Universidade Estadual de Londrina

PAULO ROBERTO GALEGO HERNANDES JUNIOR

DETECÇÃO DE ANOMALIAS COM ASSINATURA DIGITAL UTILIZANDO ALGORITMO GENÉTICO E ANÁLISE DE FLUXOS IP

Londrina
2016

PAULO ROBERTO GALEGO HERNANDES JUNIOR

DETECÇÃO DE ANOMALIAS COM ASSINATURA DIGITAL UTILIZANDO ALGORITMO GENÉTICO E ANÁLISE DE FLUXOS IP

Dissertação apresentada ao Programa de Mestrado em Ciência da Computação da Universidade Estadual de Londrina para obtenção do título de Mestre em Ciência da Computação.

Orientador: Prof. Dr. Mario Lemes Proença Junior

Londrina
2016

Paulo Roberto Galego Hernandes Junior

Detecção de anomalias com assinatura digital utilizando algoritmo genético e análise de fluxos IP/ Paulo Roberto Galego Hernandes Junior. – Londrina–PR, 2016-

107 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Mario Lemes Proença Junior

– Universidade Estadual de Londrina, 2016.

1. Redes de computadores. 2. Anomalias. I. Proença Junior, Mario Lemes. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação. IV. Título

CDU 02:141:005.7

PAULO ROBERTO GALEGO HERNANDES JUNIOR

DETECÇÃO DE ANOMALIAS COM ASSINATURA DIGITAL UTILIZANDO ALGORITMO GENÉTICO E ANÁLISE DE FLUXOS IP

Dissertação apresentada ao Programa de Mestrado em Ciências da Computação da Universidade Estadual de Londrina para a obtenção do título de Mestre em Ciência da Computação.

BANCA EXAMINADORA

Orientador: Prof. Dr. Mario Lemes Proença Junior
Universidade Estadual de Londrina – UEL

Prof. Dr. Cinthyan Renata Sachs C. Barbosa
Universidade Estadual de Londrina – UEL

Prof. Dr. Bruno Bogaz Zarpelão
Universidade Estadual de Londrina – UEL

Prof. Dr. Lourival Aparecido de Góis
Universidade Tecnológica Federal do Paraná – UTFPR, Campus de Ponta Grossa

Londrina, 05 de maio de 2016.

À minha mãe Lúcia, pelo exemplo de vida e humildade. Ao meu pai Paulo e meu irmão Fernando, pelo apoio incondicional. À Amanda, minha noiva e companheira pela paciência e apoio em todos os momentos. Às minhas avós Etelvina e Eliza, que eu perdi durante este trajeto mas me deram muita força nesta caminhada.

AGRADECIMENTOS

Agradeço a Deus pela vida, pela saúde e pela proteção nas inúmeras viagens à Londrina.

Agradeço a minha família, meus pais e irmão, que sempre estiveram ao meu lado em todas as decisões que eu tomei. À Amanda, minha noiva, que sempre acreditou em mim e me apoiou durante essa dura jornada.

Ao meu orientador, professor Dr. Mario L. Proença Jr., pela motivação, empenho e pela confiança no meu trabalho.

Agradeço à professora Dra. Cinthyan Renata Sachs C. Barbosa, por me apresentar ao professor Mario, por acreditar no meu trabalho e por todas as palavras de incentivo.

Ao professor Dr. Bruno Zarpelão pelo empenho nas aulas e por compartilhar seu conhecimento de forma excepcional com todos seus alunos.

Aos amigos do grupo de redes, Gilberto, Eduardo, Marcos, e em especial ao Luiz, que sempre esteve disponível, auxiliando, apoiando e compartilhando seu conhecimento, sem o qual eu não conseguiria completar esta tarefa.

Aos amigos e companheiros de mestrado, Fadir, Alex e João, pelo incentivo durante as muitas horas na estrada entre Ourinhos e Londrina.

À Faculdade de Tecnologia (Fatec) Ourinhos, sua diretoria, seus professores, funcionários e alunos, onde já estudei e hoje leciono, por toda importância que representam na minha vida profissional e pessoal.

A todos que contribuíram para que pudesse realizar esse trabalho.

"Um homem precisa viajar. Por sua conta,
não por meio de histórias, imagens, livros ou TV.
Precisa viajar por si, com seus olhos e pés,
para entender o que é seu. (...)
Um homem precisa viajar para lugares que não conhece
para quebrar essa arrogância,
que nos faz ver o mundo como o imaginamos,
e não simplesmente como é ou pode ser;
que nos faz professores e doutores do que não vimos,
quando deveríamos ser alunos, e simplesmente ir ver."
Amyr Klink, navegador e escritor brasileiro

HERNANDES JR., PAULO R. GALEGO. **Detecção de anomalias com assinatura digital utilizando algoritmo genético e análise de fluxos IP**. 107 p. Dissertação de Mestrado (Mestrado em Ciência da Computação) – Universidade Estadual de Londrina, Londrina–PR, 2016.

RESUMO

Detectar anomalias em redes de computadores com precisão e rapidez é fundamental para que o diagnóstico e a solução dos problemas sejam encontrados rapidamente. Este trabalho apresenta a ferramenta *Genetic Algorithm for Digital Signature* (GADS), que identifica padrões anômalos no comportamento de uma rede através do uso de uma assinatura digital de segmento de rede, chamada *Digital Signature of Network Segment using Flow Analysis* (DSNSF). Foram usados dados reais extraídos de duas universidades, adquiridos por meio das técnicas de fluxos IP, usando os padrões Netflow e sFlow. São usados seis atributos desses dados: bits/s, pacotes/s, os endereços IP de origem e destino e as portas de origem e destino. Para a organização dos dados e geração das assinaturas digitais, é usado o Algoritmo Genético, uma ferramenta para solucionar problemas de otimização. Para a detecção de anomalias são usadas três técnicas que já foram abordadas em conjunto com DSNSFs, o *Adaptive Dynamic Time Warping*, o uso de limiares (*thresholds*) e a lógica paraconsistente. As três técnicas são comparadas e testadas para a avaliação da eficácia destas com o GADS usando dados reais de dois cenários diferentes, a Universidade Estadual de Londrina e a Universidade Tecnológica Federal do Paraná, campus Toledo. Nos testes são buscadas as melhores taxas de detecção de anomalias, para que o administrador não seja sobrecarregado com falsos alarmes.

Palavras-chave: Assinatura digital. Detecção de anomalias. sFlow. Netflow. Algoritmo Genético. DSNSF. Fluxos IP.

HERNANDES JR., PAULO R. GALEGO. **Anomaly detection through digital signature using genetic algorithm and IP flows analysis.**. 107 p. Master's Thesis (Master in Science in Computer Science) – State University of Londrina, Londrina–PR, 2016.

ABSTRACT

In computer networks, detecting anomalies accurately and in a fast way is fundamental to rapid diagnosis and repair of problems. This work presents the Genetic Algorithm for Digital Signature (GADS), a tool that identifies network anomalous behavior through the use of a Digital Signature of Network Segment using Flow Analysis (DSNSF). A real data set extracted from two universities was used, through the use of IP data flow techniques, using the Netflow and sFlow patterns. Six attributes from these flows were used: bits/s, packets/s, source and destination of IP addresses and source and destination of logical ports. To organize these data and generate the digital signatures, the Genetic Algorithm was used, which is a tool used to solve optimization problems. For the anomaly detection, three techniques were used, and all of them have already approached the DSNSF. The first uses the Adaptive Dynamic Time Warping (ADTW), the second uses simple thresholds and the third, the Paraconsistent Logic. These three techniques were compared and tested to evaluate their efficiency combined with GADS, using a real data set of two different scenarios extracted from the State University of Londrina and the Federal University of Technology Paraná, city of Toledo. Tests were performed to find the best anomaly detection rate, so that the administrators are not overloaded with a high false positive rate.

Keywords: Digital Signature. Anomaly Detection. sFlow. Netflow. Genetic Algorithm. IP Flows. DSNSF.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

ACO	*Ant Colony Optimization*
ACODS	*Ant Colony Optimization for Digital Signature*
ADTW	*Adaptive Dynamic Time Warping*
ARIMA	*Autoregressive Integrated Moving Average*
CC	Coeficiente de Correlação
DoS	*Denial of Service*
DSNS	*Digital Signature of Network Segment*
DSNSF	*Digital Signature of Network Segment using Flow Analysis*
DTW	*Dynamic Time Warping*
GA	*Genetic Algorithm* (Algoritmo Genético)
GADS	*Genetic Algorithm for Digital Signature*
HWDS	*Holt-Winters for Digital Signature*
IPS	*Intrusion Prevention System*
LP	Lógica Paraconsistente
MPC	Máquina Paraconsistente Correlacional
NIDS	*Network Intrusion Detection Systems*
NMSE	*Normalized Mean Square Error*
PCA	*Principal Component Analysis*
PCADS	*Principal Component Analysis for Digital Signature*
ROC	*Receiver Operating Characteristic*
SNMP	*Simple Network Management Protocol*
SUS	*Stochastic Universal Sampling*
UEL	Universidade Estadual de Londrina
UTFPR	Universidade Tecnológica Federal do Paraná

LISTA DE SÍMBOLOS

J	Distância Euclidiana
K	Número de clusters
P_c	Ponto de *crossover*
P_i	Número de indivíduos da população inicial
X	Vetor representando a quantidade de dados de cada atributo
$f(s)$	Função de Silhouette
γ	Número de iterações do algoritmo
ρ	População inicial do GA
σ	Coeficiente de significância para valores de limiares
τ	População após aplicação de *crossover*

SUMÁRIO

1 INTRODUÇÃO

Redes de computadores são dinâmicas e complexas, assim como a sua configuração e manutenção. Tipicamente, estas redes tem uma grande quantidade de roteadores, switches, firewalls, servidores e inúmeros outros elementos, com muitos tipos de eventos ocorrendo simultaneamente. Com um número cada vez mais crescente de redes interconectadas, a gerência destes serviços vem se tornando uma tarefa cada vez mais complexa e de grande responsabilidade.

Por muitos anos, administradores de redes buscaram informações sobre como gerenciar seus recursos usando o protocolo *Simple Network Management Protocol* (SNMP). Entretanto, este protocolo não era mais adequado para entregar todas as informações necessárias, face aos novos requisitos de qualidade de serviços requeridos pelas atuais redes de banda larga e serviços integrados. Além disso, serviços como voz sobre IP (VoIP - *Voice over IP*), e vídeo sob demanda estão cada vez mais acessíveis e com menores custos.

Com o surgimento dos fluxos IP, essa demanda pode ser suprida, proporcionando aos administradores as informações requeridas, de acordo com as necessidades de gerência impostas para os novos paradigmas de redes de serviços integrados. Com esse novo paradigma, a área de gerência de redes ganhou mais informações para a tomada de decisão, no que tange ao aumento da integridade, disponibilidade e confiabilidade dos dados trafegados. Para assegurar a qualidade nos serviços de suas redes, os atuais operadores devem assegurar a precisão do monitoramento do tráfego, fornecendo diagnósticos imediatos mediante quaisquer anomalias encontradas.

Uma anomalia é qualquer desvio de comportamento da rede em análise. Anomalias podem ser intencionais, como ataques deliberados, ou não intencionais, como um aumento abrupto no uso da Internet, por conta de um resultado de um concurso público por exemplo, neste caso conhecido como *Flash Crowd*. Como ataques intencionais, destacam-se os ataques *Denial of Service* (DoS), *Distributed Denial of Service* (DDoS), *port scans*, espionagem digital e até guerras cibernéticas têm sido travadas nos últimos anos [4].

Para conter o avanço dessas ameaças, são encontradas ferramentas como sistemas de detecção de intrusão baseados em redes, os *Network Intrusion Detection Systems* (NIDS), antivirus, firewalls e sistemas de prevenção de intrusão, os *Intrusion Prevention Systems* (IPS), porém o uso apenas dessas ferramentas não torna nenhuma rede imune a esses perigos [5].

Diante deste cenário, vê-se que o estudo de técnicas que identifiquem e classifiquem as anomalias de redes, de forma proativa e não supervisionada, é cada vez mais desafiador, dada a heterogeneidade dos dados trafegados ocasionada pela quantidade de serviços disponíveis nas atuais redes. Na literatura, em geral são encontradas duas formas de classificar anomalias em redes. A primeira engloba os métodos baseados em assinaturas de ataque, trabalhando de forma análoga aos atuais antivírus, obrigando o administrador a atualizar sua base de informações constantemente. A segunda compreende as técnicas baseadas na criação de perfis de rede, classificando o tráfego de acordo com seu histórico, permitindo a identificação de desvios de comportamento através de alterações nas estimativas de padrões pré estabelecidas [6].

Criar este perfil da rede, ou caracterizar o tráfego, é definir qual o comportamento esperado daquela rede em relação a certos parâmetros, que podem ser a quantidade de bits trafegados por segundo, por exemplo. Após a fase de caracterização do tráfego, poderão ser aplicadas métricas que identifiquem desvios de comportamentos nesta rede, e que serão o objeto de estudo nesta dissertação.

Este trabalho apresenta um sistema de detecção de anomalias em redes, que usa o *Genetic Algorithm* (GA), conhecido como Algoritmo Genético, e análise de fluxos IP. Para isso serão usados seis atributos fornecidos pelos padrões de fluxos IP, sendo dois atributos quantitativos, que são a quantidade de bits e pacotes por segundo, e quatro atributos qualitativos, que são as entropias de endereços IP e Portas de origem e destino, onde será usada a Entropia de Shannon [7]. O sistema apresentado, que irá realizar esse processo recebeu o nome de *Genetic Algorithm for Digital Signature* (GADS), e deverá trabalhar em duas etapas.

A primeira etapa do GADS faz a caracterização do tráfego, usando os seis atributos já citados para gerar a assinatura digital deste segmento de rede, que vai descrever o comportamento normal esperado para essa rede. Essa assinatura será nomeada como *Digital Signature of Network Segment using Flow Analysis* (DSNSF). Esse processo começa com a extração das informações e seu tratamento, como transformar atributos descritivos, como números de portas e endereços IP, em informações quantitativas, para que possam ser aplicadas as mesmas regras que os atributos bits e pacotes. Depois os dados são clusterizados, utilizando o GA para o otimização do processo de clusterização.

