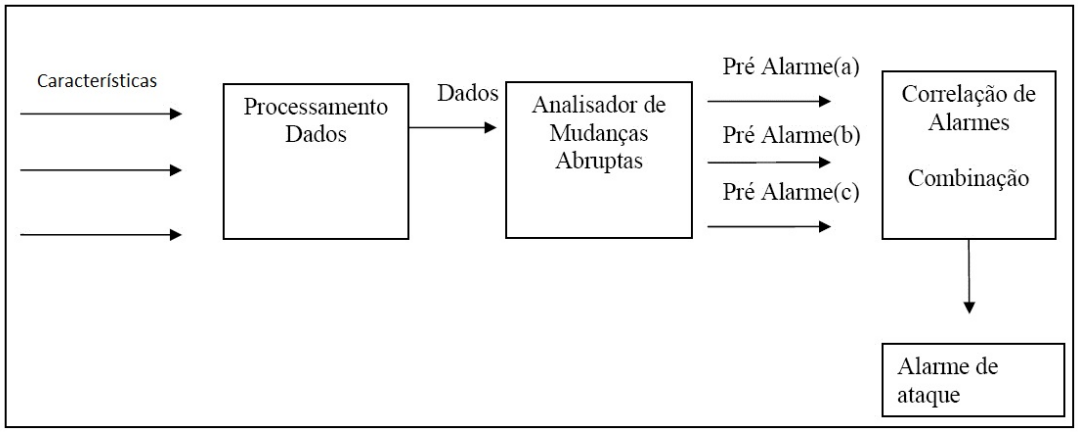

Figura 4.4: Modelo de detecção de anomalias adotado, proposto por Thottan e Ji (1998)

O algoritmo de detecção de anomalias realiza a comparação do comportamento de rede, dado como comportamento normal e o referente a comunicação realizada no intervalo de avaliação. O comportamento normal da rede é dado pela série $L(t) = L(t1), L(t2), L(t3), ...L(tn)$ e o tráfego de rede para avaliação é dado por $S(t) = s(t1), s(t2), s(t3), ...s(tn)$. Mudanças abruptas são detectadas pela comparação da variância de resíduos obtidos a partir de duas janelas adjacentes de dados, representadas por L e S. Estes resíduos são obtidos pela imposição de um modelo autoregressivo nas séries temporais em cada janela.

O algoritmo de detecção analisa a variância dos resíduos em relação ao comportamento normal. A determinação de dados anômalos é identificada com a utilização de testes de hipóteses, baseado nos resultados do *Generalized Likelihood Ratio* (*GLR*) (4.2). Este cálculo é utilizado para representar a saúde da rede com parâmetros entre 0 e 1.

$$-\ln \lambda = \acute{N}_R(\ln \hat{\sigma}_P - \ln \hat{\sigma}_R) + \acute{N}_S(\ln \hat{\sigma}_P - \ln \hat{\sigma}_S) \tag{4.2}$$

As hipóteses são formadas por duas etapas: pela avaliação do distúrbio da variância e pelo teste de força. A primeira etapa do teste de hipóteses é marcada pela avaliação da variância entre os dados referentes ao comportamento normal (L) e o tráfego para avaliação (T). Os

parâmetros que descrevem o modelo de cada uma das séries, dado por um modelo AR. Na segunda etapa, tem-se conhecimento das variações ocorridas no tráfego e analisa-se o nível de perturbação identificado por meio do teste *GLR* e a comparação com o comportamento padrão da rede.

H0	H1
$AR(R) = AR(S)$ $\hat{\sigma}_R = \hat{\sigma}_S = \hat{\sigma}_P$	$AR(R) \neq AR(S)$ $\hat{\sigma}_R \neq \hat{\sigma}_S$

Tabela 4.1: Hipóteses de verificação de perturbação do comportamento (4.3)

H0	H1
$-\ln \lambda \leq h$	$-\ln \lambda > h$

Tabela 4.2: Hipótese que avalia o teste de força (4.4)

Quando identificada uma amostra anômala ela é inserida em um vetor de anormalidade. Este evento poderá ser considerado um ataque se sua repercussão for dada por um período de tempo maior que uma observação, evitando detecções isoladas. Isto poderia gerar falsos positivos de ataques inviabilizando o algoritmo proposto. Esta tarefa é feita em um filtro de duração que controla a avaliação das amostras anômalas.

4.3.1 Correlação de variáveis descritivas de tráfego para detecção

Foram utilizados como base da correlação o valor *GLR* obtido em (4.2). Para cada uma das informações referentes ao tráfego, os valores obtidos estão compreendidos em intervalo de 0 a 1, onde 0 é considerado sem anomalias e 1 é considerado um estado anômalo da rede. Os valores pertencentes ao intervalo descrevem qual o grau de saúde para o fluxo de dados percebido na variável.

A utilização de uma variável por vez permite descrever as anomalias de forma individual para determinação de uma intrusão com melhor qualidade, isto é, com redução de falsos positivos através de correlação de alarmes. Assim, é necessário que seja feita a correlação dos alarmes gerados através das variáveis com o objetivo de encontrar atividades de intrusão. Para correlacionar as variáveis observadas, é utilizada a Equação (4.5), que demonstra a regra geral de correlação dos vetores de anormalidade.

$$f\left(\overrightarrow{\psi}\,(t)\right) = \overrightarrow{\psi}\,(t)\, A\, \overrightarrow{\psi}\,(t) \tag{4.5}$$

A matriz formada pelo operador A é dada por (MxM). Este operador Matriz foi projetado com base em autovetores. A operação da matriz emprega um bloco diagonal da matriz superior com um bloco inferior. Os elementos do bloco superior seguem duas regras: $I = J$ (4.6) e $I \neq J$. Para $I \neq J$ (4.7), é a média do conjunto dos pontos de correlação que possuem cruzamento espacial no vetor de anomalias estimadas no tempo.

$$Aupper\,(i,j) = 1 - \sum_{i \neq j} A\,(i,j) \tag{4.6}$$

$$Aupper\,(i,j) = \frac{1}{T} \sum_{t=1}^{T} \psi_i\,(t)\,\psi_j\,(t) \tag{4.7}$$

O resultado do emprego das regras gera um coeficiente entre 0 e 1, responsável por indicar qual o grau de anomalia que está presente nas variáveis observadas. Para avaliar o resultado obtido aplica-se duas regras de avaliação do resultados. Esta regras são definidas pela Equação (4.8) que avalia se é maior que 0 e a Equação (4.9) que avalia se o valor obtido é menor quer 1.

$$t_a = inf\{t : \int (\overrightarrow{\psi}(t)) \geq \lambda_N\} \tag{4.8}$$

$$\int (\overrightarrow{\psi}(t)) \leq \lambda_M = 1 \tag{4.9}$$

Este resultado mapeado no intervalo entre 0 e 1, representa o resultado final da detecção de uma intrusão, obtido através da correlação de alarmes individuais das variáveis referentes ao tráfego de rede. Quanto mais próximo a 1, maior o grau de variação e consequentemente do grau da intrusão. Diferentemente, quanto mais próximo a 0, menor foi a variação gerada pelo ataque e o grau desta intrusão é considerado baixo.

4.4 Conclusões Parciais

Neste capítulo foram apresentados aspectos relativos a proposta de seleção de características, o SDCorr. A metodologia proposta é desempenhada por um processo semi-automático que se diferencia dos propostos na literatura por realizar análise de correlação de maneira dinâmica, ou seja, avalia a correlação de segmentos das variáveis de interesse (janelas), possibilitando melhor adaptação a variabilidade de características no tráfego de rede. O mecanismo utiliza como avaliador o método de correlação de Pearson, um método simples e eficaz para avaliar a correlação linear entre variáveis.

A partir de uma variável que possua mudança abrupta (potencialmente provocada por uma intrusão), chamada de pivô, as variáveis são selecionadas com base no teste de correlação. Deste modo, as características mais prováveis de estarem envolvidas com a intrusão devem ser escolhidas (processo de seleção de variáveis).

Os aspectos de implementação também foram abordados no capítulo. O projeto do algoritmo e o diagrama de classes referentes ao funcionamento do mecanismo de seleção do SDCorr foram abordados, possibilitando a identificação da simplicidade do método. O detalhamento do detector de intrusões baseado em análise de sinais abruptos, explica o detector escolhido para integração com o SDCorr, sendo considerado o cenário de uso do selecionador nos experimentos do capítulo seguinte.

5 EXPERIMENTOS E RESULTADOS

Neste Capítulo são apresentados os experimentos realizados para testar e avaliar o método de seleção de variáveis SDCorr juntamente com a técnica de detecção de intrusão através de análise de sinais abruptos. Os experimentos utilizam ataques da base Darpa 99 (seção 5.2) e realizam a detecção de sinais abruptos sem e com o método de seleção de variáveis (seções 5.3 e 5.4, respectivamente). A discussão em relação aos resultados é realizada em conjunto com a apresentação dos mesmos. Os detalhes do ambiente considerado para os testes, assim como definições importantes para realização dos experimentos em questão, são apresentados na seção 5.1.

5.1 Ambiente e Definições

Para a avaliação da proposta de seleção SDCorr foi utilizado o detector de intrusão baseado em anomalia proposto em Thottan e Ji (2003) e a base de ataques Darpa 99 (DARPA, 1999). O detector realiza a análise da ocorrência de sinais abruptos em diferentes variáveis de rede no mesmo período de tempo. Alarmes individuais são gerados para cada variável e então combinados por uma matriz de correlação, resultando em alarmes de intrusões. As variáveis da base Darpa 99 utilizadas correspondem a contadores de pacotes referentes aos protocolos de rede tradicionais envolvidos na comunicação. Foram utilizadas 13 variáveis que são: TCP, UDP, IP, ICMP, ARP, TCP-SYN, TCP-ACK, TCP-FIN, TCP-URG, TCP-CWR, TCP-PSH, TCP-ECE e TCP-RST. Os testes foram realizados em um computador *notebook*, *Intel(R) Core(TM) i7* de frequência 2.67Ghz, com 4 GB de memória RAM. Este computador opera sobre o sistema operacional *Fedora Linux*, com a ferramenta estatística *R* instalada.

Neste cenário, foram realizados dois experimentos: um com o detector operando sem o método de seleção de variáveis (utiliza o conjunto completo de 13 variáveis) e outro com o detector operando com o método de seleção SDCorr. Metodologia similar foi adotada em (SHANMUGAM; IDRIS, 2009) e (SUEBSING; HIRANSAKOLWONG, 2009), onde a proposta de seleção é comparada à um caso que não usa seleção (considera o conjunto completo neste caso).

Dentre os ataques existentes na base Darpa 99, foram selecionados 8 contidos na segunda semana (os documentados e que possibilitam avaliar a precisão do detector). Nos testes foram utilizados 4 tamanhos de janelas deslizantes: 32, 64, 128, 256 elementos. O menor tamanho foi definido em função da quantidade mínima de amostras exigida pela modelagem de séries

temporais AR (*Auto-Regressive*) utilizada na detecção por sinais abruptos. O maior tamanho foi definido em função da diminuição da precisão do detector, apontando que janelas maiores não eram necessárias. Para determinação da correlação significativa de duas séries analisadas, utilizou-se intervalos de confiança para apontar quando há correlação não nula, de acordo com o índice de 95% de confiança, que é muito utilizado em intervalos de confiança para inferências estatísticas (GUJARATI, 2004).

5.2 Ataques

Nesta Seção os ataques da base DARPA 99 considerados para os experimentos são explicados de forma a entender suas características em relação a detecção de intrusão e o método de seleção de variáveis. O comportamento em relação ao tráfego de rede é apresentado através de gráficos.

5.2.1 Neptune

O ataque *Neptune* (PILLI; JOSHI; NIYOGI, 2011) (TUPAKULA; VARADHARAJAN; PANDALANENI, 2009), também nomeado *TCP-SYN Flood*, inicia com a realização de várias conexões com a vítima através do envio de pacotes *TCP* com a *flag SYN* (*synchronization*) ativada. Estas conexões não são completadas pelo retorno de pacotes com a *flag ACK* (*acknowledgement*). Como o servidor mantém um *buffer* com todas as conexões iniciadas e não completadas, este *buffer* fica sobrecarregado levando ao sistema ao estado de negação de serviço.