Na segunda etapa será feita a identificação do tráfego anômalo, quando o tráfego real será comparado com o DSNSF, ou seja, o tráfego previsto será confrontado com o atual e serão aplicadas técnicas que identifiquem o desvio de comportamento. Serão comparadas três técnicas já estudadas pelo Grupo de Pesquisa em Redes de Computadores e Comunicação de Dados da Universidade Estadual de Londrina (UEL),

e já descritas em trabalhos anteriores [8] [9] [10] [11]. A primeira utiliza a abordagem *Adaptive Dynamic Time Warping* (ADTW), uma modificação da técnica do tipo *pattern matching*, desenvolvida inicialmente para o reconhecimento de padrões de voz. A segunda abordagem utiliza *thresholds* simples, ou limiares, que vão definir o limite entre o tráfego considerado normal e o anômalo. A terceira utiliza a lógica paraconsistente, que considera evidências favoráveis e contrárias a uma determinada proposição, utilizando uma sequência de procedimentos para extrair informações evidenciais derivadas do contraste entre cada DSNSF e as leituras reais de tráfego.

O GADS foi submetido a testes para a validação de seus resultados. Para tanto, foram utilizados dados reais extraídos da Universidade Estadual de Londrina (UEL) e da Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Toledo. Além disso, o tráfego real foi submetido à injeção de tráfegos anômalos para que o estudo fosse complementado, verificando a eficácia do GADS na deteção de ataques.

O trabalho está organizado da seguinte forma: no capítulo 2 são apresentados os trabalhos relacionados ao tema proposto, como caracterização do tráfego, detecção de anomalias e algoritmos genéticos. O capítulo 3 detalha a caracterização do tráfego utilizando o *Genetic Algorithm for Digital Signature*, as três técnicas usadas na deteção de anomalias e quais a métricas usadas para a validação dos resultados. O capítulo 4 apresenta os resultados das validações realizadas que comprovam a eficácia do sistema em ambientes reais e simulados. Por fim, o capítulo 5 apresenta as conclusões e projetos futuros.

2 REFERENCIAL TEÓRICO

Neste capítulo serão apresentados os artigos que foram usados como fonte de pesquisa para a elaboração do presente trabalho. Esses artigos apresentam, entre outras informações, conceitos referentes ao uso de fluxos de dados para caracterização do tráfego. Esses conceitos são usados para a criação de modelos de detecção de anomalias. Ainda será apresentado o algoritmo genético, bem como pesquisas relacionadas ao seu uso.

2.1 Detecção de anomalias em redes

De acordo com Lee e Yun [12] *Network Intrusion Detection Systems* (NIDS) ou sistemas de detecção de intrusão baseados em redes, são componentes vitais para a infraestrutura de segurança de redes. São eles que monitoram o tráfego e alertam os administradores em caso de atividades suspeitas. Um NIDS normalmente é instalado na borda da rede e analisa cada pacote que entra nesta rede, procurando por assinaturas conhecidas de ataques. Uma assinatura é um conjunto de regras que identifica um ataque e deve ser constantemente atualizada.

Patcha e Park [13] classificam os sistemas de detecção de anomalias em redes de três formas: baseados em assinatura (*signature based*), baseados em perfil (*profile based*) e híbridos. Um sistema baseado em assinatura, assim como os sistemas de intrusão baseados em rede, é norteado por um conjunto predefinido de assinaturas (padrões) de ataques conhecidos, onde pacotes são inspecionados para verificar se correspondem aos padrões das assinaturas. Sistemas baseados em perfil buscam criar uma representação do que é considerado um comportamento normal, identificando como anômalo qualquer desvio desse perfil. O terceiro sistema é denominado híbrido e agrega características das duas técnicas descritas anteriormente.

As três técnicas têm suas vantagens e desvantagens. Os modelos que se baseiam em assinaturas de ataques conseguem identificar as investidas rapidamente tendo baixa taxa de falsos positivos. Em contrapartida, estes sistemas têm a desvantagem de lidarem apenas com anomalias que já tenham sido relatadas previamente. Ou seja, deve haver uma assinatura daquele tipo de ataque. Se esse ataque não foi descrito previamente, a assinatura para ele ainda não existe e este não será identificado. Os sistemas baseados na caracterização do tráfego têm como vantagem a habilidade de identificarem ataques desconhecidos até então, pois sua estratégia é retratar o comportamento normal de uma rede para identificar desvios de comportamento. Outra vantagem é a capacidade de criar perfis personalizados de acordo com as características da rede, e mesmo se essas características forem se alterando, esses sistemas se adaptam. Como desvantagem,

modelos baseados na caracterização do tráfego apresentam uma alta taxa de falsos positivos e uma alta complexidade.

O protocolo SNMP foi usado por muitos anos e tem sido gradualmente substituído pela análise dos fluxos IP, pois estes têm informações mais detalhadas sobre os processos a que se deseja serem analisados em uma rede. Com isso, a gerência de redes passa a ganhar mais itens a serem levados em consideração na análise de seu comportamento. Um fluxo é definido por uma sequência de dados unidirecional, passando por um mesmo ponto na rede durante um certo intervalo de tempo, que contém características em comum e que tem seu início e fim claramente definidos. Em um fluxo, podem-se identificar quais protocolos estão sendo mais usados e quanto de tráfego cada aplicativo ou usuário está consumindo [14].

No trabalho de Sperotto *et al.* [15] são mostradas algumas técnicas de detecção de intrusão baseadas em fluxos IP. Com o uso crescente das redes de computadores e o consequente aumento no tráfego dessas redes, aumentam as dificuldades dos administradores. Ainda segundo os autores, fluxos IP podem ser especialmente úteis em redes de alta velocidade, pois permitem que a rede seja analisada como um todo ao invés de ter que se analisar cada pacote individualmente, o que poderia ocasionar uma lentidão excessiva, devido à enorme quantidade de dados trafegados. Com isso, cada par de endereços IP pode agregar informações como o horário que a comunicação ocorreu, a quantidade de bytes transmitido e o número de pacotes enviados. Apesar de fluxos se limitarem às interações que ocorrem nas redes, o reconhecimento dos padrões dessas comunicações pode ser suficiente para identificar muitos tipos de ataques.

Lakhina *et al.* [1] usou a entropia como uma ferramenta de classificação para fluxos IP, usando dados como endereços IP e portas, habilitando assim uma classificação automática de tipos de anomalias usando um aprendizado não supervisionado. A entropia mede o grau de dispersão ou concentração dentro de uma distribuição de dados e foram identificados alguns tipos de ataques intencionais e não intencionais.

Ataques *Denial of Service* (DoS) ou em português, ataques de negação de serviço, são causados quando um atacante resolve enviar uma grande quantidade de solicitações a um serviço com o intuito de derrubá-lo. Essa grande quantidade de solicitações vai sobrecarregar o servidor que não vai conseguir responder a todas requisições, deixando estes recursos inacessíveis aos usuários legítimos. No trabalho de Rahmani *et al.* [16] é apresentado um modelo para identificar ataques *Distributed Denial of Service* (DDoS), em português ataques distribuídos de negação de serviço, que adicionam a dimensão de "um para vários" ao ataque DoS, ou seja, o ataque não é realizado por apenas uma pessoa, mas é coordenado por muitos atacantes diferentes contra apenas um sistema, visando sobrecarregar os serviços atacados mais rapidamente. Pela natureza desses ataques, torna-se um desafio distinguir o tráfego normal de um tráfego gerado por um atacante. Os

autores propõem uma abordagem dividida em duas fases, sendo que a primeira cuida da detecção de fluxos com base na evolução da distribuição no tamanho das conexões ao longo do tempo e a segunda fase faz a diferenciação com base na estimativa do grau de congestionamento. Essa técnica mostrou-se mais eficiente que as técnicas baseadas em entropia.

Uma extensa pesquisa sobre técnicas de detecção de anomalias é apresentada por Monowar *et al.* [17] e também por Ahmed *et al.* [4], onde são apresentados e comparados os modelos mais eficientes. Dentre os modelos apresentados, destacam-se os baseados na caracterização do tráfego, como o de Deljac *et al.* [18], que usa filtros bayesianos, um método estatístico que classifica dados por categorias, para detectar falhas em equipamentos de grandes redes, ou o trabalho de Ahmed e Mahmood [19] que usa um algoritmo de clusterização para classificação dos dados e consegue detectar ataques DoS. Estes modelos adaptam-se facilmente aos mais diversos tipos de redes, inclusive redes com alto tráfego, como backbones.

A técnica do uso de assinatura digital de segmento de rede, também referida por alguns autores como *baseline*, para fazer a caracterização do tráfego, tem sido utilizada pelo Grupo de Pesquisa em Redes de Computadores e Comunicação de Dados do Departamento de Computação da Universidade Estadual de Londrina (UEL) [20], tendo obtido resultados promissores. A seguir serão descritos trabalhos pertencentes a este grupo, do qual o algoritmo GADS e seu autor fazem parte.

A técnica de criação de uma assinatura digital da rede é discutida por Proença Jr. *et al.* em [21], onde a utilização de um *Digital Signature of Network Segment* (DSNS) serve como base para a criação da assinatura e consequente caracterização do tráfego. Neste trabalho os dados são obtidos através do protocolo SNMP e a assinatura digital é gerada baseada no tráfego apresentado por cada dia da semana individualmente.

Um DSNS também é usado no trabalho de Adaniya *et al.* [22] em conjunto com o algoritmo de clusterização *K-Harmonic Means* e a heurística *Firefly Algorithm*. Neste trabalho foram usados dados da UEL obtidos através do protocolo SNMP, apresentando como resultados taxa de detecção de 80% e falso positivo de 25%.

No trabalho de Zacaron *et al.* [23] a assinatura de rede é obtida através de fluxos IP e denominada *Digital Signature of Network Segment using Flow Analysis* (DSNSF). Em seu trabalho os dados foram coletados na UTFPR (Universidade Tecnológica Federal do Paraná) no Campus Toledo e foi usado o protocolo de exportação de fluxos Netflow. O modelo de clusterização *K-means* foi aplicado para a caracterização do tráfego e geração das assinaturas digitais, que descrevem o comportamento esperado para aquele segmento de rede.

Com o advento dos fluxos IP para a coleta de informações sobre o comportamento

da rede, outros trabalhos surgiram com o objetivo de caracterização do tráfego de rede para a busca de anomalias. No trabalho de Assis *et al.* [10] é proposta uma análise em sete dimensões, usando os atributos: bits/s, pacotes/s, fluxos/s, endereços IP de origem e destino e portas de origem e destino. Para a caracterização do tráfego foi usada uma versão aprimorada do tradicional método Holt-Winters, chamada de *Holt-Winters for Digital Signature* (HWDS), e através de limiares é feita a detecção de qualquer desvio do que a assinatura digital descreveu como comportamento normal. Se há um desvio de comportamento em ao menos quatro dos sete atributos, um alarme é gerado para o administrador da rede. O autor se baseia na utilização de intervalos de confiança (os limiares ou *thresholds*), descritos por Brutlag [24] que detectam comportamentos diferentes do esperado na rede analisada.

No trabalho de Fernandes Jr. *et al.* [11] um DSNSF é criado com base na Análise de Componentes Principais, a *Principal Component Analysis* (PCA) e é chamado de *Principal Component Analysis for Digital Signature* (PCADS), um método estatístico para redução de dimensionalidade. É usado pelo autor um *threshold* (limiar) nos atributos quantitativos (bits, pacotes e fluxos por segundo), para delimitar o tráfego normal do anômalo. Se identificada alguma anomalia, o administrador da rede é avisado e também são passados os atributos qualitativos (entropias de endereços IPs e portas de origem e destino).

O modelo *Autoregressive Integrated Moving Average* (ARIMA) é usado por Pena *et al.* [9] para a criação do DSNSF. O autor introduz o uso da Lógica Paraconsistente para otimizar o emprego do DSNSF, tentando simular o conhecimento de administradores de rede e como eles se beneficiam do conhecimento histórico dos segmentos para lidar com eventos danosos na infraestrutura de rede.

O algoritmo *Ant Colony Optimization* (ACO) é utilizado por Carvalho *et al.* [8] em conjunto com o modelo *Dynamic Time Warping* (DTW) para a detecção de anomalias em redes de computadores. Foi criado o modelo *Ant Colony Optimization for Digital Signature* (ACODS), que é usado no processo de clusterização dos dados adquiridos através de fluxos de dados, que vai gerar o DSNSF. Após esse processo de caracterização do tráfego, é proposta uma adaptação no modelo DTW, chamada *Adaptive Dynamic Time Warping* (ADTW), usada para reconhecimento de padrões, onde alinhando-se o DSNSF gerado com o tráfego real, é possível identificar padrões anômalos na rede.

Como nota-se pelo resultados dos trabalhos elaborados pelo grupo de pesquisas da UEL, a técnica do uso de DSNSF tem se mostrado eficiente na caracterização do tráfego de rede. A Tabela 1 resume os trabalhos deste grupo cujos dados foram adquiridos por meio das técnicas de fluxos de dados. Na tabela são apresentados os nomes dos modelos, qual algoritmo é usado, como é feita a detecção de anomalias, quantos atributos obtidos por meio dos fluxos de dados e qual o tempo usada para o treinamento do algoritmo, ou

Tabela 1 – Resumo dos trabalhos do grupo de pesquisa em redes da UEL

Modelo	**Algoritmo**	**Detecção de anomalia**	**Atributos**	**Tempo para treinamento**
ACODS	ACO	ADTW	7	4 semanas
HWDS	Holt Winters	Thresholds	7	1 dia
ARIMA	ARIMA	Lógica Paraconsistente	6	3 semanas
PCADS	PCA	Thresholds	7	5 semanas
DSNSF-KM	K-means	Não faz	3	7 semanas
GADS	GA	ADTW; Thresholds; Lógica Paraconsistente	6	3 semanas

seja, quantos dias são necessários para a geração do DSNSF.

Muitos algoritmos já foram usados na otimização da geração de assinaturas digitais de redes e outras técnicas empregadas na a detecção de anomalias nestas redes foram analisadas. O GADS utiliza-se de uma técnica de mineração de dados, chamada clusterização para obtenção do DSNSF. A mineração de dados é o processo de explorar grandes quantidades de dados procurando padrões ou relacionamentos entre os dados. Ela tem sido usada nos mais diversos meios, inclusive no processo de clusterização, que é o agrupamento não supervisionado de dados de acordo com algum grau de semelhança. Cada agrupamento é chamado *cluster*. Dados dentro de um *cluster* devem ter um alto grau de semelhança e dados entre esses devem ter pouca semelhança.

Segundo Dubes e Jain [25] as técnicas de clusterização tem o objetivo de organizar pontos em um espaço multimensional, de forma que todos os pontos em um mesmo grupo tenham alguma relação e pontos entre grupos diferentes tenham alguma diferença. Maulik e Bandyopadhyay [26] definem a clusterização como uma importante técnica de classificação não supervisionada, onde um conjunto de padrões, normalmente vetores em um espaço multimensional, são agrupados em *clusters* (grupos) de forma que em um mesmo *cluster* estes padrões sejam similares em algum sentido e padrões entre diferentes *clusters* sejam diferentes no mesmo sentido. Quem usa uma técnica de clusterização está tentando compreender um conjunto de dados e descobrir quaisquer estruturas que residam nestes dados. Técnicas de clusterização são ferramentas para descoberta, ao invés de apenas apresentarem uma solução final, portanto permitem aos

seus usuários formularem questões para futuros estudos.

Existem na literatura diferentes abordagens para técnicas de clusterização, cada uma com suas características, vantagens e desvantagens. Destacam-se entre essas técnicas, as que usam metodologias metaheurísticas como o algoritmo genético (GA) para a solução do problema.

O GA é um algoritmo proposto por Holland [27] [28] que simula a evolução natural das espécies, através de operadores como seleção, cruzamento (*crossover*) e mutação. O GA é conhecido por ser utilizado na otimização de resolução de problemas com uma grande variedade de soluções, como por exemplo a clusterização, onde várias técnicas podem ser aplicadas nesta solução.

No trabalho de Sheikh *et al.* [29] é apresentada uma extensa pesquisa das técnicas que usam GA no processo de clusterização. Em [30] os autores propõem uma técnica que seleciona os centros dos *clusters*, chamados centróides, a partir do conjunto de dados, e usa uma representação binária ao invés de números inteiros para representar o valor dos centróides.

No trabalho de Maulik e Bandyopadhyay [26] uma técnica de clusterização usando algoritmo genético é proposta e usa a distância Euclidiana como função objetivo para calcular em qual *cluster* cada dado deve ser colocado. No referido trabalho, os cromossomos são representados como números reais e são apresentados quatro exemplos artificias e três exemplos práticos para a técnica apresentada pelos autores. Os valores dos centróides são escolhidos de forma aleatória de acordo com os valores dos dados a serem clusterizados, portanto não usa representação binária. Com isso, os valores dos centróides são também os valores dos cromossomos do algoritmo genético. Essa técnica se mostrou bastante eficaz de acordo com as implementações feitas pelos autores e os resultados obtidos foram comparados com o algoritmo de clusterização *K-means*. A técnica usando GA foi altamente superior para a classificação dos dados.

Na seção a seguir será descrito o funcionamento do GA, bem como de seus operadores, que serão usados pelo GADS para otimizar o processo de clusterização para a obtenção da assinatura digital da rede, chamada de DSNSF. Essa assinatura será usada para detecção de anomalias naquele segmento de rede.

2.2 Algoritmo Genético

O GA foi desenvolvido por John Holland, seus colegas e alunos da Universidade de Michigan e seu objetivo era estudar os processos da natureza e desenvolver modelos computacionais que simulassem os mecanismos de adaptação natural. Esses algoritmos são especialmente eficazes na resolução de problemas complexos, onde não se tem muita informação a respeito do problema mas existem muitas soluções possíveis [27] [31] [32].

Ao contrário de outros modelos baseados em computação evolutiva, o objetivo principal do GA não era desenvolver algoritmos para resolver problemas específicos, mas ao invés disso, estudar formalmente os fenômenos da adaptação que ocorrem na natureza, e desenvolver caminhos nos quais os mecanismos de adaptação natural pudessem ser importados para os sistemas de computação e fossem usados para a resolução de qualquer problema matemático. GAs tem uma grande aplicação em áreas científicas, como otimização de soluções, aprendizado de máquina, análise de modelos econômicos, problemas de engenharia, entre outros [33].

Nos GAs, as variáveis que poderão ser usadas na resolução dos problemas são representadas como genes em um cromossomo. Este é designado como um indivíduo e o conjunto de indivíduos será a população. Como na natureza, indivíduos competem entre si dentro de uma população para que possam reproduzir-se e passar os seus genes adiante para a gerações futuras. Os mais adaptados ao meio são os que têm maior probabilidade de propagarem os seus genes para os seus filhos e consequentemente terão um número maior de descendentes nas populações futuras. O processo que classifica os indivíduos desta população como mais aptos ou menos aptos é chamado *seleção*.

O GA se baseia no princípio fundamental que o conjunto de soluções do problema proposto seja a *população*, onde cada solução isolada seria um *indivíduo*. Essa população irá evoluir a cada nova geração, e como os indivíduos mais adaptados deverão propagar os seus genes, espera-se, a cada nova geração, indivíduos mais aptos ao meio. Também deverão ser introduzidos periodicamente quaisquer mecanismos que alterem, a qualquer momento, soluções dentro da população. Isso irá garantir a diversidade desta população e também poderá produzir soluções inovadoras, que serão avaliadas no processo de seleção, e classificadas como boas soluções ou não.

De acordo com Holland [27], para que seja possível a aplicação do GA para a resolução de um problema, deve-se necessariamente seguir alguns preceitos:

- uma população inicial deve ser estabelecida. Normalmente esses indivíduos são criados aleatoriamente;
- deve-se estabelecer um critério para avaliar essa população, que vai definir o quão esses indivíduos estão adaptados ao meio. Esse critério é denominado função de aptidão ou *fitness*;
- o próximo passo é usar os dados obtidos através da função *fitness* para escolher quais são os melhores indivíduos desta população. Esse operador é chamado de seleção e é fundamental para a criação das novas gerações;
- para obtenção dessa nova geração, os indivíduos mais aptos e escolhidos para a reprodução deverão reproduzir-se e gerar filhos. É o processo de cruzamento

(*crossover*), que é vital, pois nesse processo ocorre a troca de genes entre os pais, o que vai garantir a diversidade da população;

- existe uma probabilidade de que um indivíduo sofra uma alteração em seus genes durante esse processo, que é chamado de mutação, e que vai garantir que novas soluções possam ser encontradas. Se ocorrer a mutação, esse indivíduo será colocado dentro da nova população e será avaliado pelo processo de seleção, onde poderá ser descartado ou usado para a reprodução nas gerações futuras;

- todos esses passos podem ser repetidos por um número fixo de vezes, ou pode ser estabelecido um critério de parada. Atingidos um ou outro, chega-se ao fim o processo e espera-se obter os melhores indivíduos das melhores gerações.

Esse ciclo é demonstrado na Figura 1 e será descrito em seus detalhes nas próximas subseções.

2.2.1 População

A população é o conjunto de indivíduos que representam cada uma das soluções possíveis para o problema em questão. A cada nova geração, a população aumenta de tamanho, e para que uma população muito grande não prejudique o desempenho do algoritmo, devem ser previstos mecanismos para a prevenção desse aumento excessivo. Ao mesmo tempo, uma população pequena reduz o espaço de busca, podendo resultar em uma solução final não ótima, seja pela pouca variedade de soluções, seja pela rápida convergência em uma solução que se acredite ser a ideal.

A população inicial é gerada aleatoriamente, baseada no espaço de busca de soluções. Esse é o primeiro passo do GA, pois a partir dessa população inicial serão aplicados todos outros operadores e serão geradas as próximas populações.

Uma população é formada por indivíduos distintos. Cada um desses indivíduos representa uma solução e serão esses indivíduos que irão ser submetidos aos processos de seleção, *crossover* e mutação em busca dos melhores, também conhecidos como mais aptos ou mais adaptados ao meio em que vivem.

Assim como na natureza, indivíduos são formados por genes. Essa representação genética deve ser planejada de forma muito cuidadosa, pois são os genes que irão configurar as soluções para os problemas propostos. A representação dos genes de um indivíduo pode ser feita usando-se a representação binária, ou ainda é possível usar qualquer outra representação, como letras, números reais ou códigos [27].

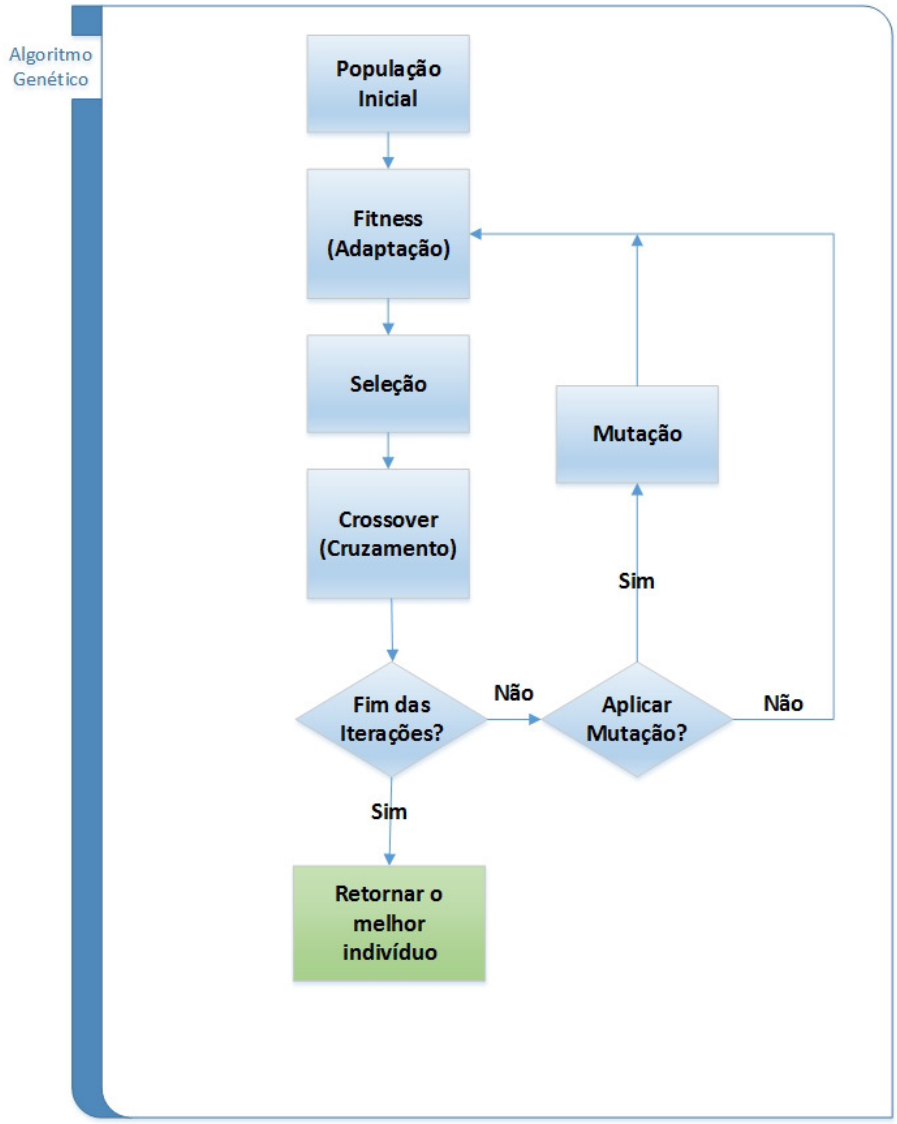

Figura 1 – Ciclo do Algoritmo Genético

2.2.2 *Fitness* (Aptidão)

Para que o processo de escolha dos indivíduos mais adaptados possa ser executado com perfeição é preciso saber com exatidão quais os pré-requisitos que um indivíduo deve ter para ser considerado apto ou inapto àquele meio que está sendo submetido.

A escolha correta da chamada *função objetivo* é um processo chave para o sucesso do algoritmo. Essa é a função que será otimizada, fornecendo meios para se chegar à resolução do problema. Para se manter a uniformidade sobre outros domínios, pode-se normalizar o valor da função objetivo para se obter um valor conveniente, conforme a equação 2.1 [34].

$$F_{(x)} = \frac{1}{f_{(x)}} \tag{2.1}$$

A função resultante $F_{(x)}$ será a *função fitness* e seu valor estará compreendido entre $0 < F_{(x)} < 1$.

Essa função *fitness* vai definir qual o comportamento de cada indivíduo na população frente a resolução do problema proposto, atribuindo a cada indivíduo um valor que corresponde ao valor do *fitness* deste. Esse valor é a representação do quão bem este indivíduo soluciona o problema, ou em outras palavras, o quanto este indivíduo está adaptado ao meio em que ele vive [33].

2.2.3 Seleção

Na natureza, o processo em que os indivíduos mais adaptados ao meio são definidos como os que tem maior probabilidade de sobrevivência, e consequentemente maior chance de se reproduzirem e passarem os seus genes adiante para as próximas gerações é chamado seleção natural.

No GA o processo de seleção vai escolher os indivíduos com um melhor valor de *fitness*, ou seja, os que estão mais próximos da solução. Como a função *fitness* já atribuiu um valor para cada indivíduo, resta fazer a escolha dos que irão se reproduzir e propagar os seus genes para as próximas populações.

Para fazer essa escolha, pode-se recorrer a um dos muitos métodos de seleção existentes. A seguir serão descritos os métodos utilizados na confecção do GADS, que em conjunto foram escolhidos para que pudessem extrair o melhor aproveitamento de cada um.

2.2.3.1 Método da Roleta *(Roulette Wheel)*

Esse método usa a proporcionalidade relativa ao *fitness* para atribuir a probabilidade que um indivíduo terá em se reproduzir, ou seja, quanto mais ótimo o valor de *fitness*, maior a probabilidade do indivíduo ser escolhido para a reprodução [33].

O método consiste em dar para cada indivíduo um valor que corresponda a uma fatia de uma roda circular conforme mostra a Figura 2. Esse valor é percentualmente proporcional ao valor de *fitness* de toda a população. Indivíduos com um maior valor de *fitness* têm maior probabilidade de serem os escolhidos.