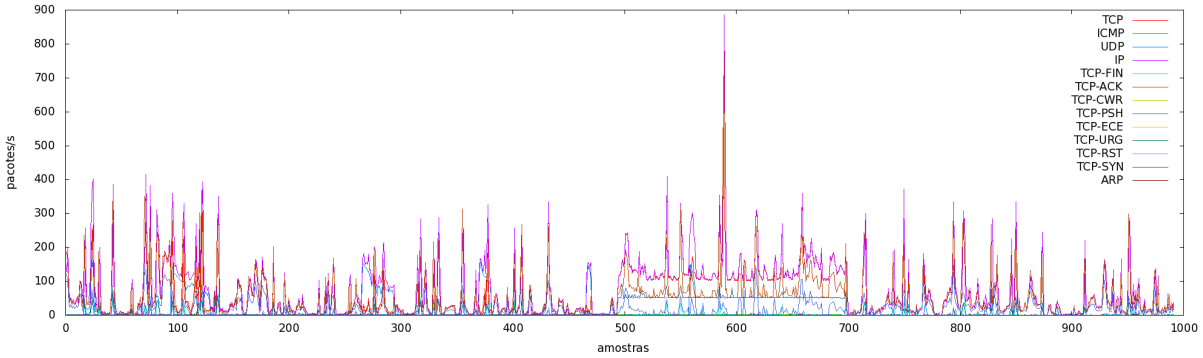

Figura 5.1: Ataque Neptune e as variáveis disponíveis no momento do ataque.

A Figura 5.1 mostra o comportamento das variáveis disponíveis no momento do ataque *Neptune*. Na figura, mesmo com a presença de tantas curvas, é possível perceber as variações e que permitem identificar o ataque. O ataque inicia próximo ao instante 500 e termina pró-

ximo ao instante 700. Analisando esta imagem é possível notar que fora do período do ataque as variáveis possuem comportamento heterogêneo. Diferentemente, entre os instantes 500 e 700 as variáveis apresentam mudanças decorrente da ameaça, mantendo um comportamento de variação possível de notar o *Neptune*.

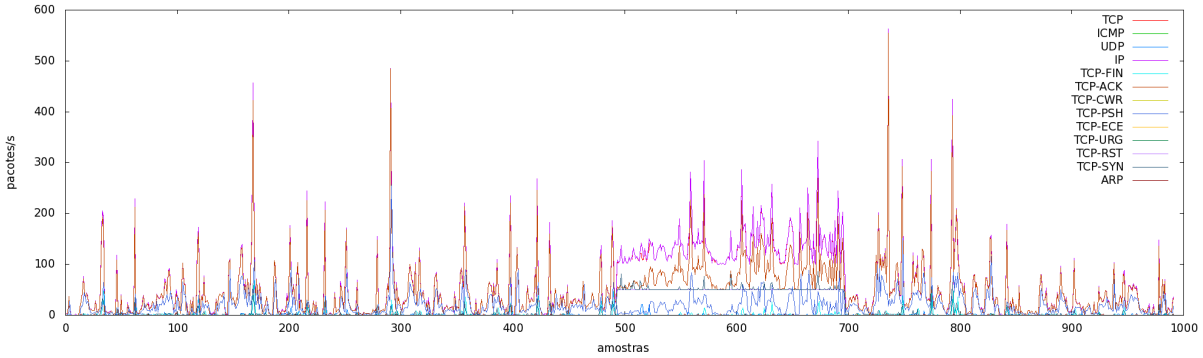

Figura 5.2: Variáveis disponíveis e o momento do ataque Neptune 2

A Figura 5.2 apresenta todas as variáveis no momento da intrusão. Através da observação deste elemento gráfico é possível identificar facilmente o início do ataque que ocorre próximo ao instante 500. As variáveis dentro do período da intrusão, possuem mudanças correlacionadas, visto que o comportamento de diversas variáveis têm praticamente o mesmo comportamento. Embora seja possível verificar alguns picos fora do período da intrusão (500 a 700), no momento do ataque os dados mostram-se parecidos em suas curvas, indicando a presença de correlações entre as variáveis.

As Figuras 5.1 e 5.2 mostram os gráficos das variáveis disponíveis nos recortes que possuem ataques do tipo Neptune. Mesmo com a presença de tantas curvas, é possível perceber as variações que representam os ataques nos dois casos. Embora os ataques sejam execuções distintas presentes na base, como é possível observar comparando as imagens, os ataques iniciam próximo ao instante 500 nas duas imagens.

As Figuras 5.3 e 5.4 apresentam os sinais provocados pelo início do ataque no instante próximo ao 500. Nestas imagens, provenientes do mesmo tipo de ataque (Neptune) mas não da mesma execução é possível verificar comportamentos praticamente iguais. Como este ataque provoca disparos de pacotes do tipo TCP-SYN e estas conexões não são finalizadas, percebe-se um aumento na quantidade de pacotes para esta variável. Antes do início do ataque a variável TCP-SYN possui comportamento variável, diferentemente, quando o ataque é iniciado a quantidade de pacotes por segundo é maior que 50 até o final da intrusão (instante 700).

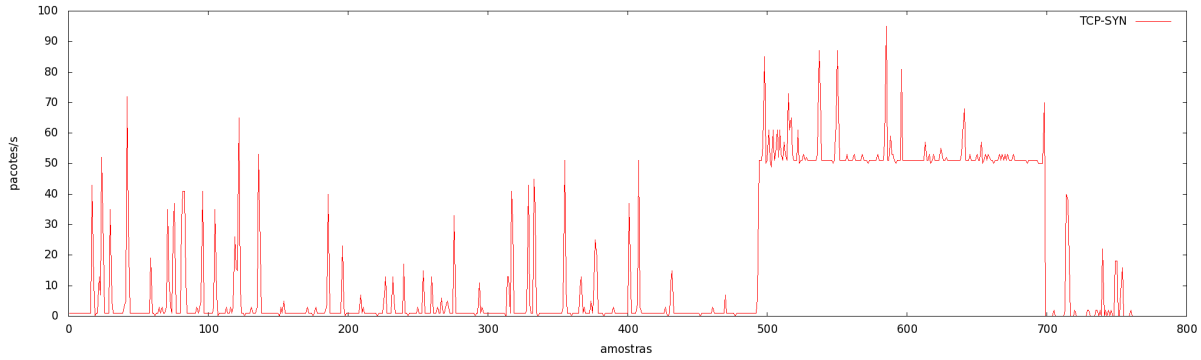

Figura 5.3: Comportamento do ataque Neptune 1

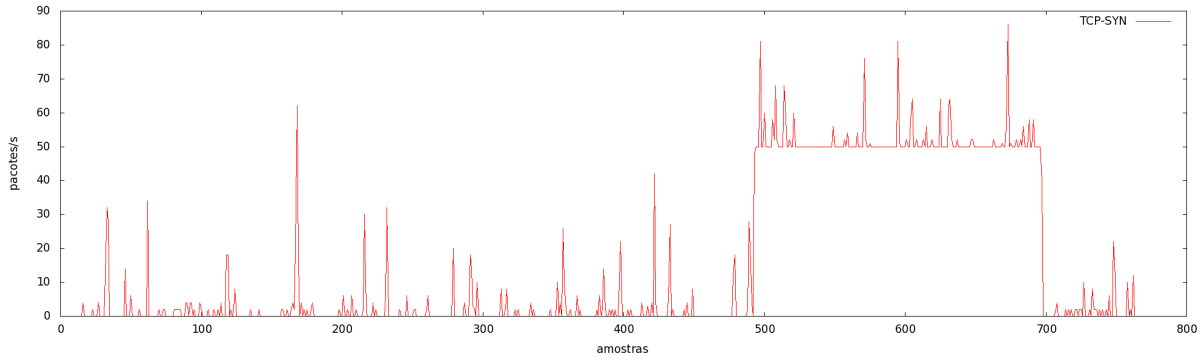

Figura 5.4: Comportamento do ataque Neptune 2

5.2.2 Mailbomb

O ataque *Mailbomb* (OSTASZEWSKI; BOUVRY; SEREDYNSKI, 2008) (DARPA, 1999) é realizado pelo envio múltiplo de mensagens de *e-mail* para um servidor, causando sobrecarga na fila de mensagens, o que causa falha no servidor.

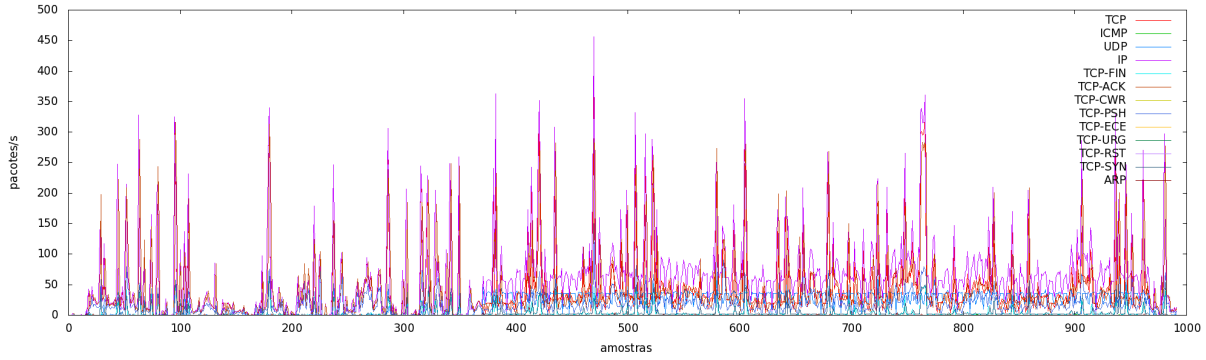

Figura 5.5: Período do ataque Mailbomb e as variáveis disponíveis.

A Figura 5.5 mostra as variáveis disponíveis em recortes que contém o ataque Mailbomb. O gráfico ilustra que quando as curvas de todas as variáveis são colocas em uma mesma imagem, não fica claro o momento que ocorre o ataque. Porém, embora não seja claro, é possível verificar

uma variação considerável próximo ao momento 400, onde, de fato, inicia o ataque Mailbomb.

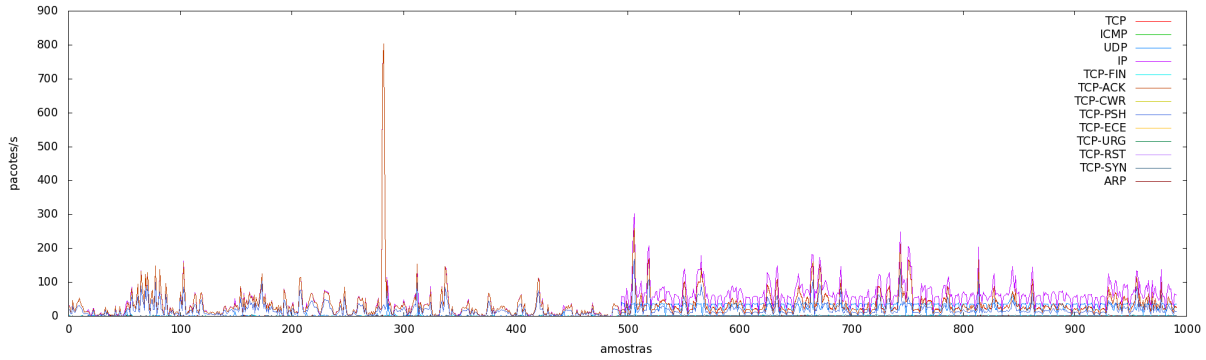

Figura 5.6: Variáveis disponíveis e o momento do ataque Mailbomb 2

A Figura 5.6 apresenta as variáveis envolvidas em uma execução do Mailbomb distinta da anterior. Nesta imagem é possível verificar uma grande variação próximo ao instante 300. Esta variação não está associada ao ataque e pode ser classificada como uma mudança abrupta isolada. Já quando olha-se para o instante 500 percebe-se uma grande aglomeração das curvas de diferentes descritores do tráfego, sendo este associado ao ataque Mailbomb.