Após a atribuição das porcentagens a que cada indivíduo tem direito na roleta, esta é "girada" o mesmo número de vezes da quantidade de indivíduos nesta população. Cada vez que o roleta é girada, um indivíduo é selecionado.

Usando-se como exemplo a Figura 2, existem quatro "fatias", sendo uma para cada indivíduo. Como são quatro indivíduos, essa roleta seria girada por quatro vezes. Como no gráfico o *Indivíduo 3* tem uma maior área, a probabilidade de a roleta "escolher" este indivíduo é maior. Como a porcentagem indicada no gráfico é

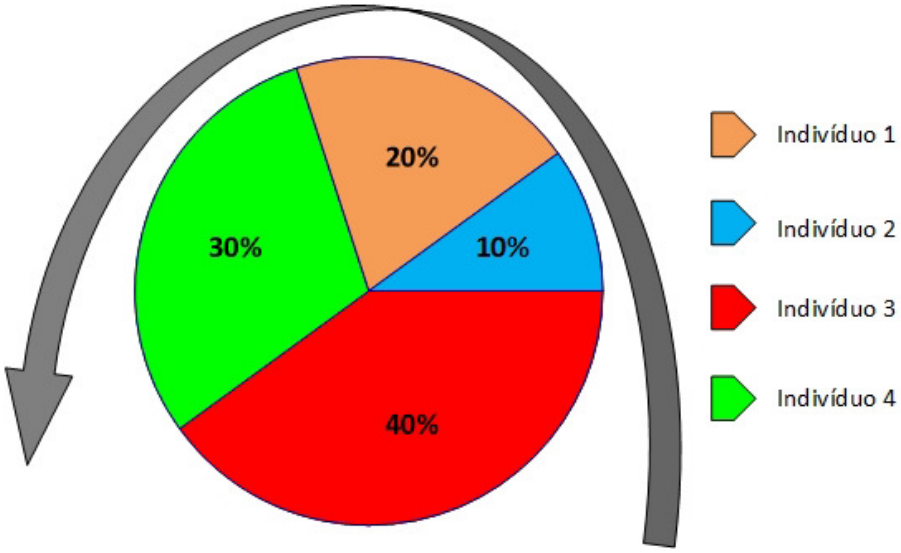

Figura 2 – Método da Roleta

proporcional ao *fitness* do indivíduo, ou seja, o quanto esse indivíduo está adaptado ao meio, percebe-se que os indivíduos mais adaptados deverão ser os mais escolhidos, e por consequência se reproduzirem mais vezes.

O método da roleta proporciona que indivíduos considerados mais adaptados se reproduzam mais vezes, fazendo com que seus cromossomos sejam replicados em seus descendentes, gerando populações cada vez mais adaptadas, consequentemente gerando soluções mais próximas das ideais. Como desvantagem, por se tratar de um método que faz com que indivíduos sejam escolhidos aleatoriamente, é possível (apesar de pouco provável), que indivíduos com valores de *fitness* baixos sejam sempre os escolhidos para a reprodução, fazendo com que soluções pouco ideais sejam geradas como descendentes. Para minimizar esse problema, o método da roleta pode ser implantado junto com algum outro método, como o elitismo.

2.2.3.2 Elitismo

Esta técnica foi introduzida por De Jong [35] e consiste em guardar uma parte das melhores soluções encontradas, ou seja, dos indivíduos mais aptos de cada geração, e introduzi-los nas gerações subsequentes. Esses indivíduos poderiam ser perdidos se não fossem selecionados para a reprodução ou se seus descendentes fossem escolhidos para a mutação. Segundo Mitchel [33], muitos pesquisadores concordam que o elitismo melhora significativamente a performance do GA.

2.2.4 Cruzamento (*Crossover*)

O operador encarregado do cruzamento, chamado de *crossover*, é responsável por fazer a combinação genética entre dois indivíduos, os pais, resultando dois novos

cromossomos, ou dois novos filhos para a população seguinte. Espera-se, em princípio, que esses dois indivíduos sejam mais aptos que os seus pais, melhorando o *fitness* da população [31].

Cada indivíduo em uma população é um ponto em um espaço de buscas. Tradicionalmente, indivíduos em um GA são representados por uma sequência de valores de tamanho n. O crossover opera em nível individual. Em um GA simples, escolhe-se aleatoriamente apenas um ponto de *crossover*, que será usado nos dois indivíduos que irão se reproduzir. Este ponto vai dividir a sequência de bits em duas partes, uma à esquerda do ponto, outra à direita, em cada um dos pais. O *crossover* então vai trocar as porções correspondentes em cada indivíduo para formar dois novos, com características genéticas distintas [36].

Como exemplo pode-se observar a Figura 3, onde os indivíduos *1* e *2* vão se reproduzir e gerar dois filhos. Os cromossomos na figura são representados por bits binários: zeros e uns. O ponto de crossover foi escolhido para o quarto bit, portanto cada novo filho deve receber metade da carga genética de cada um dos pais.

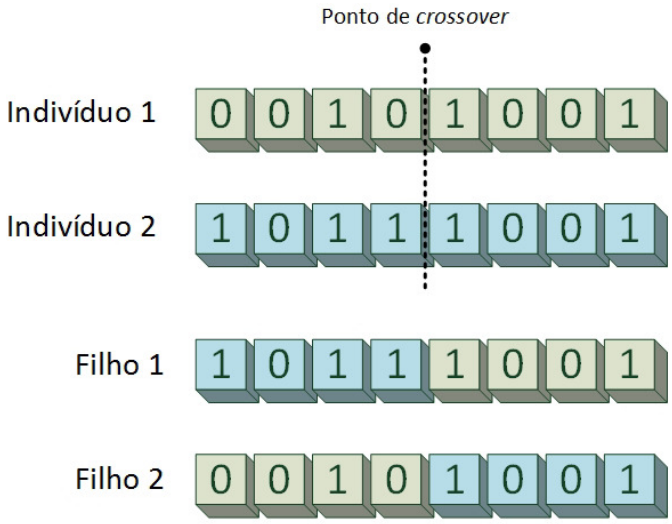

Figura 3 – *Crossover* em um único ponto

O cruzamento vai gerar novos indivíduos que serão incorporados à população em busca das soluções ideais. Além disso, assim como na natureza, o *crossover* nem sempre é aplicado. Após a escolha dos indivíduos selecionados para reprodução pelos métodos de seleção descritos anteriormente, o operador *crossover* só é executado se um número randômico, gerado no intervalo 0 a 1, for maior que a probabilidade p_c, ou seja, esta taxa p_c, chamada taxa de *crossover*, é a probabilidade de ocorrer o cruzamento [34].

2.2.5 Mutação

O operador da mutação é o responsável por introduzir um grau a mais de diversidade genética na população. Ele permite introduzir indivíduos com características

distintas à uma população, na intenção de encontrar soluções que pelos métodos tradicionais não seriam possíveis. Isso garante que o espaço de busca em um GA tenha uma grande probabilidade de explorar todas as soluções possíveis, mesmo as soluções menos prováveis. De acordo com Srinivas e Patnaik [37], a mutação no GA permite recuperar soluções perdidas ou inexploradas, prevenindo convergências prematuras ou soluções abaixo do ideal.

A mutação ocorre isoladamente em um indivíduo, através da alteração de alguma característica genética nos cromossomos deste indivíduo. Ela também é determinada por uma probabilidade de ocorrência, em geral baixa. Normalmente, após o *crossover*, cada indivíduo é submetido à probabilidade de ocorrer a mutação [38].

Quando o indivíduo é escolhido para a mutação será escolhido aleatoriamente neste individuo o ponto que será alterado, e será feita a alteração de acordo com o espaço de buscas de soluções. Por exemplo, se a população conta com indivíduos formados por cromossomos binários, a alteração de apenas um bit de 0 para 1 já torna esse indivíduo completamente diferente de seus progenitores. A Figura 4 exemplifica o processo de mutação alterando o valor do quinto bit de 0 para 1.

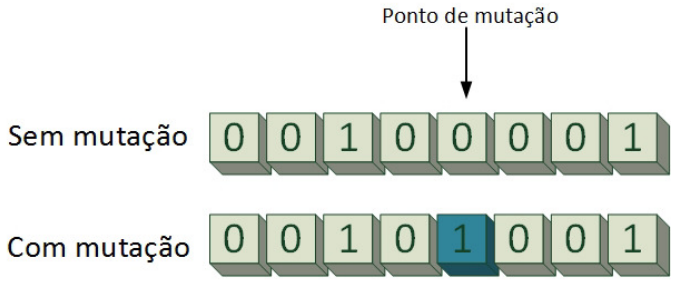

Figura 4 – Mutação ocorrendo em um indivíduo

2.2.6 Critérios de parada

Com a aplicação de todos os operadores já abordados, o algoritmo irá selecionar os indivíduos mais adaptados segundo os critérios definidos pelo programador. Estes indivíduos selecionados para a reprodução serão submetidos ao *crossover*, e após ele será aplicada a mutação nos que forem escolhidos de acordo com a probabilidade. Esse processo irá gerar novas populações que substituirão as antigas por completo ou não, também de acordo com os critérios do programador. Essas novas populações serão responsáveis por gerar outras sucessivamente até que seja atingido um critério de parada.

Esse critério pode ser definido como um número máximo de iterações do algoritmo por exemplo, especificando quantas gerações serão produzidas a partir da população inicial. Também pode ser especificado um valor de *fitness* que quando alcançado determina que a solução já foi encontrada. Outro critério pode ser se todas as

soluções em uma população atingirem um certo grau de homogeneidade (se um grande número de valores de cromossomos têm as posições idênticas) [34].

Os critérios de parada dependem diretamente do propósito do algoritmo genético e vão influenciar tanto no tempo e custo computacional do algoritmo, como na eficácia da solução encontrada.

2.3 Conclusão do capítulo

Seguindo a linha dos trabalhos pesquisados, em especial os trabalhos que usam a técnica de DSNSF, pretende-se criar um sistema de detecção de anomalias em redes, usando os fluxos de dados para a aquisição dos dados a serem analisados, pois estes serão clusterizados e na otimização do processo de clusterização será usado o GA. Espera-se com essa abordagem obter uma caracterização do tráfego mais precisa, e como consequência uma alta taxa de sucesso na deteção de anomalias.

3 SISTEMA DE DETECÇÃO DE ANOMALIAS GADS

Caracterizar o tráfego de rede é um dos grandes desafios de ferramentas e sistemas de gerência de redes, que tem o objetivo de detectar anomalias. A criação de um perfil que descreva o comportamento esperado de uma rede é de extrema importância para que se possa definir, como anômalo ou normal, o tráfego desse segmento.

Neste capítulo será descrito o algoritmo usado para a construção da assinatura digital da rede, denominada *Digital Signature of Network Segment using Flow Analysis* (DSNSF), que é a definição do que pode ser considerado o tráfego normal da rede, e é a base para o sistema de detecção. Para a criação do DSNSF será usado o *Genetic Algorithm* (GA), em português algoritmo genético, uma metaheurística capaz de otimizar processos de resolução de problemas que tem uma grande diversidade de soluções. Também serão apresentadas as técnicas usadas para a detecção de anomalias e as métricas usadas para a validação dos resultados

3.1 A caracterização do tráfego com GADS

Organizar dados em clusters ou grupos, requer o uso de alguma técnica para que os dados organizados sejam distribuídos corretamente entre os clusters. Quanto mais homogêneos os dados dentro de um cluster, e mais heterogêneos os dados entre os clusters estiverem, mais eficiente foi o processo de divisão dos dados entre os clusters.

O objetivo do GA neste trabalho é otimizar o processo de clusterização de dados. Os dados em análise serão referentes a semanas anteriores ao dia para o qual será gerada a assinatura digital, cuja quantidade de semanas será abordada em uma das seções subsequentes. Por exemplo, para uma segunda-feira, serão usados dados de segundas-feiras anteriores. O resultado final desse processo será a assinatura digital do segmento de rede, chamada DSNSF, cujo propósito é a caracterização do tráfego e será usada para a detecção de anomalias neste segmento de rede.

No presente trabalho é apresentada uma proposta para o uso do GA durante a clusterização de um conjunto de dados, cujo resultado será a criação do DSNSF, que é a assinatura digital do perfil da rede, descrevendo o comportamento padrão normal de tráfego daquele segmento. A Figura 5 ilustra todas as atividades do sistema GADS, como a geração do DSNSF e a detecção de anomalias.

A janela de tempo usada para fazer a caracterização do tráfego foi de um minuto. Devido aos altos índices de dados trafegados nas redes atuais, este tempo mostrou ser suficiente para que após o alerta, o administrador da rede possa tomar alguma atitude para conter o problema.

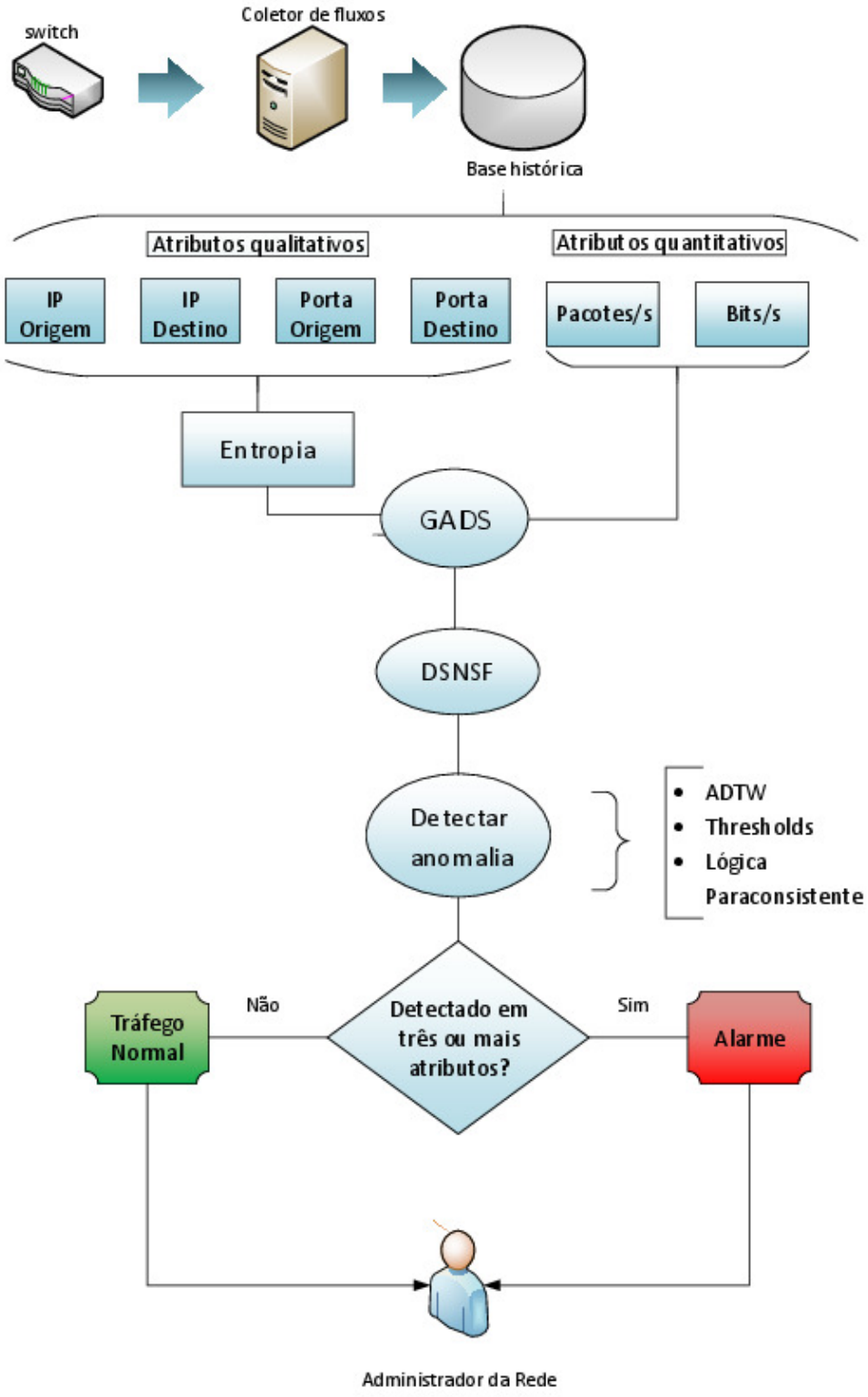

Figura 5 – Diagrama geral do GADS

Após essa etapa da caracterização do tráfego é apresentada a abordagem para detecção de anomalias, onde serão confrontadas a análise do comportamento normal, com o comportamento real da rede em busca de divergências que indiquem um comportamento

anômalo.

3.1.1 Aquisição e tratamento dos dados

A primeira etapa do GADS é a aquisição dos dados da rede, que serão usados para a criação dos DSNSF's. Esses dados foram coletadas da UEL e da UTFPR campus Toledo para os três cenários que serão usados nos testes.

Esses dados foram escolhidos para que possam ser comparados ambientes heterogêneos, avaliando a capacidade do sistema GADS em se adaptar a diferentes tipos de redes, bem como comparando o GADS com outra técnica similar. Mais detalhes sobre os cenários escolhidos serão dados no capítulo subsequente.

Para que todo tráfego das universidades seja avaliado, os dados são coletados nos *switches* centrais das redes e em seguida encaminhados a outro dispositivo chamado coletor. O coletor recebe o dados dos fluxos ainda em formato binário, e guarda as informações sobre o tráfego da rede de cada minuto em arquivos separados, um para cada minuto. A ferramenta *nfdump* [39] é usada para a transformação desses dados de fluxos em arquivos formato texto.

Após essa transformação, as informações ainda são reagrupadas em forma de arquivos, sendo estes ordenados por dia para que possam ser processados pelo GADS. Cada um dos seis atributos em análise vai gerar um arquivo texto diferente, onde cada linha nesse arquivo terá o valor do dado em questão para um segundo, totalizando 86400 linhas no arquivo. Por exemplo, para o atributo número de pacotes trafegados por segundo, será gerado um arquivo correspondente ao dia em questão, e este arquivo vai conter 86400 linhas, onde cada linha corresponde à soma do número de pacotes que trafegou naquela rede naquele segundo.

3.1.2 População do GADS

Todo algoritmo genético necessita de uma população para que seja dado início aos operadores de seleção, *crossover* e mutação. Essa população é formada por indivíduos que irão se reproduzir fazendo-a crescer a cada nova geração. Esses indivíduos são formados por cromossomos, que contém a representação genética (os genes) desse indivíduo. Como o GA irá otimizar o processo de clusterização dos dados, os genes dos indivíduos serão representados pelo valor dos centróides dos clusters, portanto o número de genes é diretamente proporcional ao número de clusters. O centróide é o centro do cluster, ou seja, é o elemento que representa melhor aquele conjunto de dados.

Após a definição dos centróides, é feita a distribuição dos elementos dentro dos clusters, de acordo com a distância do elemento até o seu centróide, usando-se a distância Euclidiana, fazendo com que cada conjunto contenha dados mais homogêneos.

O método empregado para validação e interpretação da quantidade de clusters usado neste trabalho é o método de Silhouette [40], que proporciona uma representação gráfica de como os elementos estão organizados dentro de cada cluster. A Figura 6 apresenta diferentes tipos de gráficos para diferentes números de clusters. Esta figura apresenta cinco testes, cada um com um número de clusters. Em cada teste, observa-se no eixo y o número total de clusters K e no eixo x encontram-se os valores da função Silhouette $f(s)$, para cada elemento i do cluster. Essa função retorna um valor compreendido entre $-1 \leq f(s) \leq 1$. Quando o valor de $f(s)$ está próximo de 1, indica que há uma alta semelhança entre esse elemento i e os outros elementos pertencentes ao cluster. Se este valor estiver próximo de zero, indica que aquele elemento poderia estar atribuído a mais de um cluster. A pior situação ocorre quando $f(s)$ apresentar um valor próximo de -1, o que indica que aquele elemento foi classificado erroneamente, e por consequência deveria estar atribuído a outro cluster. Nota-se na figura, na cor verde, a melhor situação acontecendo quando são usados quatro clusters, onde não são encontrados valores negativos, nem valores zero.

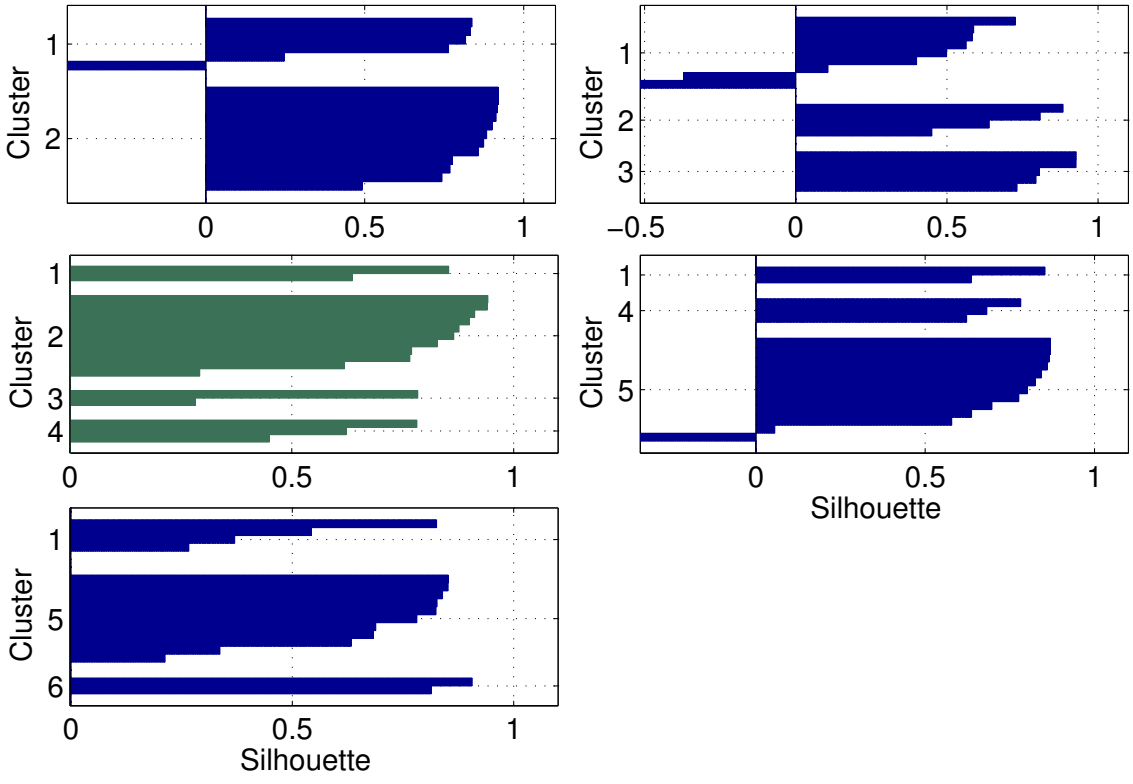

Figura 6 – Gráfico de Silhouette

A população inicial P_i é gerada de forma aleatória, porém o número de indivíduos da população inicial deve ser escolhido com cuidado. Os valores dos genes serão compreendidos entre 0 e o maior valor encontrado dentre os dados que serão clusterizados. Segundo [31], estudos empíricos definem que o tamanho da população inicial esteja compreendido entre 30 e 100 indivíduos. Foram realizados testes para

valores de P_i para o GADS, levando em consideração o custo computacional e a eficácia do sistema. É interessante notar que como a população cresce exponencialmente, deve-se haver uma preocupação com a população inicial, para que um número total de indivíduos muito baixo não deixe o espaço de busca reduzido, e um número muito alto não aumente consideravelmente o tempo para a geração da assinatura digital.

Com base nestas informações, foram realizados testes com valores para P_i entre 10 e 100 indivíduos. Foram executados os mesmos testes da seção 3.2, o erro quadrático médio normalizado (NMSE) e o Coeficiente de Correlação (CC), com base nos DSNSFs gerados em um dia para cada um dos valores de P_i. Os valores para NMSE e CC encontram-se na Tabela 2. Pode-se notar nesta tabela valores muito próximos para os resultados, por isso foi calculada a média para ambos os valores e foram destacados os valores mais próximos dessa média. Os valores de P_i para 40 e 60 indivíduos foram os mais próximos dos valores ideais. Como a diferença entre ambos é mínima, para preservar a complexidade computacional do sistema foi optado pelo número $P_i = 40$.

Tabela 2 – Valores de NMSE e CC para P_i

P_i	NMSE	CC
10	0.0435	0.9226
20	0.0455	0.9190
30	0.0440	0.9201
40	0.0446	0.9203
50	0.0445	0.9195
60	0.0446	0.9203
70	0.0447	0.9220
80	0.0449	0.9188
90	0.0442	0.9211
100	0.0443	0.9203
Média	**0.0445**	**0.9204**

Devido ao crescimento exponencial da população, a cada geração vão ser implementados mecanismos que descartem os indivíduos com valores menores de *fitness*, para que o tamanho total da população não acabe por prejudicar o desempenho do algoritmo.

3.1.3 Número de semanas para treinamento

Para a geração do DSNSF são usados dados das últimas três semanas. Para a aferição do número de semanas foram feitas análises de NMSE e CC para a geração de um DSNSF usando os atributos quantitativos bits e pacotes, com números de semanas entre uma e dez semanas de histórico de dados.

Os resultados para o Coeficiente de Correlação para bits e pacotes são apresentados nas Figuras 7 e 8, onde é possível notar números muito próximos para todos os períodos testados, tanto para bits quanto para pacotes, sendo a diferença da CC para bits de apenas 0.034 entre o maior e o menor valor, e para pacotes de apenas 0.039.

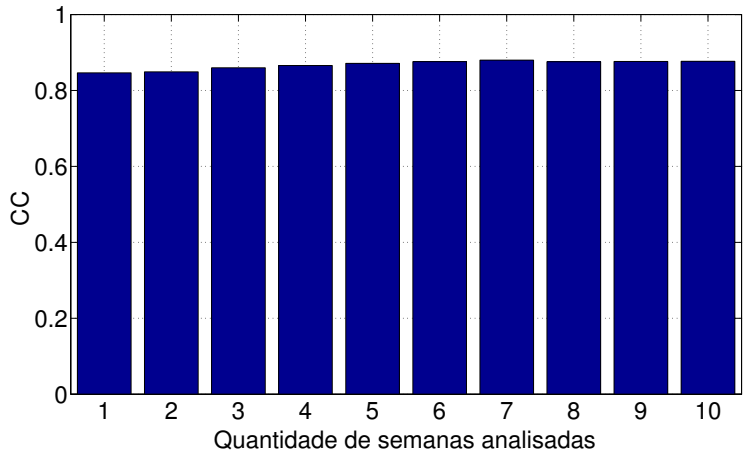

Figura 7 – CC para bits, para número de semanas no histórico de dados

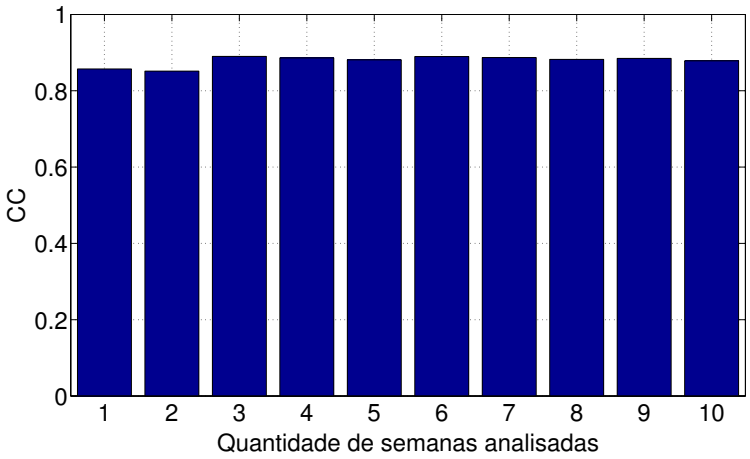

Figura 8 – CC para pacotes, para número de semanas no histórico de dados

Os resultados para NMSE também para bits e pacotes são apresentados nas Figuras 9 e 10, onde notam-se os melhores valores, ou seja, os mais próximos de zero, para o número de três semanas. Com esse resultado, será usado o número de três semanas para a base de dados histórica dos dias que irão compor o DSNSF.