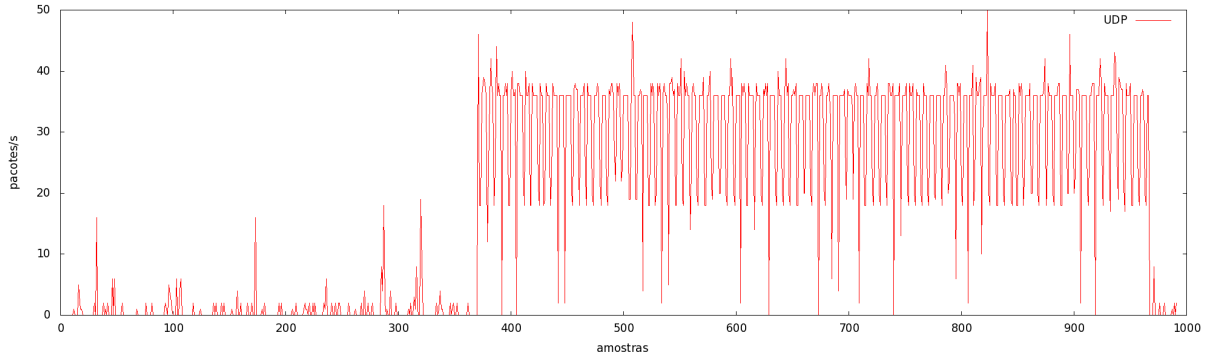

Figura 5.7: Comportamento do ataque Mailbomb 1

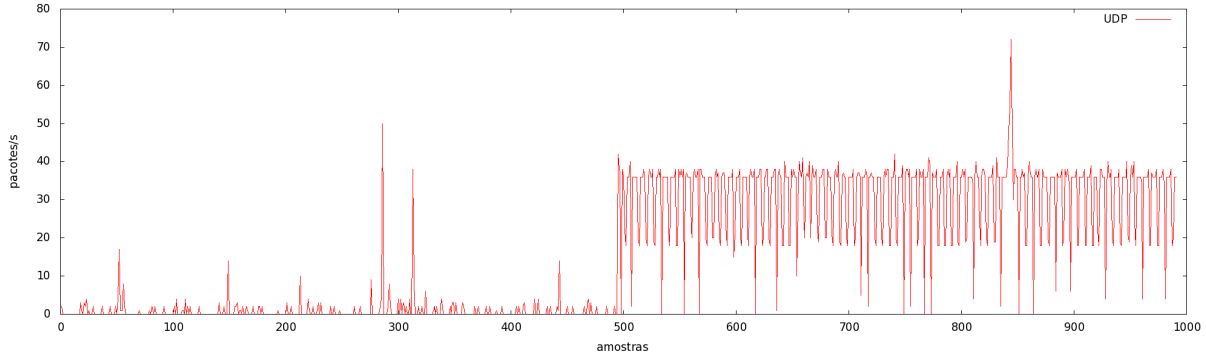

Figura 5.8: Comportamento do ataque Mailbomb 2

Uma análise detalhada da variável UDP, conforme ilustram as Figuras 5.7 e 5.8, torna claro o momento em que iniciam os ataques Mailbomb. Na primeira execução a variação incomum inicia próximo ao instante 400 e, na segunda execução a variação abrupta inicia próximo ao instante 500. Analisando as curvas, percebe-se que antes do momento da intrusão, a quantidade de pacotes por segundo possui variação, sempre voltando à níveis pequenos. A partir dos instantes referenciados como início das intrusões, a quantidade de pacotes por segundo oscila entre 20 e 40, caracterizando este tipo de ataque.

5.2.3 Ping da Morte

O Ping da Morte (KR; INDRA, 2010) (PENG; LECKIE; RAMAMOHANARAO, 2007), também conhecido como *Ping of Death - PoD*, explora vulnerabilidades do protocolo ICMP. O ataque consiste no envio de pacotes *ICMP* com tamanho maior que 65.536 bytes cada, excedendo o tamanho máximo permitido para um pacote *IP*, como resultado ele causa negação de serviço.

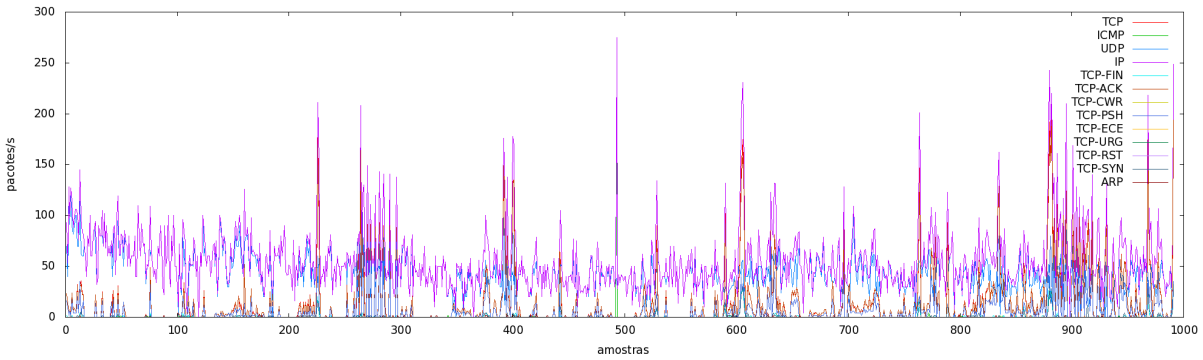

Figura 5.9: O momento do ataque Ping da Morte e as variáveis disponíveis.

A Figura 5.9 ilustra um período de ataque PoD. Embora existam alguns momentos de variação decorrente das diversas variáveis disponíveis colocadas em uma mesma imagem, é possível destacar uma variação mais significativa próximo ao instante 500, que corresponde ao ataque Ping da Morte.

Na Figura 5.10 o momento do ataque fica mais claro, também próximo ao 500 e que corresponde ao Ping da Morte 2. Diferentemente dos outros ataques considerados para os experimentos, é possível perceber que o Ping da Morte não provoca variações abruptas em diversas variáveis, sendo associado à apenas esta variação demonstrada próximo ao instante 500, para os dois casos.

Já através da observação das Figuras 5.11 e 5.12 as variações decorrentes dos ataques do

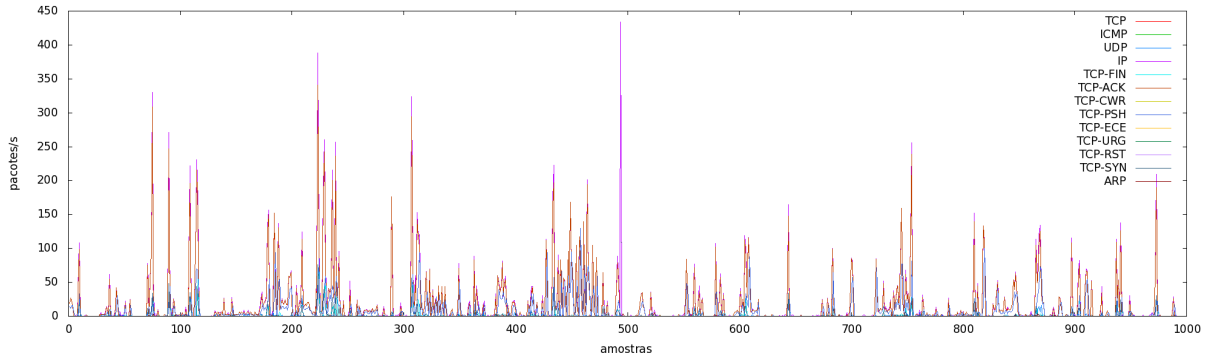

Figura 5.10: Variáveis disponíveis e o momento do ataque Ping da Morte 2

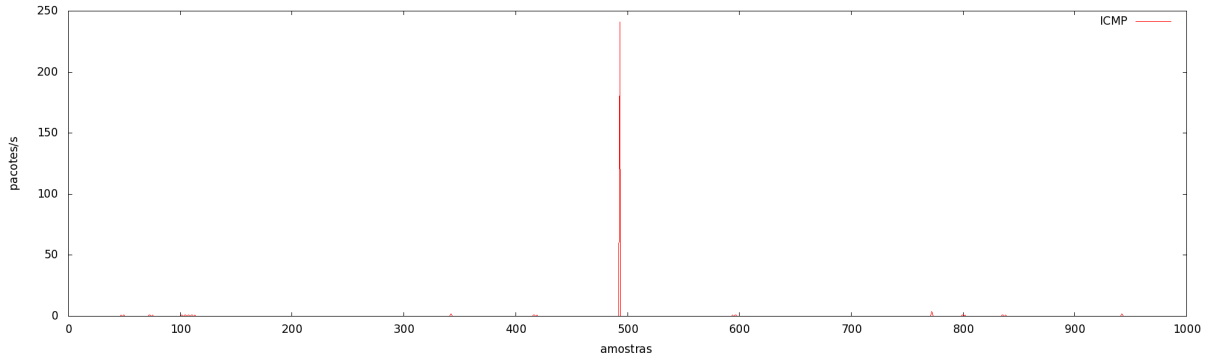

Figura 5.11: Comportamento do ataque Ping da Morte sobre a variável ICMP.

tipo Ping da Morte ficam claras, quando a variável ICMP é colocada separadamente das demais. Este ataque não gera inundação no tráfego, como pode-se verificar, apenas uma variação abrupta é associada ao ataque.

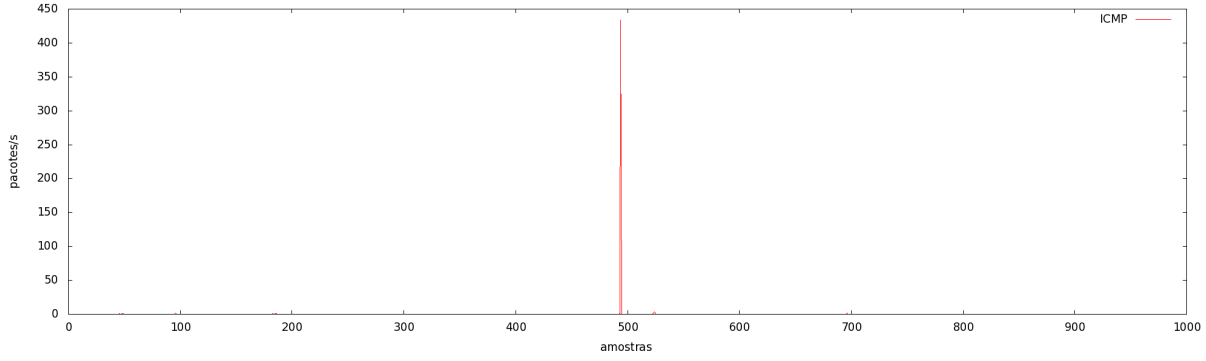

Figura 5.12: Comportamento do ataque ataque Ping da Morte 2

5.2.4 SATAN

O *SATAN (Security Administrator Tool for Analyzing Networks)* (DARPA, 1999) é uma ferramenta de análise de rede que permite testar o ambiente em busca de vulnerabilidades. O

ataque SATAN possui três níveis: leve, moderado, pesado. Estes níveis são referentes ao tipo de análise realizado.

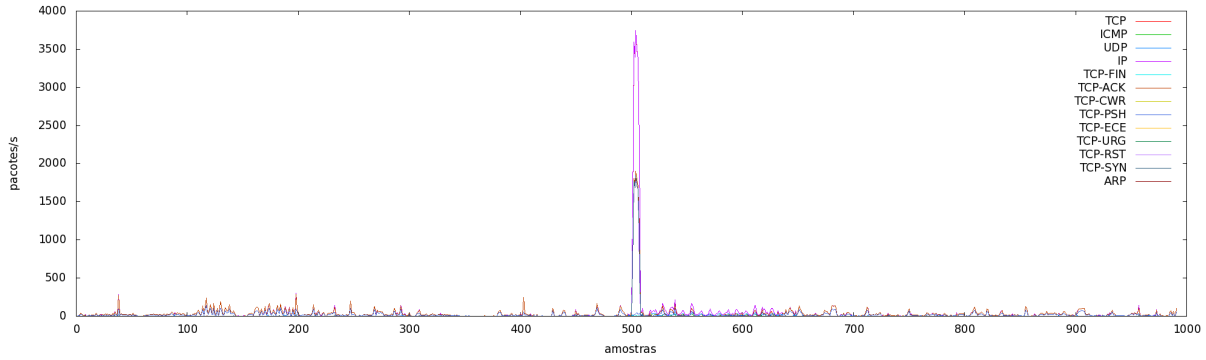

Figura 5.13: Variáveis disponíveis e o momento do ataque Satan 1

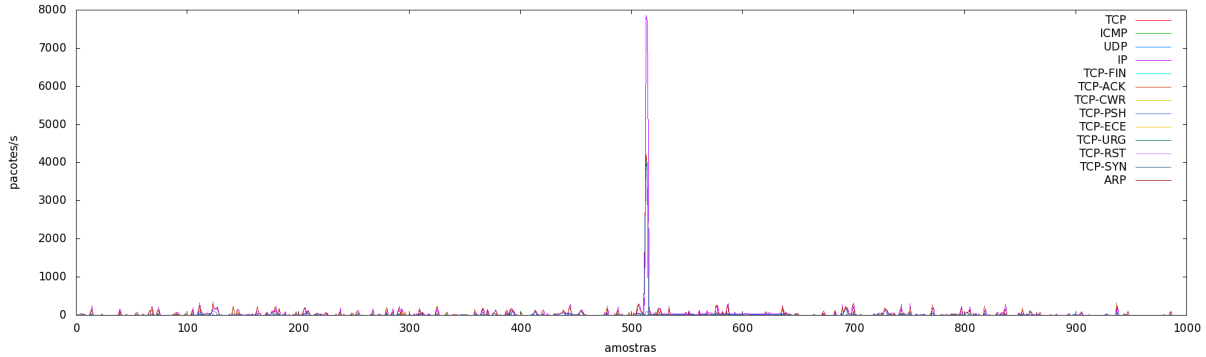

Figura 5.14: Variáveis disponíveis e o momento do ataque Satan 2

Tanto através da observação das Figuras 5.13 e 5.14 que apresentam todas as variáveis como nas Figuras 5.15 e 5.16, é possível definir o momento em que ocorrem os ataques SATAN 1 e 2, próximo ao instante 500 nos dois casos.

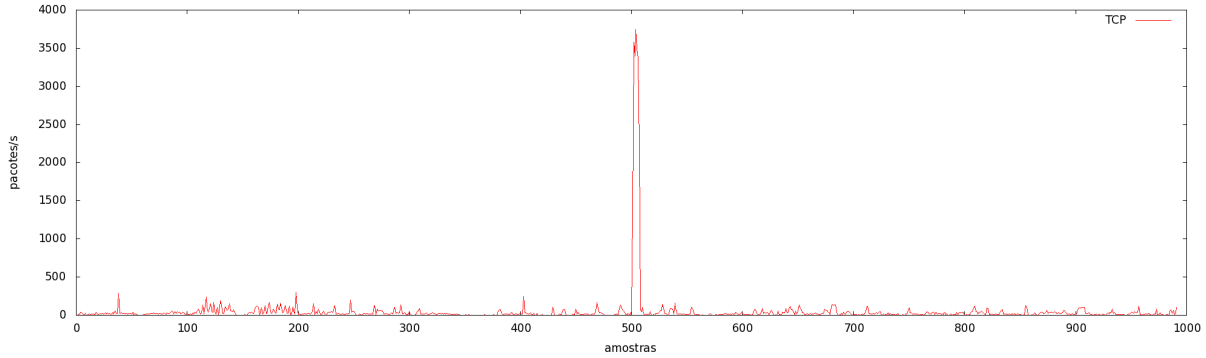

Figura 5.15: Comportamento do ataque Satan 1

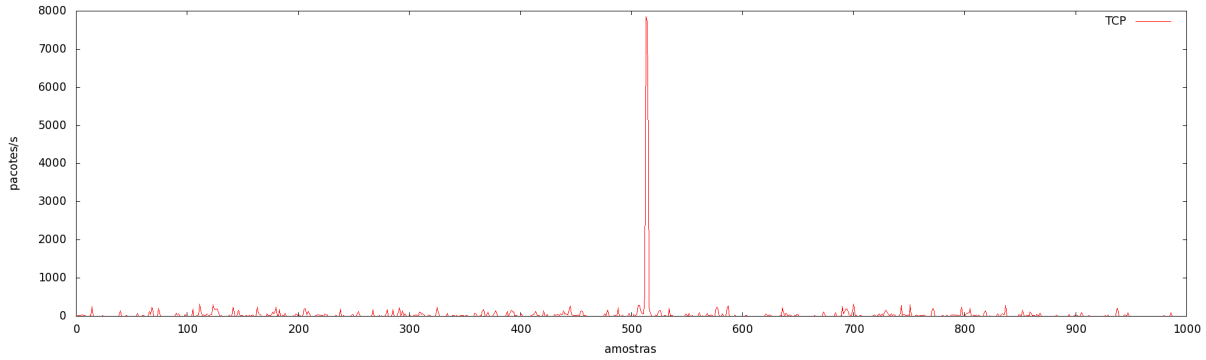

Figura 5.16: Comportamento do ataque Satan 2

5.3 Experimento Sem Seleção de Características

O primeiro experimento foi realizado com o detector de intrusão sem o método de seleção SDCorr. Em outras palavras, o detector recebe como entrada todas as variáveis disponíveis. Todos os tamanhos de janelas de detecção (32, 64, 128, 256 elementos) foram utilizadas e avaliadas.

Considerando todos os 8 ataques executados com os diferentes tamanhos de janelas deslizantes, neste experimento não houve nenhum tipo de alarme, zerando o número de falsos positivos e verdadeiros positivos. Isto decorre do tipo de análise utilizado na detecção das diversas variáveis. Cada uma das variáveis envolvidas são analisadas a fim de descobrir alarmes individuais. Após isto, os alarmes são colocados em uma matriz de correlação que tem como objetivo gerar os alarmes finais.

A quantidade de dados (13 variáveis) envolvidas pode ter diminuído as chances de que a matriz de correlação produzisse resultados positivos. Isto porque as variáveis podem possuir alarmes individuais em instantes diferentes, e mesmo no momento de uma intrusão, as variáveis podem inserir ruídos que provocam o resultado nulo da matriz de correlação.

Como não houve utilização de um método de seleção das variáveis, a taxa de utilização das variáveis envolvidas neste experimento foi de 100% para todas as 13 envolvidas na execução, em todas as janelas deslizantes. A Tabela 5.1 mostra o tempo total de execução na detecção dos 8 ataques para os 4 tamanhos diferentes de janelas no experimento utilizando o conjunto completo de variáveis. Os dados nesta tabela representam médias de trinta execuções e o desvio padrão pode ser acompanhado através da terceira coluna.

Tabela 5.1: Tempos de Execução no experimento Sem Seleção de Variáveis

	Tempo (média em segundos)	Desvio Padrão
Janela 32	184,347	2,253
Janela 64	146,924	1,679
Janela 128	113,486	1,434
Janela 256	94,108	1,862

5.4 Experimento Com Seleção pelo SDCorr

Esta seção apresenta o experimento em que foi utilizado o método SDCorr em conjunto com a técnica de detecção de intrusão através de análise de sinais abruptos.

Neste experimento, os parâmetros utilizados são os mesmos do experimento anterior, ou seja, 4 tamanhos diferentes de janelas e 13 variáveis de rede, mas com utilização do método seleção de variáveis proposto.

A cada movimentação da janela deslizante uma variável pivô é testada com as demais variáveis disponíveis em busca de correlação através do método de *Pearson*. Assim, é feita a seleção, escolhendo aquelas mais apropriadas de acordo com os dados da janela em questão. É importante notar que a seleção é realizada a cada movimentação da janela para que se adapte ao comportamento dinâmico dos dados.

A Tabela 5.2 apresenta a quantidade de verdadeiros positivos, falsos positivos e falsos negativos para os diferentes tamanhos de janelas utilizados neste experimento.

Tabela 5.2: Resultados do experimento com o SDCorr

	Verdadeiros Positivos	Falsos Positivos	Falsos Negativos
Janela 32	5	279	3
Janela 64	4	426	4
Janela 128	0	342	8
Janela 256	0	0	8

Através da análise da Tabela 5.2 é possível notar que o melhor caso, em que 5 dos 8 ataques foram detectados ocorre com janela de tamanho 32. Conforme o tamanho da janela vai crescendo, a precisão vai diminuindo, ou seja, isto decorre da possibilidade do teste de correlação de *Pearson* ser mais preciso quando o tamanho da janela é menor. Quando a janela é maior, a seleção verifica a presença de correlação em toda a extensão da janela, podendo em determinados momentos não serem correlacionadas, o que prejudica a detecção. Desta forma, no caso

das janelas maiores de 32 pode existir resposta de correlação, mas isto não implica que todas as curvas da janela analisada sejam correlacionadas, provocando alarmes nulos na utilização da matriz de correlação por parte do detector de intrusão.

Tabela 5.3: Tempos de execução no experimento Com o SDCorr

	Tempo (segundos)	Desvio Padrão	Economia de Tempo
Janela 32	61,078	2,446	66,86%
Janela 64	51,429	2,328	64,99%
Janela 128	43,204	1,859	61,93%
Janela 256	31,512	2,625	66,51%

A Tabela 5.3 mostra o tempo de execução na detecção dos 8 ataques para os 4 tamanhos diferentes de janelas neste experimento, que utiliza o método de seleção SDCorr. É possível observar que o tempo de execução em relação ao experimento sem seleção foi consideravelmente menor em todos os tamanhos de janela, alcançando até 66,86% de economia no caso de janela de tamanho 32. A menor taxa de economia de tempo obtida foi de 61,93% com janela de tamanho 128. Os dados referem-se à médias de trinta execuções dos testes e o resultado do desvio padrão para cada janela pode ser acompanhado por meio da terceira coluna desta tabela.

Tabela 5.4: Taxa de Seleção das Variáveis Disponíveis

	Janela 32	Janela 64	Janela 128	Janela 256
ICMP	7,66%	14,16%	21,42%	12,5%
TCP-SYN	30,64%	40%	41,07%	20,83%
TCP	67,33%	64,16%	58,92%	54,16%
ARP	3,62%	9,16%	5,35%	4,16%
UDP	12,5%	19,16%	16,07%	16,66%
TCP-PSH	33,06%	43,33%	46,42%	25%
IP	64,51%	65%	66,07%	54,16%
TCP-FIN	32,66%	37,5%	35,71%	20,83%
TCP-URG	0,40%	0,83%	0%	0%
TCP-ECE	0%	0%	0%	0%
TCP-CWR	0%	0%	0%	0%
TCP-RST	3,23%	3,33%	5,35%	4,16%
TCP-ACK	41,53%	48,33%	48,21%	20,83%

A Tabela 5.4 apresenta a taxa de seleção das variáveis disponíveis. A partir das informações apresentadas nesta tabela é possível afirmar que o selecionador SDCorr promoveu economia de recursos que refletiu menor tempo de processamento, devido a escolha de diferentes variáveis a

cada janela analisada. Como as variáveis foram escolhidas a cada movimento da janela, pode-se verificar que as taxas de utilização são distintas para cada variável apresentada. Em comparação ao caso que o conjunto completo foi utilizado, ou seja, utilização de 100% de todas as variáveis, e economia de recursos computacionais fica clara analisando a Tabela 5.4 assim como a Tabela 5.3.

5.4.1 Gráficos dos Resultados para os Ataques

O experimento com seleção pelo SDCorr resultou na identificação de alarmes, diferentemente do que ocorreu com o experimento sem seleção (utilização de todas as variáveis). Os gráficos apresentados e discutidos a seguir são referentes ao teste que obteve melhor resultado, ou seja, com janela de tamanho 32.