3.1.4 Função *fitness*

No GA, a função *fitness* é a fórmula que vai mensurar o quanto um indivíduo está adaptado ao meio em que vive. Sabe-se que o escopo deste trabalho é fazer a classificação

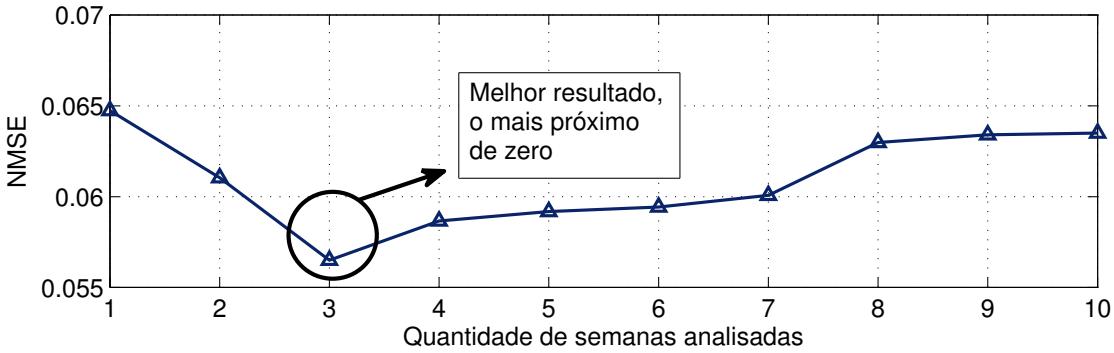

Figura 9 – NMSE para bits, para número de semanas no histórico de dados

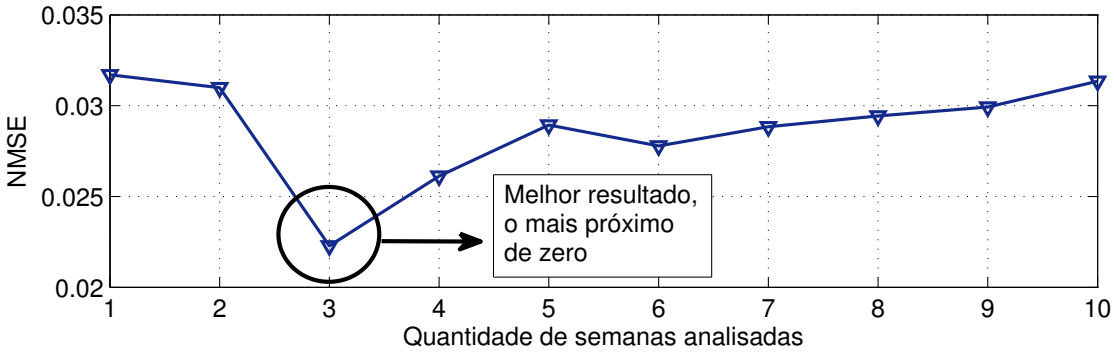

Figura 10 – NMSE para pacotes, para número de semanas no histórico de dados

dos dados obtidos através das técnicas de fluxos IP, usando o GA neste processo de clusterização.

Como método para definir o quanto os dados estão bem organizados dentro de um cluster, optou-se por fazer a soma total das distâncias entre cada ponto e seu respectivo centróide. Para tanto foi usada a distância Euclidiana, que é dada por:

$$J = \sum_{i=1}^{E} \sum_{j=1}^{K} \sqrt{\sum_{n=1}^{N} (x_{in} - c_{jn})^2} \tag{3.1}$$

onde K é o número total de clusters, E representa a quantidade de fluxos para serem clusterizados, e N indica as dimensões que serão dadas, por exemplo quantos atributos dados pelos fluxos serão usadas (bits, pacotes, etc). A variável x_{in} indica o valor dos atributos n dos fluxos i e c_{jn} guarda o valor do centro do cluster j na dimensão n.

Fazendo-se a soma das distâncias entre cada elemento de um cluster e seu centro, é possível aferir se dados de um cluster estão mais bem organizados que outro, fazendo a comparação dos valores de ambos, pois se os dados estão muito dispersos, somando-se a distância entre cada elemento e seu centróide o valor será mais alto do que um cluster que tem seus elementos próximos ao centro do cluster.

Com essas métricas, nota-se que quanto menores os valores encontrados, melhor os dados estarão organizados dentro de um cluster, pois os dados neste cluster vão estar mais próximos de seus centros. Os valores obtidos pela função *fitness* serão essenciais para o próximo operador, a seleção.

3.1.5 Seleção

Após a aplicação da função *fitness* para cada indivíduo, é necessário escolher aqueles que estão mais bem adaptados para que possam se reproduzir. A escolha é feita com base nos valores absolutos de *fitness*, onde quanto menor o valor, melhor organizados os dados estão naquele cluster.

Foram escolhidos dois métodos de seleção para o uso neste trabalho, que foram o método da roleta e o elitismo. A combinação destes dois métodos foi feita com base nas características de cada um deles, para que o melhor de cada um fosse extraído. A escolha do método da roleta se deu pela sua característica de percorrer todos os indivíduos avaliando todos os valores de *fitness*. Como o objetivo do GA é encontrar clusters onde os indivíduos estejam mais bem organizados, não deve-se correr o risco de descartar indivíduos mais adaptados. Por essa razão também foi implantado o método do elitismo, para que os indivíduos mais aptos não fossem descartados.

A roleta vai percorrer todo o espaço de busca daquela população e escolher os indivíduos mais adaptados, que estarão aptos à reprodução, através do operador de *crossover*.

3.1.6 *Crossover*

É na reprodução dos indivíduos de uma população que ocorre a diversificação da espécie, onde dois pais vão passar suas características genéticas à sua prole. Após a escolha pelos métodos da roleta e elitismo, dois pais são submetidos ao operador de *crossover*, onde haverá a troca de genes entre o par, para a geração de dois filhos.

Como cada cromossomo vai ter a mesma quantidade de genes quanto for a quantidade de clusters, o ponto de *crossover* será escohido no intervalo $P_c = 1 < P_c < j - 1$, onde j é o numero de clusters e P_c é o ponto de *crossover* que estará compreendido entre 1 e o número de clusters menos 1.

O cruzamento entre os indivíduos vai permitir a diversificação genética da população. Como a busca do GA é pelo cluster mais bem organizado, no momento da troca dos genes será gerado um indivíduo com valores diferentes para cada centro de cada cluster. A partir desses valores de centróides, para a próxima geração será obtido um valor de *fitness* baseado na função *fitness*, que irá definir, baseado nos valores destes centróides, através da distância entre cada ponto no cluster e seu centro, se os dados

estão ou não bem organizados. Esse valor vai indicar se aquele indivíduo vai permanecer na população podendo se reproduzir, ou se vai ser descartado.

3.1.7 Mutação

O último operador a entrar em ação é a mutação. Ela ocorre em taxas bem pequenas, e visa introduzir novas possibilidades ao espaço de busca que podem ter sido perdidas em outras gerações, ou que não haviam sido geradas pelos outros operadores.

Em estudos anteriores as probabilidades de ocorrerem a mutação em um indivíduo tem se mantido entre 0.01 e 0.05 (o que equivalem a 1% e 5% respectivamente). Em [34] foram encontradas taxas de mutação ideais entre 0.01 e 0.05 para populações entre 30 e 200 indivíduos. Já em [41], foi usada uma taxa de mutação de 0.01 para um GA utilizado na clusterização. Para o mesmo propósito, [42] usou uma taxa de mutação e 0.01 para uma população de tamanho fixo de 800 indivíduos. Em [43], também foi usada a taxa de 0.05 em um GA para o mesmo processo de organizar dados em clusters.

Diante dos resultados obtidos na literatura estudada, foi definido que a mutação vai ocorrer com uma probabilidade de 5%, que mostrou ser um valor que não vai destruir possíveis boas soluções por ser muito alto, nem deixar de auxiliar a diversidade das soluções por ser muito baixo. Além disso, o propósito deste trabalho é usar o GA e seus operadores para gerar uma assinatura digital da rede, e não fazer um estudo aprofundado nas melhores taxas de mutação. Caso o indivíduo seja escolhido para a mutação, o valor do gene é alterado e ele é reinserido na população. Para isso será gerado um valor usando a Fórmula 3.2

$$New_i = Old_i + (\theta * Old_i) \tag{3.2}$$

onde o novo valor é chamado New_i, o velho valor é chamado Old_i e θ é o número randômico gerado no intervalo $0 < n < 1$ que determinou se a mutação ocorreu ou não. Este novo valor de gene será usado nas próximas gerações.

3.1.8 Critério de parada

Outro fator a ser levado em consideração é o número de iterações γ do algoritmo, ou seja, quantas gerações de filhos o algoritmo vai explorar em busca de soluções. Em [26], o número máximo de 100 iterações é considerado como ideal, por não causar um excessivo custo computacional ao algoritmo. Já no trabalho de Ribeiro *et al.* [44], o número máximo de iterações foi definido em 50. Para definir o melhor número de iterações, neste trabalho foram feitos testes com números entre 10 e 100, e depois foram geradas as taxas de NMSE e CC para cada um dos valores. A Tabela 3 apresenta esses valores.

Tabela 3 – NMSE e CC para γ no GADS

γ	NMSE	CC
10	0.0457	0.9177
20	0.0445	0.9202
30	0.0459	0.9180
40	0.0463	0.9173
50	0.0448	0.9201
60	0.0450	0.9203
70	0.0458	0.9184
80	0.0464	0.9187
90	0.0447	0.9200
100	0.0450	0.9208
Média	**0.0454**	**0.9191**

É possível observar nos valores da tabela que houve pouca variação para ambas taxas. Em busca dos melhores resultados, foram destacados os valores mais próximos da média em cada cálculo, de acordo com a variância entre os valores máximos e mínimos de γ. Com base nesses resultados, e mais uma vez buscando valores que também levem em consideração o menor custo computacional, foi definido o valor de $\gamma = 50$ iterações.

O pseudo-código do algoritmo é apresentado no Algoritmo 1. Os cromossomos são vetores que contém os valores dos centróides de cada cluster. O tamanho do cromossomo é definido por $N * K$, onde N é o número de dimensões do nosso espaço de buscas e K é o número de clusters. Como foram definidos 4 clusters, os dados foram distribuídos entre esses quatro clusters, de acordo com a proximidade do centro do cluster. Cada valor de centróide será um gene no cromossomo.

Ao final de cada iteração será obtida a melhor população. Desta população será escolhido o melhor indivíduo, ou seja, aquele que obteve o menor valor de *fitness*, que representa que os dados dentro dos clusters estão bem organizados, com seus elementos próximos de seu centro, uma vez que esse valor é a soma das distâncias entre cada elemento e seu centróide.

Depois é calculada a média entre os quatro valores dos centróides. Esse número final vai representar um ponto no DSNSF, e deverão ser calculados 1440 pontos, uma vez que foi optado por representar cada minuto de um dia da semana no gráfico.

Ao final desse processo, como foram usados dados dos três dias da semana anteriores correspondentes ao dia analisado, será obtida a assinatura digital para esse dia ou o DSNSF. Essa assinatura é a representação da previsão do comportamento da rede para aqueles atributos analisados e será usada no processo de detecção de anomalias nesta rede.

Algoritmo 1 GADS para a criação do DSNSF

Entrada: Conjunto de dados (atributos) obtidos através de fluxos de dados (bits, pacotes, entropias, etc) através de dados históricos; número de clusters; número de iterações.
Saída: X: Vetor representando a quantidade de dados dos atributos, organizados em 1440 intervalos, um para cada minuto do dia. O DSNSF.

```
1: para: i = 1 até 1440 faça:
2:    Inicializando a população ρ
3:    para: t = 1 até γ, o número total de gerações faça:
4:       Computar o fitness para ρ
5:       Selecionar os mais adaptados através dos métodos da roleta e elitismo
6:       Aplicar crossover para gerar a população τ
7:       Aplicar a mutação se necessário
8:       Avaliar soluções através da função fitness (3.1)
9:       Guardar o melhor indivíduo daquela população
10:   fim-para:
11:   Calcular o centro de cada cluster da melhor solução encontrada (o mais adaptado)
12:   X_i ← média entre os centróides
13: fim-para:
14: Retornar: X
```

3.2 Métricas para avaliação

Os atributos foram divididos em dois grupos, de acordo com suas características. No grupo dos atributos quantitativos foram colocados os bits e pacotes por segundo, já que eles determinam a quantidade de tráfego da rede. Os outros atributos foram classificados como atributos qualitativos, por descreverem o comportamento deste tráfego, com as entropias de portas e endereços IPs de origem e destino.

Serão usadas duas métricas de avaliação para mensurar a qualidade das previsões do DSNSF para os atributos quantitativos. A primeira é o erro quadrático médio normalizado (NMSE) e a segunda é o Coeficiente de Correlação (CC). É importante ressaltar que alguns atributos qualitativos, quando analisados individualmente, podem não refletir de forma correta o comportamento da rede, mas são extremamente importantes quando analisados em conjunto com outros atributos. É o caso, por exemplo, da entropia de porta de origem, que pelo grande número de portas de comunicação TCP/IP (mais de 60 mil portas), os índices de CC seriam ruins pela dificuldade do DSNSF acompanhar o movimento do tráfego real. Por esse motivo, os atributos qualitativos serão avaliados em seus detalhes quando forem usados para a detecção de anomalias.

Para mensurar a eficiência do GADS na detecção de anomalias serão usadas algumas métricas de desempenho descritas por *Witten et al.* [45]. Para esta análise foram usados todos os seis atributos em conjunto. As métricas de avaliação de desempenho usadas serão a acurácia, o *recall*, a precisão, a especificidade e a *F-measure*. Também foi

usada a curva ROC (*Receiver Operating Characteristic*) para ilustrar o desempenho da detecção de anomalia.

Para o cálculo dessas métricas, foi elaborada a matriz de confusão, identificando todas as quatro possibilidades de comparação entre o DSNSF e o tráfego real. Na matriz de confusão são identificadas as quantidades de taxas de Verdadeiro Positivo (*True Positive, TP*), ou seja, todas as vezes que o tráfego foi classificado como anômalo e realmente apresentou uma anomalia. As taxas de Verdadeiro Negativo (*True Negative, TN*), quando o tráfego foi classificado como normal, e assim era, sem anomalias. As taxas de Falso positivo, (*False Positive, FP*), quando o tráfego foi identificado como anômalo e era normal. E as taxas de Falso negativo (*False negative, FN*), quando o tráfego foi identificado como normal e era anômalo. A matriz de confusão é apresentada de acordo com a Tabela 4.

Tabela 4 – Matriz de Confusão

		Previsão	
Tráfego Real		**Anomalia**	**Normal**
	Anomalia	TP	FN
	Normal	FP	TN

Todas as métricas de avaliação e desempenho usadas neste trabalho serão descritas a seguir.

3.2.1 Erro quadrático médio normalizado (NMSE)

O NMSE, do inglês *Normalized Mean Square Error*, tem o objetivo de fazer a validação da diferença entre os valores esperados e os que foram obtidos realmente [46]. Para os cenários propostos, onde a caracterização do tráfego é uma das tarefas principais, a análise de performance dessa caracterização tem como propósito mensurar o quão distante o DSNSF gerado está, na média normalizada, do tráfego real observado naquele período de tempo.

Para o cálculo do NMSE entre essas duas séries temporais, ou seja, o DSNSF e o tráfego real daquele dia, deve ser usada a seguinte equação:

$$NMSE = \frac{1}{n} \sum_{i=1}^{n} \frac{(D_i - T_i)^2}{\overline{DT}} \tag{3.3}$$

A equação 3.3 descreve o cálculo da NMSE, onde as duas variáveis discretas D e T representam os vetores que contém os valores dos DSNSF e do tráfego real, respectivamente. O i, em termos computacionais, representa um índice que varia de 1 até n, onde $i = 1, 2, 3, ..., n$.

Tabela 5 – Resultados para CC

CC	Significado
1	As variáveis são iguais (há correlação)
0	As variáveis são completamente diferentes (não há correlação)
-1	As variáveis são opostas (não há correlação)

Essa métrica possui o limite de 0 para seus resultados, ou seja, quanto mais próximos de 0 forem os resultados, indica que os valores esperados são mais próximos dos verificados. Da mesma forma, valores elevados para NMSE indicam resultados mais distantes dos que eram esperados.

3.2.2 Coeficiente de Correlação (CC)

O CC mensura a relação entre as tendências apresentadas pela assinatura e pelo movimento real do tráfego do dia comparados. Quanto mais o coeficiente de correlação se aproxima do valor 1, maior a correlação entre o DSNSF e o tráfego analisado, ou seja, maior a similaridade entre o tráfego previsto e o real. Na caracterização do tráfego, o CC indica se o DSNSF acompanha as tendências de crescimento e de diminuição do movimento gerado pelo tráfego real, mostrando se o método usado para a previsão deste movimento é eficiente ou não.

Para o cálculo do coeficiente de correlação entre duas variáveis quaisquer, usa-se a fórmula:

$$CC = \frac{\sum_{i=1}^{n}(D_i - \overline{D}) * (T_i - \overline{T})}{\sqrt{\sum_{i=1}^{n}(D_i - \overline{D})^2 * \sum_{i=1}^{n}(T_i - \overline{T})^2}} \tag{3.4}$$

Na equação 3.4, verifica-se como é feito o cálculo do coeficiente de correlação para as variáveis D e T, que assim como na equação 3.3, representam o DSNSF e o tráfego real, respectivamente.

Os resultados para o CC podem variar de -1 a 1, onde cada valor representa a análise das variáveis. Pode-se ver esses valores e seus significados na Tabela 5. Na caracterização do tráfego, espera-se que os valores obtidos para o DSNSF sejam o mais próximo possível dos valores reais do tráfego analisado, ou seja, espera-se que os valores de CC sejam o mais próximo possível de 1, o que indica que o tráfego previsto é igual ao observado.

3.2.3 Acurácia

A acurácia mede a porcentagem de amostras positivas e negativas (TP e TN), que foram classificadas de forma correta, em relação à soma de todas amostras negativas e positivas (TP, TN, FP, FN). Ela é expressa e acordo com a Equação 3.5:

$$Acurácia = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.5}$$

Essa métrica tem o objetivo de calcular a probabilidade do sistema gerar resultados corretos, ou seja, disparar um alarme quando naquele intervalo houver uma anomalia, e não disparar qualquer alarme quando o intervalo representar um tráfego normal.

3.2.4 *Recall*

O *recall* avalia a porcentagem de amostras positivas (TP) classificadas corretamente, sobre o total de amostras realmente positivas, ou seja, que eram realmente uma anomalia (TP, FN). O cálculo é dado pela Equação 3.6:

$$Recall = \frac{TP}{TP + FN} \tag{3.6}$$

3.2.5 Precisão

A precisão calcula a porcentagem de amostras positivas (TP) classificadas corretamente, sobre o total de amostras que foram classificadas como positivas, ou seja, que o DSNSF identificou como anomalia. Ela é dada pela Equação 3.7:

$$Precisão = \frac{TP}{TP + FP} \tag{3.7}$$

3.2.6 Especificidade

A especificidade calcula a porcentagem de amostras negativas identificadas corretamente (TN), sobre o total de amostras negativas (TN + FP), ou seja, as que não foram classificadas como anomalias e as que realmente não eram anomalia. É dada pela Equação 3.8:

$$Especificidade = \frac{TN}{TN + FP} \tag{3.8}$$

3.2.7 *F-measure*

A *F-measure* é a média entre a precisão e o *recall*. Ela é dada pela Equação

$$F - Measure = \frac{2 * (Precisão * Recall)}{Precisão + Recall} \tag{3.9}$$

3.2.8 Curva ROC

A curva ROC (*Receiver Operating Characteristic*) é uma ilustração do *trade-off* do sistema em avaliação. Ela caracteriza a relação entre as taxas de intervalos classificados corretamente como anômalos e a taxa de falsos positivos. Os dados serão representados tendo as taxas de verdadeiros positivos (TPR) no eixo vertical, e as taxas de falso positivo (FPR) no eixo horizontal [45]. Essas duas taxas serão dadas pelas Equacões 4.1 e 4.2:

$$TPR = \frac{TP}{TP + FN} \tag{3.10}$$

$$FPR = \frac{FP}{FP + TN} \tag{3.11}$$

O resultado será uma curva, cuja origem será o menor valor dessas duas métricas, e cujo destino será o maior desses valores. Assim, quanto mais a ROC se parecer com uma reta, pior será o desempenho do sistema, e quanto mais acentuada for a curva, quer dizer que o sistema de aproxima das melhores taxas de verdadeiros positivos.

3.3 Detecção de anomalias

Um dos grandes desafios em sistemas de detecção de anomalias é encontrar valores de *thresholds* (limiares entre o tráfego anômalo e o normal) que se adequem a diferentes situações. Esses sistemas não podem gerar muitos alarmes falso positivos, nem deixar de soar o alarme no momento de uma anomalia, gerando falsos negativos. Assim sendo, um dos intuitos dos vários testes realizados, é encontrar um valor de *threshold* que seja o mais eficiente possível.

Serão usados neste trabalho três técnicas de detecção de anomalias já propostas, e que serão comparadas na busca das mais eficientes para o modelo GADS proposto. As três serão descritas a seguir e serão avaliadas para cada cenário proposto.

3.3.1 ADTW

O *Adaptive Dynamic Time Warping* (ADTW) é uma adaptação da técnica DTW [47], desenvolvida inicialmente para reconhecimento de padrões de comportamento de vozes e usado na detecção de anomalias por Fernandes Jr. *et al.* [11]. Esta técnica foi usada no referido trabalho para comparar o DSNSF gerado pelo GADS com o tráfego real na busca de padrões de comportamento anômalos. A assinatura digital e o tráfego real são submetidos à análise do ADTW em busca de diferenças nos padrões de crescimento

ou diminuição da quantidade de tráfego, para que, se encontradas tendências desiguais, a anormalidade seja identificada e relatada.

Essa busca é feita de acordo com os coeficientes de significância σ (*thresholds*) definidos previamente. Na busca dos melhores valores para σ, foram feitos os mesmos testes de avaliação de desempenho e também a curva ROC, que através da identificação de valores de falsos positivos e verdadeiros positivos, mostrará as melhores taxas deste coeficiente.

3.3.2 Limites de confiança (*thresholds*)

O DSNSF descreve uma estimativa do comportamento do tráfego de rede de um determinado segmento. Portanto, é levado em consideração que sempre que o tráfego real apresentar algum desvio em relação à estimativa do DSNSF, esse intervalo de tempo deve ser considerado como anômalo. Contudo, a estimativa do DSNSF não é perfeita, podendo ocorrer pequenos desvios de comportamento, que não são anomalias, visto que o tráfego de rede pode oscilar ocasionando pequenos desvios, que não podem ser considerados alarmes. Então, para que sejam considerados alarmes apenas desvios mais críticos, e o administrador da rede não seja incomodado com uma excessiva quantidade de alarmes, são definidos limites aceitáveis entre o tráfego real e a assinatura digital, sendo considerados alarmes, apenas o tráfego real que superar (para cima ou para baixo) esses limites definidos.

Brutlag [24] descreve os limiares ou intervalos de confiança, no presente trabalho identificados como *thresholds*, como uma abordagem de sucesso na detecção de comportamentos anômalos em redes. Para isso são estabelecidos limites superiores e inferiores, e quando o tráfego real ultrapassa quaisquer desses limites, um alarme deve ser gerado.

3.3.3 Lógica Paraconsistente

De acordo com Costa *et al.* [48], a Lógica Paraconsistente (LP) originou-se de estudos que investigam a possibilidade de restringir ou rejeitar um dos fundamentos básicos da lógica clássica, o princípio da não contradição. De forma geral, esse princípio é pontuado pela afirmação que uma fórmula e sua negação não podem ser ambas verdadeiras, ou ainda, que dada uma proposição com uma conotação verdadeira, ela não pode ser falsa, ou ainda que uma conotação falsa não pode ser verdadeira. A LP é uma lógica não clássica, que aceita a contradição e que percorre as características evidenciais de uma determinada proposição.

A técnica de lógica paraconsistente aplicada na detecção de anomalias em redes foi usada no trabalho de *Pena et al.* [49], que a utiliza para decidir sobre a seguinte proposição, “A rede está comprometida por uma anomalia”. Para tanto, uma ferramenta

denominada Máquina Paraconsistente Correlacional (MPC) recebe como entrada os DSNSFs e as leituras reais do tráfego de rede e utiliza uma sequência de procedimentos para extrair informações evidenciais derivadas do contraste entre cada DSNSF e as leituras reais. Essas informações foram utilizadas em uma formalização das indefinições e contradições relativas ao comprometimento da rede. Anomalias foram identificadas quando todas as informações evidenciais apontaram para a mesma resposta (presença de uma anomalia no tráfego), com o mínimo de contradição.

3.4 Regras para alarmes

Outro grande desafio, após a geração dos alarmes individuais em cada um dos seis atributos dos fluxos IP, é definir quando enviar um alarme para o administrador da rede. Segundo Lakhina *et al.* [1], cada anomalia relatada afeta de uma forma diferente a distribuição dos atributos do tráfego observado. Por exemplo, durante um ataque do tipo DoS, os atributos de endereço de IP de origem e porta de destino podem se tornar mais concentrados, enquanto que durante um evento de *flash crowd*, os endereços IP de origem e as portas de origem podem se tornar mais dispersos. Seguindo a linha deste trabalho, e de outros trabalhos semelhantes [2] [3], foi elaborada a Tabela 6 que lista as anomalias que serão usadas neste trabalho e quais as suas consequências na distribuição dos endereços IP e números de portas de origem e destino.

Tabela 6 – Anomalias e atributos de tráfego afetados. Adaptado de [1], [2] e [3].

Anomalia	**Descrição**	**Atributos afetados**
DoS	Ataque de negação de serviço. Um atacante tem o objetivo de derrubar um alvo através do envio de grande quantidade de tráfego.	Endereços IP de origem, IP de destino, portas de destino e atributos quantitativos.
DDoS	Ataque de negação de serviço distribuído. O mesmo objetivo do DoS, porém o ataque parte de vários destinos para um único alvo.	Endereços IP de origem, IP de destino, portas de origem e de destino e atributos quantitativos.
Flash crowd	Um aumento repentino no tráfego para um determinado destino a partir de várias origens. Não configura um ataque necessariamente, pode ser tráfego legítimo.	Endereços IP de destino, portas de destino e atributos quantitativos.

Devido a esse comportamento durante as anomalias, será usado no presente trabalho o número de três atributos para que seja gerado o alarme para o administrador

da rede, ou seja, se no mesmo intervalo a análise de quaisquer três atributos for considerada anômala, aquele intervalo de tempo será considerado uma anomalia e será gerado um alarme para o administrador da rede.

3.5 Conclusão do capítulo

Foram descritos neste capítulo todos os operadores que serão usados pelo GA para a obtenção da assinatura digital da rede. Também foram apresentados quais serão os testes usados para a validação da caracterização do tráfego e quais as métricas usadas para avaliar a detecção de anomalias nas três técnicas descritas. O próximo capítulo fará a apresentação dos resultados obtidos nos testes de validação nos cenários propostos.

4 TESTES E RESULTADOS

Neste capítulo serão descritos os testes realizados nos três cenários propostos, que serão descritos a seguir. Estes testes visam avaliar a capacidade do sistema GADS de realizar a previsão do comportamento da rede de forma eficiente, sendo assim apto a permitir seu uso para detecção de anomalias em redes. Além disso, serão discutidos os resultados desses testes na caracterização do tráfego e na detecção de anomalias em todos os cenários.