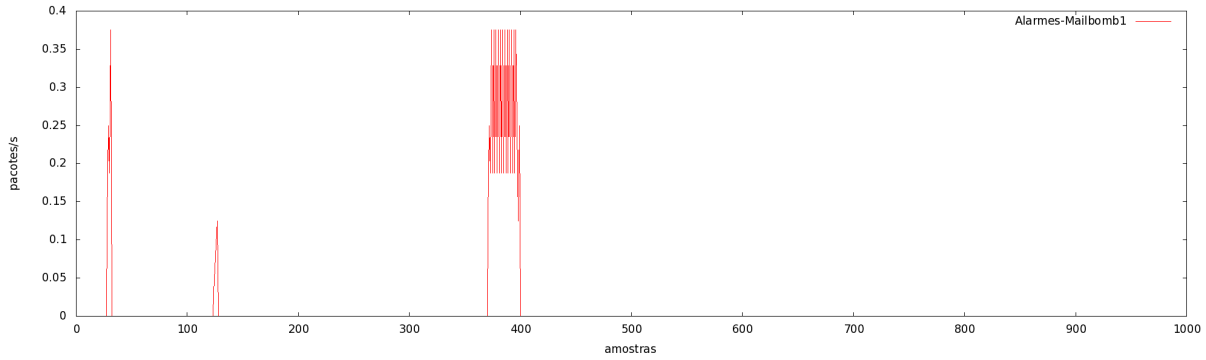

Figura 5.17: Alarmes Gerados na Detecção do Ataque *Mailbomb* 1

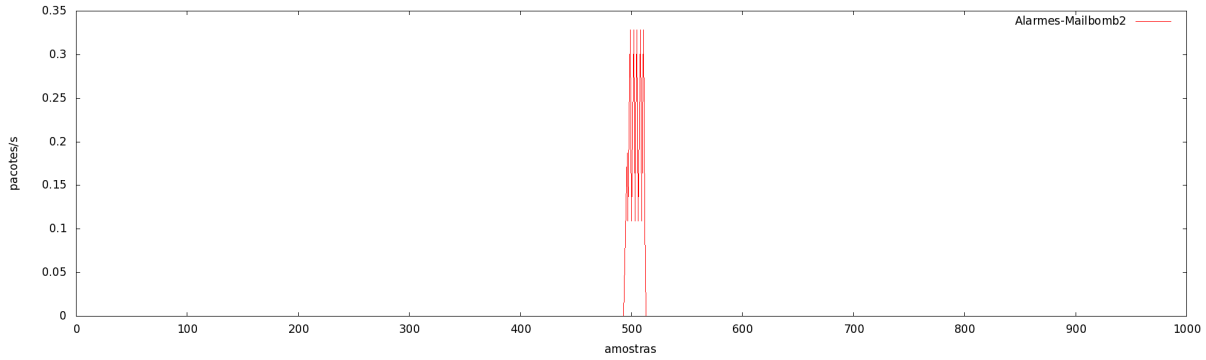

Figura 5.18: Alarmes Gerados na Detecção do Ataque *Mailbomb* 2

As Figuras 5.17 e 5.18 mostram os alarmes referentes à detecção dos ataques Mailbomb 1 e Mailbomb 2. É possível observar que os alarmes gerados próximos ao momento 370 na Figura 5.17 e próximo ao 500 na Figura 5.18 indica o momento em que o detector sinaliza a presença dos ataques. As demais variações percebidas nos gráficos referem-se a falsos positivos.

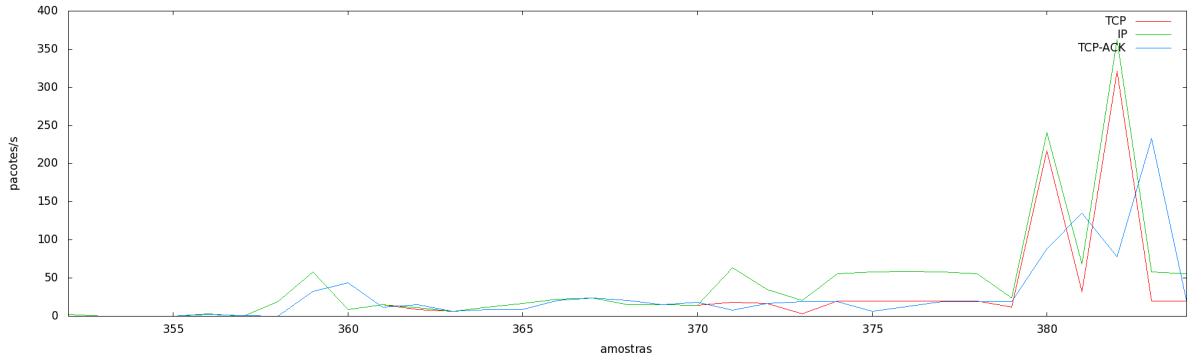

Figura 5.19: Variáveis selecionadas no momento do ataque Mailbomb 1

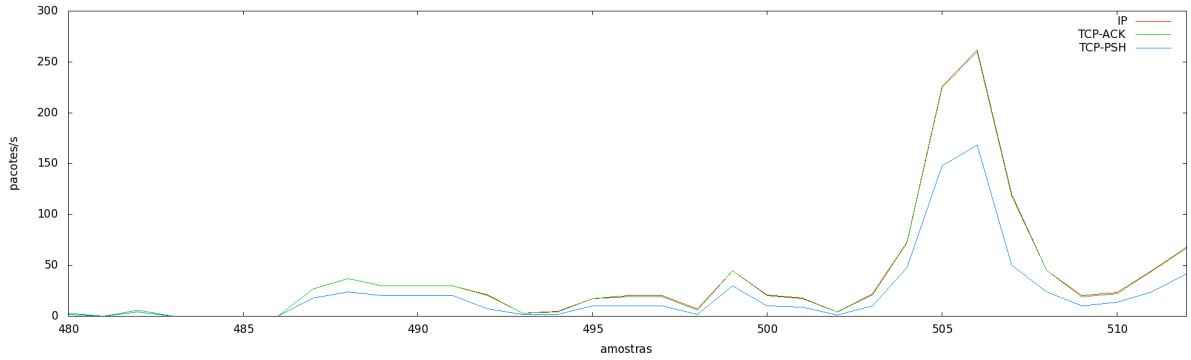

Figura 5.20: Variáveis selecionadas no momento do ataque Mailbomb 2

Nas Figuras 5.19 e 5.20 são mostradas as características selecionadas no momento dos ataques. Para que o alarme tenha sido emitido, todas estas características emitiram alarmes nos mesmos instantes, resultando nas taxas de detecção. Na Figura 5.19 observa-se a presença de 3 características selecionadas no momento do ataque, que são: TCP, IP e TCP-ACK. As variáveis TCP e IP mostram um comportamento muito similar, já a variável TCP-ACK possui curvas que possuem correlação à estas. Na Figura 5.20, 3 características foram selecionadas no instante do ataque: IP, TCP-ACK e TCP-PSH. As variáveis IP e TCP-ACK possuem curvas muito próximas, inclusive quando observa-se a quantidade de amostras, que no gráfico resulta em uma sobreposição. A variável TCP-PSH apresenta comportamento similar mas tênue, indicando correlacionamento destes descritores no momento da intrusão.

As Figuras 5.21 e 5.22 mostram os alarmes gerados na detecção dos ataques do tipo Neptune. É possível observar que próximo ao momento 500 para os dois casos, há alarmes que detectam a intrusão, sendo as demais variações decorrentes de falsos positivos, ou seja, quando o detector erra ao identificar uma intrusão.

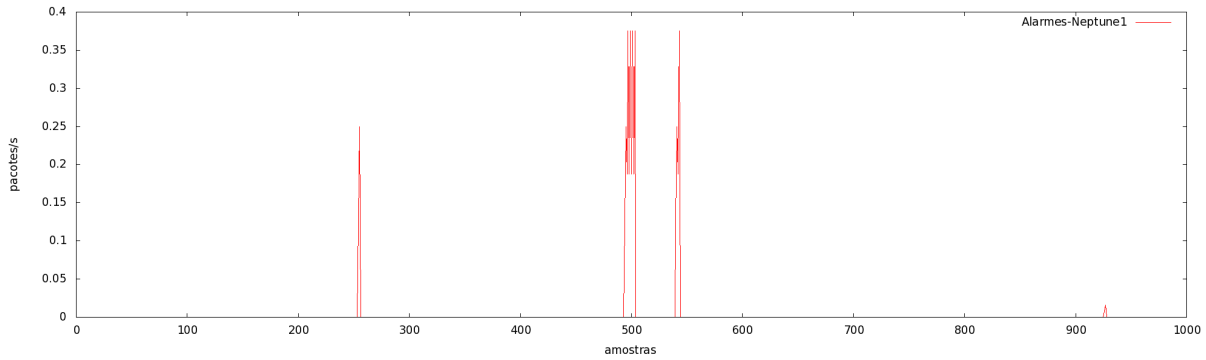

Figura 5.21: Alarmes Gerados na Detecção do Ataque *Neptune* 1

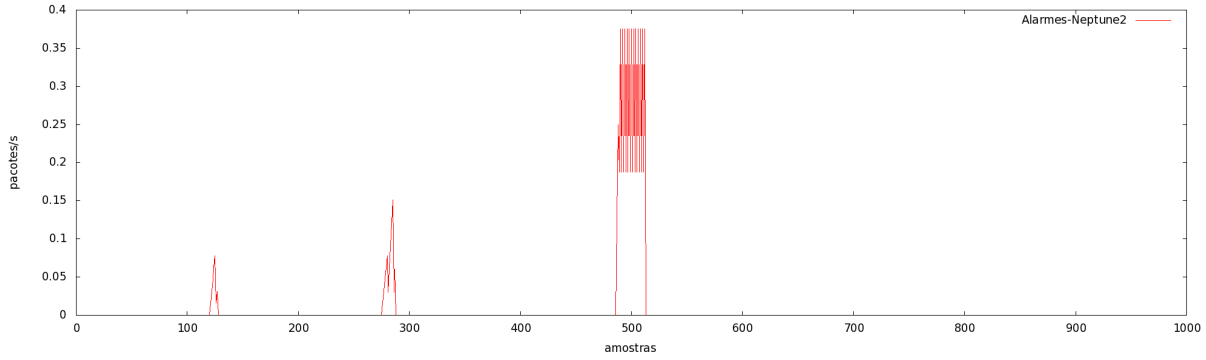

Figura 5.22: Alarmes Gerados na Detecção do Ataque *Neptune* 2

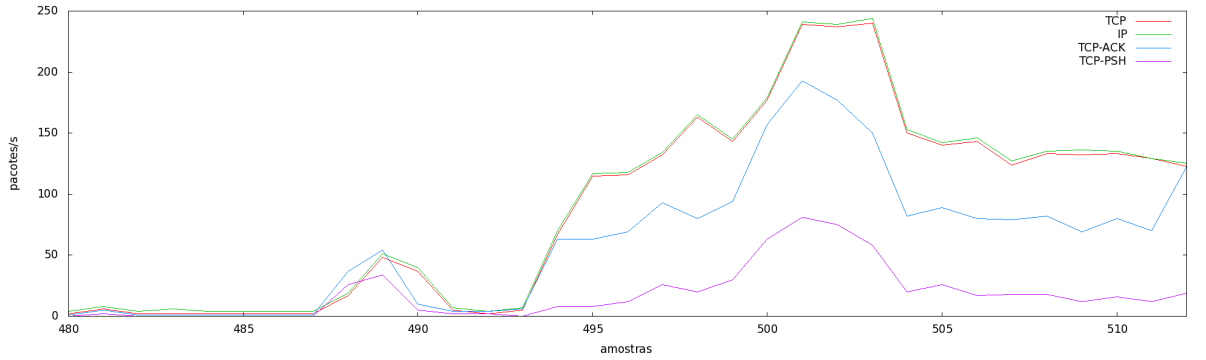

Figura 5.23: Variáveis selecionadas no momento do ataque Neptune 1

Através das Figuras 5.23 e 5.24 é possível identificar as características selecionadas na janela referente ao instante do ataque. Na Figura 5.23 foram selecionadas 4 características: TCP,

IP, TCP-ACK e TCP-PSH. Através da análise desta figura é possível perceber que as curvas entre as variáveis escolhidas possuem comportamento correlacionado, pois apresentam similaridade. A quantidade de pacotes por segundo atinge níveis próximos a 250.

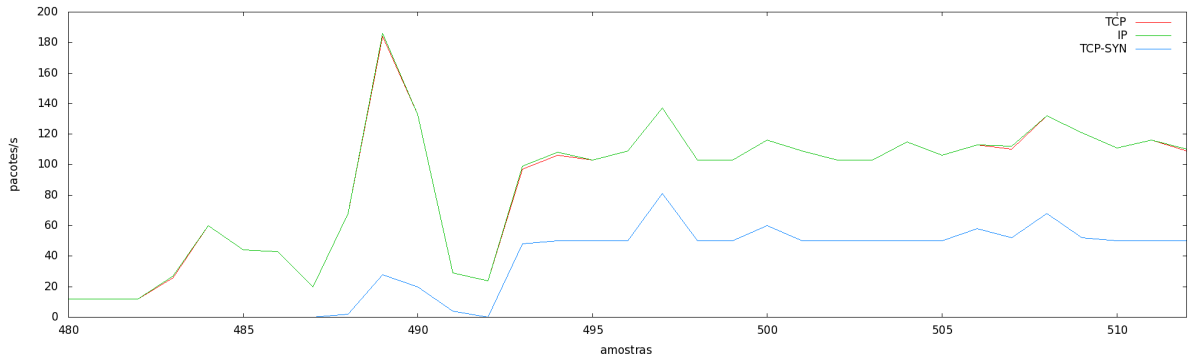

Figura 5.24: Variáveis selecionadas no momento do ataque Neptune 2

Na Figura 5.24 foram selecionadas 3 características: TCP, TCP-SYN e IP. Embora tenha sido selecionado um número menor, também é possível verificar a semelhança entre as variáveis selecionadas dentro do período da intrusão. Destaca-se o comportamento das variáveis TCP e IP que são muito parecidas e marca nível próximo a 200 pacotes por segundo, seguido da variável TCP-SYN que acompanha o tipo de mudanças, com máxima próximo de 80 pacotes por segundo.

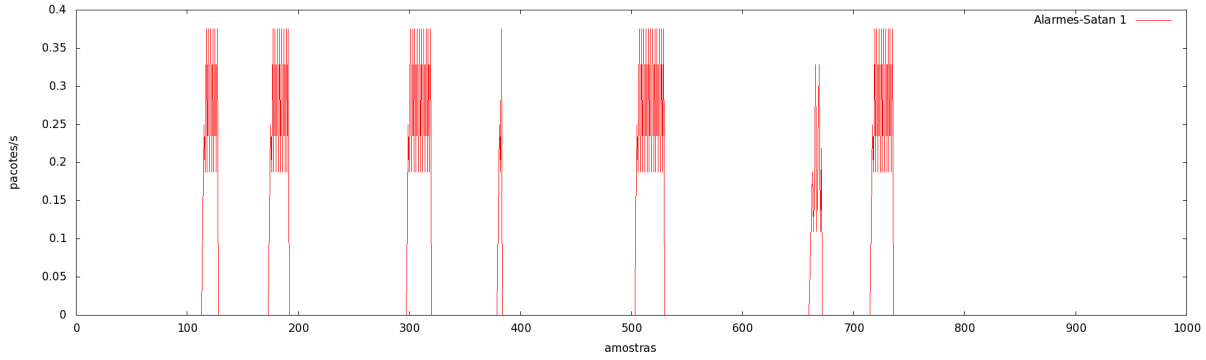

Figura 5.25: Alarmes Gerados na Detecção do Ataque *Satan* 1

As Figuras 5.25 e 5.26 mostram os alarmes gerados na detecção dos ataques SATAN 1 e SATAN 2. Na Figura 5.25 próximo ao instante 500 pode-se associar o alarme gerado ao momento do ataque, sendo falsos positivos os demais alarmes mostrados.

Na Figura 5.26 não foi possível associar alarmes ao momento do ataque, ou seja, esta ameaça não foi detectada pelo IDS, sendo verificados apenas alarmes associados a falsos positivos.

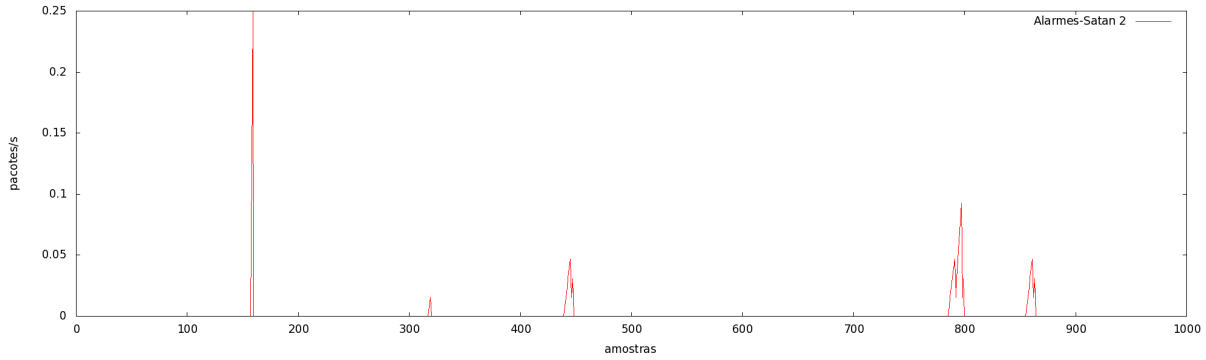

Figura 5.26: Alarmes Gerados na Detecção do Ataque *Satan* 2

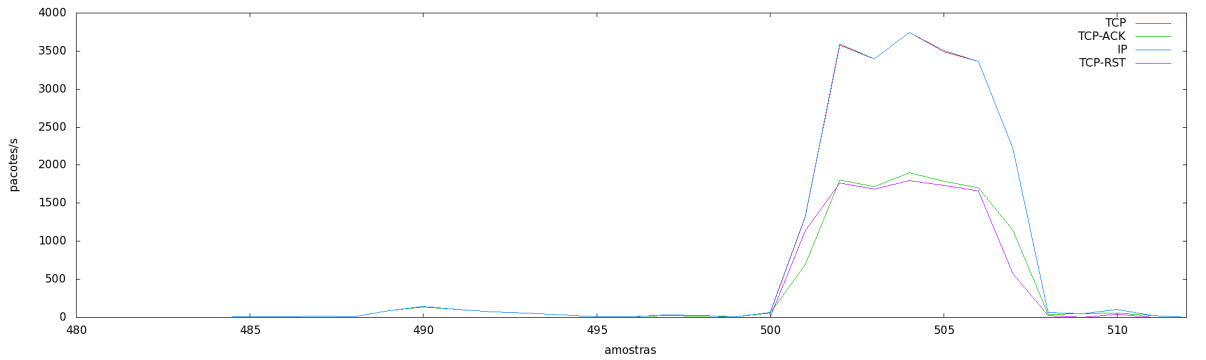

Figura 5.27: Variáveis selecionadas no momento do ataque Satan 1

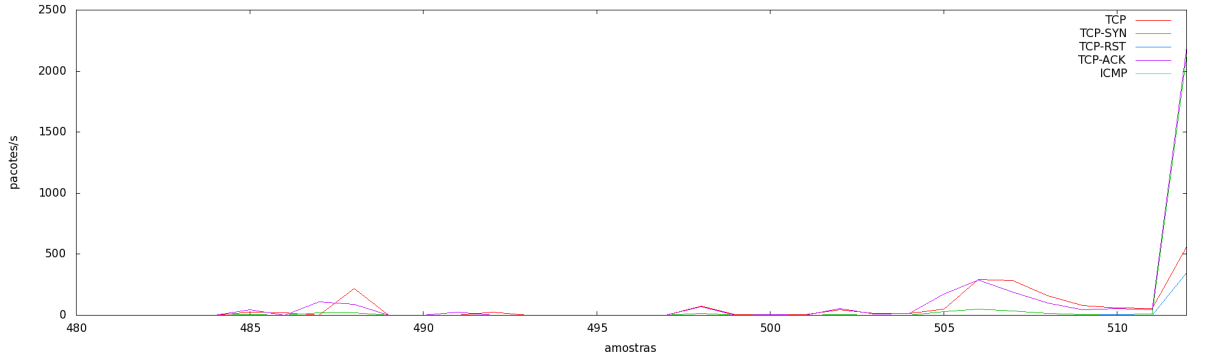

Figura 5.28: Variáveis selecionadas no momento do ataque Satan 2

Nas Figuras 5.27 e 5.28 observa-se as características selecionadas no momento dos ataques. A Figura 5.27 mostra que no ataque SATAN 1 foram selecionadas 4 características: TCP, IP, TCP-ACK e TCP-RST. É possível notar que o comportamento das curvas das variáveis envolvidas é similar, onde as variáveis TCP e IP tem comportamento de sobreposição na imagem alcançando níveis próximos à 4000 pacotes por segundo. As outras duas variáveis, TCP-ACK e TCP-RST atingem cerca de 2000 pacotes por segundo. Na Figura 5.28 é possível acompanhar as características selecionadas no momento do ataque SATAN 2, embora o detector não

tenha identificado este ataque. Foram selecionadas 5 características: TCP, TCP-SYN, TCP-RST, TCP-ACK e ICMP. Verifica-se níveis próximo a 2500 pacotes por segundo das variáveis TCP-SYN e TCP-ACK.

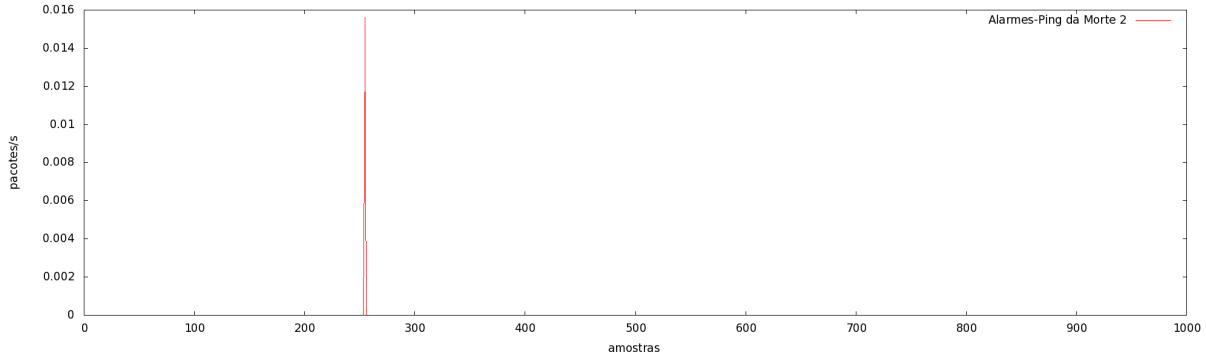

Figura 5.29: Alarmes Gerados na Detecção do Ataque Ping da Morte 2

Na Figura 5.29 é possível observar os alarmes gerados na detecção do ataques Ping da Morte 2. Nota-se que de acordo com o gráfico mostrado anteriormente que descreve o comportamento do protocolo ICMP neste caso, o momento do ataque é próximo ao instante 500. Através da análise da Figura 5.29 não existem alarmes neste momento, portanto o detector não foi capaz de identificar esta intrusão. A detecção do Ping da Morte 1 não produziu alarmes, ou seja, também não foi detectada pelo IDS. Este comportamento ausente do IDS justifica-se pelo comportamento do Ping da Morte, que não propaga variações abruptas em diferentes descritores do tráfego, dificultando a identificação pelo método utilizado.

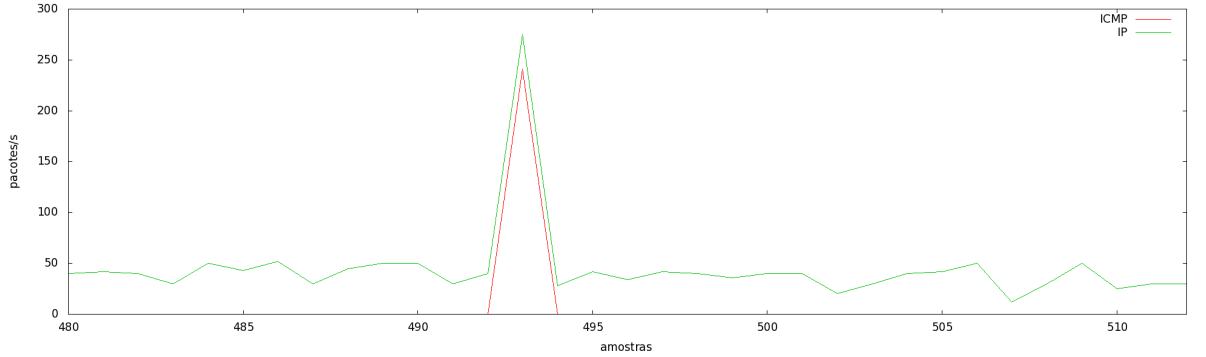

Figura 5.30: Variáveis selecionadas no momento do ataque Ping da Morte 1

As características selecionadas nos ataques Ping da Morte 1 e 2, podem ser observadas através das Figuras 5.30 e 5.31. Nos dois casos, apenas duas característica foram selecionadas: ICMP, IP.

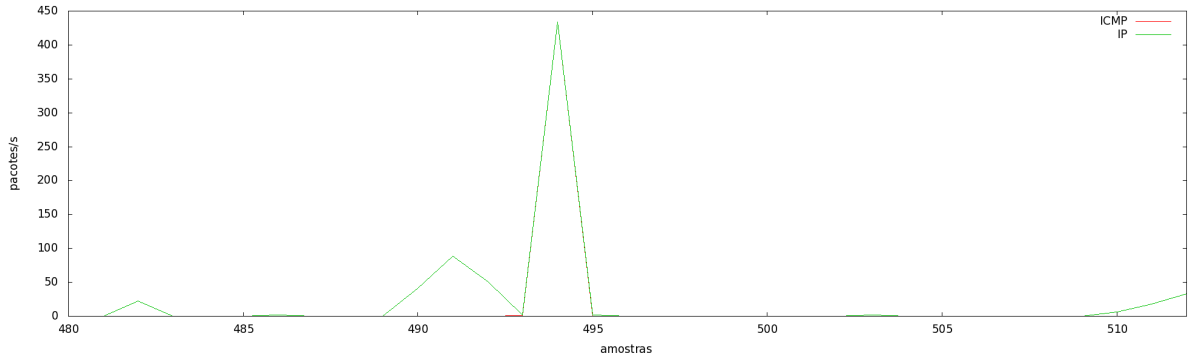

Figura 5.31: Variáveis selecionadas no momento do ataque Ping da Morte 2

5.5 Discussão dos Resultados Com e Sem o Método de Seleção SDCorr

O experimento utilizando o conjunto completo de variáveis (13 variáveis) não apresentou nenhum alarme, demonstrando ser ineficaz na detecção, devido à presença de variáveis não adequadas. No experimento em que foi usado o método de seleção proposto SDCorr, o detector apresentou resultados positivos em comparação ao experimento anterior. A Tabela 5.5 apresenta o resumo dos resultados obtidos, para o melhor caso, de janela com 32 elementos. Nesta tabela, os resultados dos dois experimentos estão presentes.

Tabela 5.5: Resumo dos Resultados obtidos

	Verdadeiros Positivos	Falsos Positivos	Falsos Negativos	Tempo (s)
SDCorr	5	279	3	61,078
Conjunto Completo	0	0	8	184,347

A Figura 5.32 mostra o gráfico de colunas dos resultados obtidos nos experimentos realizados. O experimento com o método de seleção SDCorr obteve uma taxa de acertos alcançou 62,5% e taxa de falsos positivos foi de 3,41%. Em comparação ao caso em que o conjunto completo de variáveis foi utilizado, o resultado é satisfatório, pois neste experimento não foram emitidos alarmes de nenhum tipo. Em relação ao tempo de execução, o experimento com o método de seleção em comparação ao experimento com o conjunto completo apresentou uma economia de tempo de 66,86%.

Diferentemente dos demais ataques, o Ping da Morte é representado por apenas uma amostra com variação significante. Enquanto os demais ataques possuem mais de uma amostra com variações que permitem a detecção através da técnica de análise de sinais abruptos. A presença de apenas uma amostra com variação não foi suficiente para que o método de detecção identifique que há uma anomalia, sendo a variação classificada como uma variação normal, uma vez

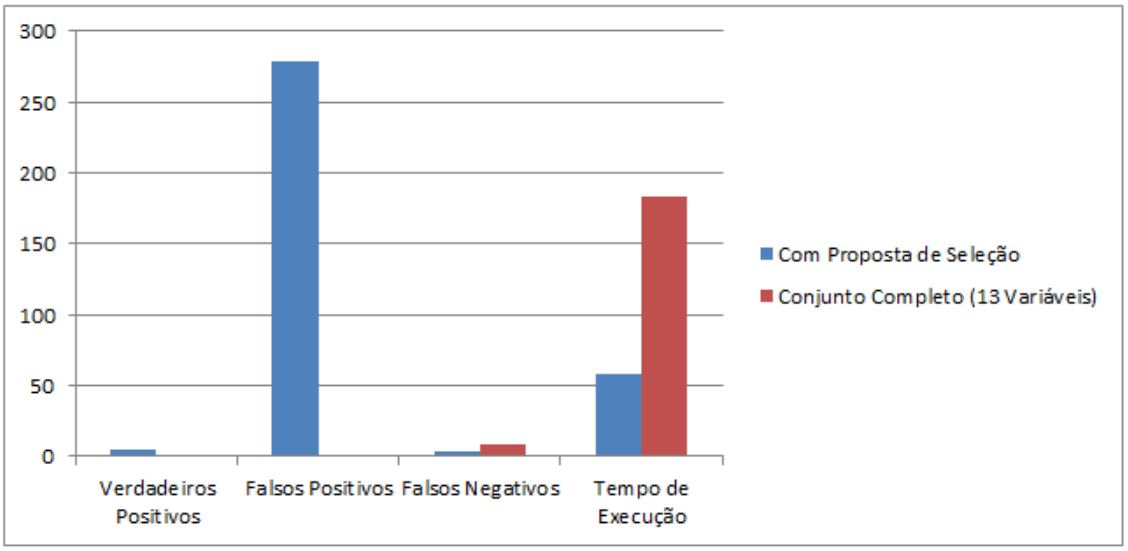

Figura 5.32: Gráfico de Comparação dos Resultados

que não se propaga para outras variáveis analisadas.

Uma questão importante é a quantidade mínima de variáveis que devem ser selecionadas para que seja possível a detecção através do método de detecção de sinais abruptos que avalia correlação. Como este método de detecção utiliza uma matriz de correlação que emprega multiplicações para filtrar as quantidades de alarmes individuais e gerar alarmes finais mais precisos, a quantidade mínima de variáveis pode ser considerada um problema a ser contornado em trabalhos futuros. Nos experimentos realizados não houve limite mínimo de variáveis a ser preenchido. Entretanto, um limitador mínimo pode diminuir a quantidade de falsos positivos. Por exemplo, considere o caso em que não foi selecionada nenhuma variável, existindo apenas a pivô informada. Ao assumir que intrusões geram perturbações significativas em mais do que uma variável, se apenas a variável pivô informada está presente, todos os alarmes gerados resultaram em alarmes finais positivos, não necessitando de matriz de correlação. Neste caso, não há filtro e a quantidade de falsos positivos pode crescer. Sendo assim, em casos em que a seleção não escolhe nenhuma ou apenas poucas variáveis, a probabilidade da taxa de falsos positivos crescer é maior. Diferentemente, no experimento com o conjunto completo não foram emitidos alarmes finais pois havia um grande número de variáveis que diminuiu as chances de que a matriz de correlação emitisse alarmes. Dessa forma, percebe-se a necessidade de equilíbrio no número de variáveis utilizadas, não podendo ser muito pequeno para que a quantidade de falsos positivos não seja alta e não podendo ser um número grande de variáveis para que a quantidade de alarmes não seja reduzida a valores nulos.

Para integração do método SDCorr aplicado em um ambiente real, algumas questões são importantes serem consideradas. É necessário que os dados coletados sejam formatados para possibilitar a verificação dos alarmes corretamente. Isto implica em utilizar um detector baseado em assinaturas para identificar as posições dos ataques e até mesmo a injeção de intrusões para ter a certeza de quantos ataques e seus momentos. Este processo permite a documentação da base, permitindo a comparação dos alarmes gerados pelo detector e assim identificar a presença dos ataques presentes. Além disto, a intrusividade do SDCorr em um ambiente real é um parâmetro importante e refere-se a interferência deste no ambiente que está inserido. O SDCorr destina-se a detecção em cenários *online*, onde a intrusividade é um ponto importante para que não atrase aplicações que estão situadas dentro do mesmo ambiente de rede. Desta forma, o SDCorr mostrou-se adequado pois provoca pouca interferência, sendo capaz de analisar uma amostra a cada 7,69 milissegundos, mostrando-se rápido e adequado a proteção de ambientes reais.

5.6 Conclusões Parciais

Neste Capítulo foram apresentados os experimentos realizados para validação da proposta do método de seleção de características SDCorr. Os experimentos utilizaram a técnica de detecção de intrusão baseada em análise de sinais abruptos. Foram executados dois experimentos: um sem seleção de características e outro com seleção de características através do SDCorr.

Os resultados mostraram que para o melhor caso, houve uma taxa de acertos (verdadeiros positivos) de 62,5%, contra 3,41% de falsos positivos. No quesito desempenho, em comparação ao experimento sem seleção de características, o teste com seleção pelo SDCorr apresentou uma economia de tempo de 66,86%. Tais resultados comprovam a eficácia do método de seleção SDCorr.

6 CONCLUSÃO

A facilidade de comunicação através da Internet trouxe ameaças, como intrusões. Uma intrusão é caracterizada como uma tentativa não autorizada de acesso ao sistema. Os IDS são utilizados em conjunto com abordagens tradicionais de proteção para aumentar a segurança. Entretanto, a precisão do IDS depende da qualidade dos dados envolvidos no processamento e a seleção de variáveis permite aumentar a precisão da detecção e diminuir o tempo de processamento.

Este trabalho propôs um método para seleção semi-automática e dinâmica de variáveis de rede com base na correlação entre variáveis chamado SDCorr (Seleção Dinâmica por Correlação). O método SDCorr baseia-se na utilização de uma variável pivô, escolhida de forma semi-automática, para selecionar variáveis de forma dinâmica através do método de correlação de *Pearson*.

O SDCorr foi testado com a utilização de um detector de intrusão baseado na análise de sinais abruptos e os experimentos realizados apresentaram resultados positivos. Houve, no melhor caso, uma economia de tempo de 66,86%, uma taxa de acertos de 62,5% e uma taxa de falsos positivos de 3,41%, quando comparado ao experimento com utilização de todas as variáveis (sem seleção). O método de seleção SDCorr também permitiu economia de tempo de processamento e aumento da taxa de detecção em comparação com o experimento sem seleção de variáveis. A metodologia mostrou-se adequada para aplicação no tratamento dos dados a serem aplicados a detectores de intrusão.

6.1 Contribuições

Uma estratégia utilizada para reduzir a taxa de falsos alarmes em IDSs baseados em anomalias é verificar se em um mesmo intervalo de tempo ocorrem mudanças abruptas em mais de uma variável de rede. Entretanto, esta estratégia assume como hipótese que as variáveis analisadas são correlacionadas, exigindo um procedimento prévio de seleção de variáveis.

Assim, este trabalho apresentou um método dinâmico de seleção de variáveis para IDS de rede, chamado SDCorr (Seleção Dinâmica por Correlação), que opera na modalidade de filtro e utiliza como avaliador o teste de correlação de Pearson. O método tem a característica de adaptar-se dinamicamente as variações do tráfego de rede por meio da seleção de novas variáveis a cada iteração com o detector, o que possibilita acompanhar as mudanças nos dados

e estabelecer correlações entre variáveis.

Como resultado, o método permitiu melhorar a precisão e desempenho do IDS, reduzindo a dimensão dos dados originais. O método selecionou variáveis dinamicamente através das correlações encontradas entre as variáveis, promoveu a redução do volume de dados através da seleção como pode ser observado pela Tabela 5.4 que mostrou a taxa de utilização das diferentes variáveis, melhorou a taxa de detecção em comparação ao caso sem seleção e permitiu economia de tempo de processamento na detecção de acordo com a Tabela 5.5, que indica uma taxa de acertos de 62,5% e economia de tempo de 66,86%.

Para a seleção dos dados, a abordagem proposta neste trabalho utilizou como pivô uma variável que teve variação abrupta decorrente de uma intrusão. As variáveis mais prováveis de conter informações relevantes foram selecionadas, refinando o conjunto de informações disponíveis e possibilitando que a detecção criasse alarmes mais precisos de acordo com correlações encontradas nas variáveis do tráfego de rede. Este trabalho apresentou uma metodologia para seleção das variáveis que pode ser utilizada em estratégias de detecção que exigem que as variáveis analisadas sejam correlacionadas, selecionando os dados dinamicamente através da escolha de novas variáveis a cada movimentação da janela deslizante.

6.2 Trabalhos Futuros

A contribuição deste trabalho permite que novos estudos sejam realizados. Como possíveis direções a serem tomadas a partir do estudo estão a utilização do método SDCorr em uma base de dados real, além de comparar o método de avaliação utilizado no selecionador com outras medidas de similaridade que possam auxiliar na análise da correlação entre variáveis de rede.

Uma outra direção pode ser a integração da proposta em um ambiente de gerenciamento de redes. Normalmente nestes ambientes, o andamento sobre os recursos é feito de forma manual com base no conhecimento e acompanhamento deste por um usuário capacitado. A característica dinâmica do SDCorr permite que o ambiente de rede seja monitorado de tempos em tempos, auxiliando o processo de conhecimento do estado atual das aplicações inseridas neste cenário.

Além disto, a escolha do tamanho das janelas foi definido manualmente. Visto que o tamanho delas pode interferir no resultado, um estudo viável aponta a adaptação dinâmica do tamanho da janela de acordo com as correlações encontradas entre as variáveis envolvidas, permitindo maior afinidade com o conjunto analisado, aumentando ou diminuindo o tamanho das

janelas automaticamente.

REFERÊNCIAS

ABOUABDALLA, O.; EL-TAJ, H.; MANASRAH, A.; RAMADASS, S. False positive reduction in intrusion detection system: a survey. **2009 2nd IEEE International Conference on Broadband Network & Multimedia Technology**, [S.l.], p.463–466, Oct. 2009.

AZEVEDO, R. P.; MOZZAQUATRO, B. A.; CAPPO, C.; NUNES, R.; SCHAERER, C. E.; KOZAKEVICIUS, A. A Bidimensional Wavelet Transform based Algorithm for DoS Attack Detection. **LADC - Fifth Latin-American Symposium on Dependable Computing, 2011, São José dos Campos.**, [S.l.], 2011.

BAI, Y.; KOBAYASHI, H. Intrusion detection system: technology and development. **Proceedings of the 17th International Conference on Advanced Information Networking and Applications**, [S.l.], p.1–6, 2003.

BOLÓN-CANEDO, V.; SANCHEZ-MAROO, N.; ALONSO-BETANZOS, A. A combination of discretization and filter methods for improving classification performance in KDD Cup 99 dataset. In: NEURAL NETWORKS, 2009. IJCNN 2009. INTERNATINAL JOINT CONFERENCE ON, 2009. **Anais...** IEEE, 2009. p.359–366.

CABRERA, J.; LEWIS, L.; QIN, X.; LEE, W.; MEHRA, R. Proactive intrusion detection and distributed denial of service attacks - a case study in security management. **Journal of Network and Systems Management**, [S.l.], v.10, n.2, p.225–254, 2002.

DARPA. **Defense Advanced Research Projects Agency**. [S.l.: s.n.], 1999. Disponível em `http://www.ll.mit.edu/IST/ideval/index.html`. Último acesso em Novembro de 2011.

DEBAR, H. An introduction to intrusion detection systems. **Proceedings of Connect2000, Doha, Qatar**, [S.l.], Apr. 2000.

EKTEFA, M.; MEMAR, S.; SIDI, F.; AFFENDEY, L. Intrusion detection using data mining techniques. In: INFORMATION RETRIEVAL & KNOWLEDGE MANAGEMENT,(CAMP), 2010 INTERNATIONAL CONFERENCE ON, 2010. **Anais...** IEEE, 2010. p.200–203.

EL-KHATIB, K. Impact of Feature Reduction on the Efficiency of Wireless Intrusion Detection Systems. **IEEE Transactions on Parallel and Distributed Systems**, [S.l.], v.21, n.8, p.1143–1149, Aug. 2010.

GRANGER, C. W. J. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. **Econometrica**, [S.l.], v.37, n.3, p.424–38, 1969.

GUJARATI, D. Basic econometrics. **McGraw Hill**, [S.l.], 2004.

GUPTA, K. K.; NATH, B.; KOTAGIRI, R. Random Fields for Intrusion Detection. **TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING**, [S.l.], v.7, n.1, p.35–49, 2010.

HENKE, M.; SANTOS, C.; NUNAN, E.; FEITOSA, E.; SANTOS, E. dos; SOUTO, E. Aprendizagem de Máquina para Segurança em Redes de Computadores: métodos e aplicações. **Capítulo 2 – Livro de Minicursos do SBSEG 2011, Brasília–DF, 51p.**, [S.l.], 2011.

HOON, C.; SHIN, S. W.; CHUNG, J. W. Network Intrusion Detection Through Genetic Feature Selection. **Seventh ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing (SNPD'06)**, [S.l.], p.109–114, 2006.

KALOUSIS, A.; PRADOS, J.; HILARIO, M. Stability of feature selection algorithms. In: DATA MINING, FIFTH IEEE INTERNATIONAL CONFERENCE ON, 2005. **Anais...** [S.l.: s.n.], 2005. p.8 pp.

KHOR, K.-C.; TING, C.-Y.; AMNUAISUK, S.-P. A Feature Selection Approach for Network Intrusion Detection. In: INFORMATION MANAGEMENT AND ENGINEERING, 2009. ICIME '09. INTERNATIONAL CONFERENCE ON, 2009. **Anais...** [S.l.: s.n.], 2009. p.133 –137.

KR, K.; INDRA, A. Intrusion Detection Tools and Techniques - A Survey. **International Journal of Computer Theory and Engineering**, [S.l.], v.2, n.6, p.901–906, 2010.

KUMAR, P. A. R.; SELVAKUMAR, S. Distributed Denial-of-Service (DDoS) Threat in Collaborative Environment - A Survey on DDoS Attack Tools and Traceback Mechanisms. **2009 IEEE International Advance Computing Conference**, [S.l.], p.1275–1280, 2009.

LAKHINA, A.; CROVELLA, M.; DIOT, C. Mining anomalies using traffic feature distributions. **ACM SIGCOMM Computer Communication Review**, [S.l.], v.35, n.4, p.217–228, 2005.

LIM, S.; JONES, A. Network anomaly detection system: the state of art of network behaviour analysis. **International Conference on Convergence and Convergence and Hybrid Information Technology 2008**, [S.l.], p.459–465, 2008.

MARHUSIN, M. F.; CORNFORTH, D.; LARKIN, H. An overview of recent advances in intrusion detection. **2008 8th IEEE International Conference on Computer and Information Technology**, [S.l.], p.432–437, July 2008.

MECHTRI, L.; DJEMILI TOLBA, F.; GHOUALMI, N. Intrusion detection using principal component analysis. In: ENGINEERING SYSTEMS MANAGEMENT AND ITS APPLICATIONS (ICESMA), 2010 SECOND INTERNATIONAL CONFERENCE ON, 2010. **Anais...** [S.l.: s.n.], 2010. p.1 –6.

MURALI, A. A survey on intrusion detection approaches. **Information and Communication**, [S.l.], p.233–240, 2005.

NAGARAJAN, R.; UPRETI, M. Correlation Statistics for cDNA Microarray Image Analysis. **IEEE/ACM Trans. Comput. Biol. Bioinformatics**, Los Alamitos, CA, USA, v.3, n.3, p.232–238, July 2006.

NETO, A.; VICTORINO, A.; FANTONI, I.; ZAMPIERI, D. Real-time dynamic power management based on Pearson's Correlation Coefficient. In: ADVANCED ROBOTICS (ICAR), 2011 15TH INTERNATIONAL CONFERENCE ON, 2011. **Anais...** [S.l.: s.n.], 2011. p.304 –309.

NGUYEN, H. T.; FRANKE, K.; PETROVIC, S. Towards a Generic Feature-Selection Measure for Intrusion Detection. **2010 20th International Conference on Pattern Recognition**, [S.l.], p.1529–1532, Aug. 2010.

NZIGA, J. Minimal dataset for Network Intrusion Detection Systems via dimensionality reduction. In: DIGITAL INFORMATION MANAGEMENT (ICDIM), 2011 SIXTH INTERNATIONAL CONFERENCE ON, 2011. **Anais...** [S.l.: s.n.], 2011. p.168 –173.

OSTASZEWSKI, M.; BOUVRY, P.; SEREDYNSKI, F. An approach to intrusion detection by means of idiotypic networks paradigm. In: EVOLUTIONARY COMPUTATION, 2008. CEC 2008.(IEEE WORLD CONGRESS ON COMPUTATIONAL INTELLIGENCE). IEEE CONGRESS ON, 2008. **Anais...** IEEE, 2008. p.2099–2108.

PENG, T.; LECKIE, C.; RAMAMOHANARAO, K. Survey of network-based defense mechanisms countering the DoS and DDoS problems. **ACM Computing Surveys**, [S.l.], v.39, n.1, p.3–es, Apr. 2007.

PILLI, E.; JOSHI, R.; NIYOGI, R. Data reduction by identification and correlation of TCP/IP attack attributes for network forensics. In: INTERNATIONAL CONFERENCE & WORKSHOP ON EMERGING TRENDS IN TECHNOLOGY, 2011. **Proceedings...** ACM, 2011. n.Icwet, p.276–283.

PUMA-VILLANUEVA, W.; SANTOS, E. dos; VON ZUBEN, F. Data partition and variable selection for time series prediction using wrappers. **The 2006 IEEE International Joint Conference on Neural Network Proceedings**, [S.l.], p.4740–4747, 2006.

R Development Core Team. **R**: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2012. ISBN 3-900051-07-0 - `http://www.R-project.org`.

SABAHI, F.; MOVAGHAR, A. Intrusion Detection: a survey. **2008 Third International Conference on Systems and Networks Communications**, [S.l.], p.23–26, Oct. 2008.

SHANKARAPANI, M.; KANCHERLA, K.; RAMAMMOORTHY, S.; MOVVA, R.; MUKKAMALA, S. Kernel machines for malware classification and similarity analysis. In: NEURAL NETWORKS (IJCNN), THE 2010 INTERNATIONAL JOINT CONFERENCE ON, 2010. **Anais...** IEEE, 2010. p.1–6.

SHANMUGAM, B.; IDRIS, N. Improved Intrusion Detection System Using Fuzzy Logic for Detecting Anamoly and Misuse Type of Attacks. In: SOFT COMPUTING AND PATTERN RECOGNITION, 2009. SOCPAR '09. INTERNATIONAL CONFERENCE OF, 2009. **Anais...** [S.l.: s.n.], 2009. p.212 –217.

SHEEN, S.; RAJESH, R. Network intrusion detection using feature selection and Decision tree classifier. **TENCON 2008 - 2008 IEEE Region 10 Conference**, [S.l.], p.1–4, Nov. 2008.

SPEROTTO, A.; SCHAFFRATH, G.; SADRE, R.; MORARIU, C.; PRAS, A.; STILLER, B. An Overview of IP Flow-Based Intrusion Detection. **IEEE Communications Surveys & Tutorials**, [S.l.], v.12, n.3, p.343–356, 2010.

SUEBSING, A.; HIRANSAKOLWONG, N. Feature Selection Using Euclidean Distance and Cosine Similarity for Intrusion Detection Model. In: INTELLIGENT INFORMATION AND DATABASE SYSTEMS, 2009. ACIIDS 2009. FIRST ASIAN CONFERENCE ON, 2009. **Anais...** [S.l.: s.n.], 2009. p.86 –91.

TANENBAUM, A. Computer Networks. **Prentice Hall Professional Technical Reference**, [S.l.], 2002.

THOTTAN, M.; JI, C. Adaptive thresholding for proactive network problem detection. **Proceedings of the IEEE Third International Workshop on Systems Management**, [S.l.], p.108–116, 1998.

THOTTAN, M.; JI, C. Anomaly detection in IP networks. **IEEE Transactions on Signal Processing**, [S.l.], v.51, n.8, p.2191–2204, Aug. 2003.

TUPAKULA, U.; VARADHARAJAN, V.; PANDALANENI, S. DoSTRACK: a system for defending against dos attacks. In: ACM SYMPOSIUM ON APPLIED COMPUTING, 2009., 2009. **Proceedings...** ACM, 2009. p.47–53.

URBANEK, S. A fast way to provide R functionality to applications. In: DISTRIBUTED STATISTICAL COMPUTING (DSC), 2003. **Anais...** [S.l.: s.n.], 2003.

VOGT, F. **Detecção de Intrusão Através de Correlação de Variáveis**. 2011. Dissertação (Mestrado) — Universidade Federal de Santa Maria – UFSM/PPGEP – Santa Maria, RS.

WANG, J. Computer network security: theory and practice. **Springer Publishing Company, Incorporated**, [S.l.], 2009.

WU, Q.; SHAO, Z. Network Anomaly Detection Using Time Series Analysis. **Joint International Conference on Autonomic and Autonomous Systems and International Conference on Networking and Services - (icas-isns'05)**, [S.l.], p.42–42, 2005.

XIONG, H.; SHEKHAR, S.; TAN, P.-N.; KUMAR, V. Exploiting a support-based upper bound of Pearson's correlation coefficient for efficiently identifying strongly correlated pairs. In: ACM

SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 2004, New York, NY, USA. **Proceedings...** ACM, 2004. p.334–343. (KDD '04).