Dados foram coletados nos três cenários propostos, os quais são:

- ***Cenário 1.*** Rede da UEL (Universidade Estadual de Londrina) do ano de 2013, início da presente pesquisa. Foram coletados dados no formato sFlow diretamente do *switch Extreme* BD8801 que fica na administração central da rede da UEL. Esse *switch* concentra o tráfego de aproximadamente seis mil computadores, mais aproximadamente quatro mil dispositivos móveis, totalizando em torno de dez mil *hosts* interligados, agregando o tráfego de toda a universidade. Para a exportação dos dados no formato sFlow [50] e para a transformação desses dados binários em texto, para que possam ser analisados, foi usada a ferramenta *nfdump* [39]. Foram realizados testes para verificar a habilidade do GADS na predição do tráfego, bem como a eficiência para detecção de anomalias usando três técnicas distintas para esse fim, já descritos na seção 3.3, que são o ADTW, os limites de confiança e a lógica paraconsistente. Os três métodos foram comparados para avaliação de seus rendimentos;

- ***Cenário 2.*** Rede da Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Toledo do ano de 2013, no formato Netflow v9 [14]. A coleta do tráfego da UTFPR é realizada no gateway principal da rede, que comporta aproximadamente quatrocentos dispositivos, sendo assim significativamente menor que a utilizada no primeiro cenário. Foi avaliada a capacidade do GADS na previsão do comportamento do tráfego, e também a sua habilidade na detecção de anomalias, usando a técnica que obteve o melhor aproveitamento no cenário anterior;

- ***Cenário 3.*** Rede da UEL no ano de 2015. Da mesma forma que no primeiro cenário foram coletados dados na administração central da rede. A partir desses dados foi feita a comparação do GADS com o ACODS (*Ant Colony Optimization for Digital Signature*) [11] [51], uma outra técnica para caracterização do tráfego de redes através da criação de DSNSF. O GADS e o ACODS foram submetidos às avaliações para aferição do desempenho de ambos na caracterização do tráfego e posteriormente na detecção de anomalias.

Serão seis os atributos colhidos nos fluxos de dados usados, sendo eles:

- número de bits trafegados por segundo;
- número de pacotes trafegados por por segundo;
- entropia de Porta de Destino, representada nas figuras como H(SrcPort);
- entropia de Porta de Origem, representada nas figuras como H(DstPort);
- entropia de endereço IP de Destino, representada nas figuras como H(DstIP);
- entropia de endereço IP de Origem, representada nas figuras como H(SrcIP).

A seguir serão descritos os resultados encontrados nas análises de avaliação para o desempenho do GADS na caracterização do tráfego e também para a detecção de anomalias para os cenários propostos.

4.1 Cenário 1

Os dados usados para a execução deste cenário compreendem o intervalo de um mês, o mês de maio de 2013, adquiridos na Universidade Estadual de Londrina (UEL), sendo que para a análise foram considerados apenas os dias úteis, enquanto que para a geração das assinaturas digitais foram usados dados das três semanas anteriores a cada dia analisado.

Para que as medidas de desempenho, elaboradas através da matriz de confusão, pudessem ser avaliadas, foi usada a ferramenta Scorpius [52], que faz a simulação de anomalias no tráfego de rede. Essa ferramenta injeta diretamente nos arquivos exportados, dados que seriam correspondentes a ataques como DoS, DDoS e *Flash crowd*. Para isso foram gerados três ataques sintéticos, todos no dia 23 de maio, que serão descritos a seguir:

- o primeiro foi um ataque de DoS, executado entre as 10 e as 11 horas;
- o segundo foi um ataque de DDoS, onde foram usados 10 IPs de origem, direcionados à porta 80 de um servidor fictício. Esse ataque teria sido realizado das 14 às 15 horas;
- o último não foi um ataque propriamente dito, mas sim uma anomalia do tipo *flash crowd*, que foi injetada entre as 16 e as 17 horas.

A seguir serão apresentados os DSNSFs gerados pelo GADS bem como os resultados obtidos nos cálculos das métricas usadas para avaliação. Foram usados os seis atributos apresentados e foram analisadas quatro semanas entre os dias 01 e 31 de maio

de 2013. Apenas os dias úteis foram considerados e foi usada a janela das 24h de cada dia. Para cada dia de cada uma das semanas analisadas foram usados dados do mesmo dia das três semanas anteriores para a geração do DSNSF. Vale lembrar também que nesse mês houve dois feriados nacionais, dia 1, feriado do dia do trabalhador e dia 30, feriado de *Corpus Christi*. No caso do último, por ser em uma quinta-feira, aconteceu o recesso, também conhecido como ponto facultativo, e também não houve atividades na universidade no dia 31.

4.1.1 DSNSFs gerados para UEL

Como optou-se pelo uso da janela de tempo de 1 minuto, o DSNSF será composto de 1440 pontos, conforme descrito na seção 3.1, que representam a quantidade total de minutos em um dia. Na Figura 11 é possível ver na cor verde o tráfego real da UEL nos dias 20 e 21 de maio de 2013 para cada uma das 24 horas do dia. A linha em vermelho representa o DSNSF gerado pelo GADS para o mesmo dia. Para a criação do DSNSF foram usados dados obtidos nas três semanas anteriores a este dia, portanto para a assinatura digital do dia 20 de maio foram usados os dados históricos dos dias 13 e 6 de maio e 29 de abril. Já para o DSNSF do dia 21 de maio foram usados os dados dos dias 14 e 7 de maio e 30 de abril.

Percebe-se que o modelo proposto conseguiu acompanhar o crescimento do tráfego em todos os atributos, bem como a diminuição do mesmo em função do horário. Nota-se que principalmente para os atributos bits e pacotes por segundo, às 7h o tráfego começa a ter um aumento devido ao início das atividades na universidade. Às 13h há uma pequena diminuição do volume de tráfego, que volta ao mesmo nível logo às 14h. A partir das 17h ocorre uma diminuição no volume trafegado em função das poucas atividades no período noturno. Nos atributos quantitativos que são apresentados nas entropias, o DSNSF consegue acompanhar o movimento do tráfego do mesmo jeito, porém de uma forma mais suave.

As Figuras 12 e 13 também apresentam os movimentos de tráfego, agora entre os dias 22 a 24 de maio de 2013 respectivamente e também para o feriado de primeiro de maio. Também nota-se que o DSNSF acompanha o movimento do tráfego de forma correta.

A Figura 13b apresenta os movimentos de tráfego para cada um dos atributos no feriado de primeiro de maio. É evidente o desacordo entre o tráfego previsto e o tráfego real, causado pelo pouco volume de tráfego em virtude da pouca atividade na universidade.

(a) Dia 20 de maio de 2013 (b) Dia 21 de maio de 2013

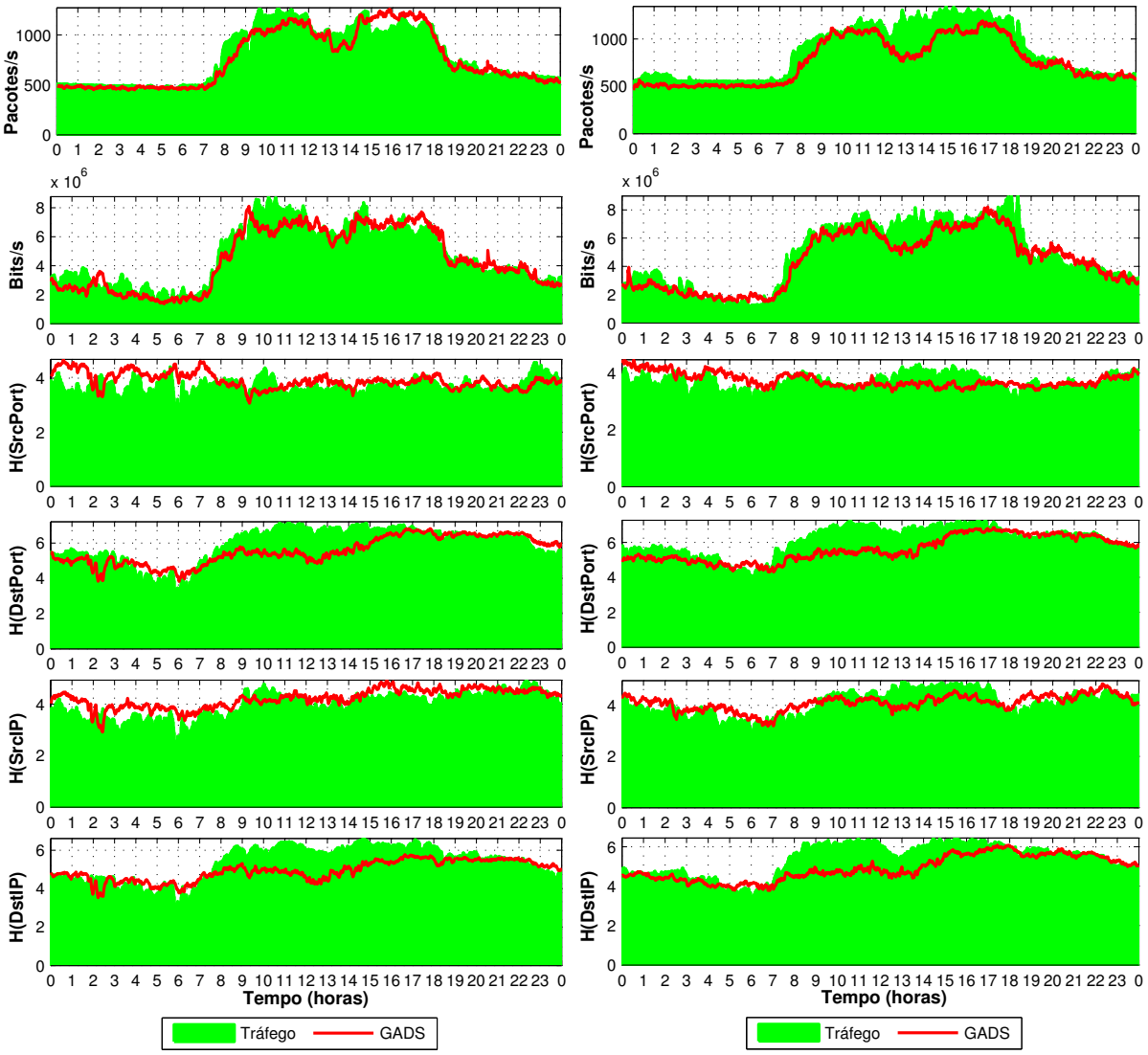

Figura 11 – DSNSF e movimentos do tráfego para os dias 20 e 21 de maio de 2013

(a) Dia 22 de maio de 2013 (b) Dia 23 de maio de 2013

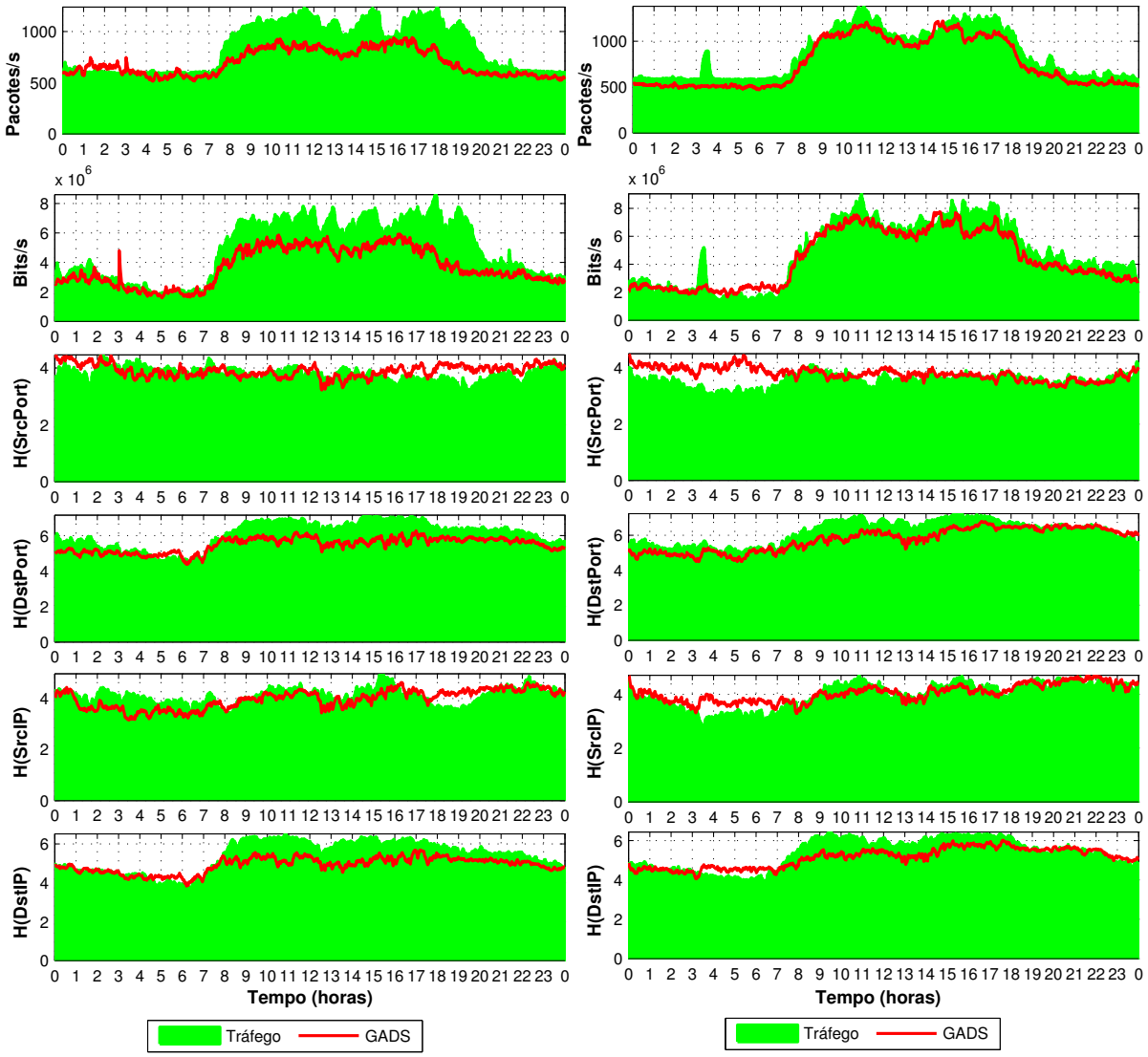

Figura 12 – DSNSF e movimentos do tráfego para os dias 22 e 23 de maio de 2013

(a) Dia 24 de maio de 2013 (b) Dia 01 de maio de 2013

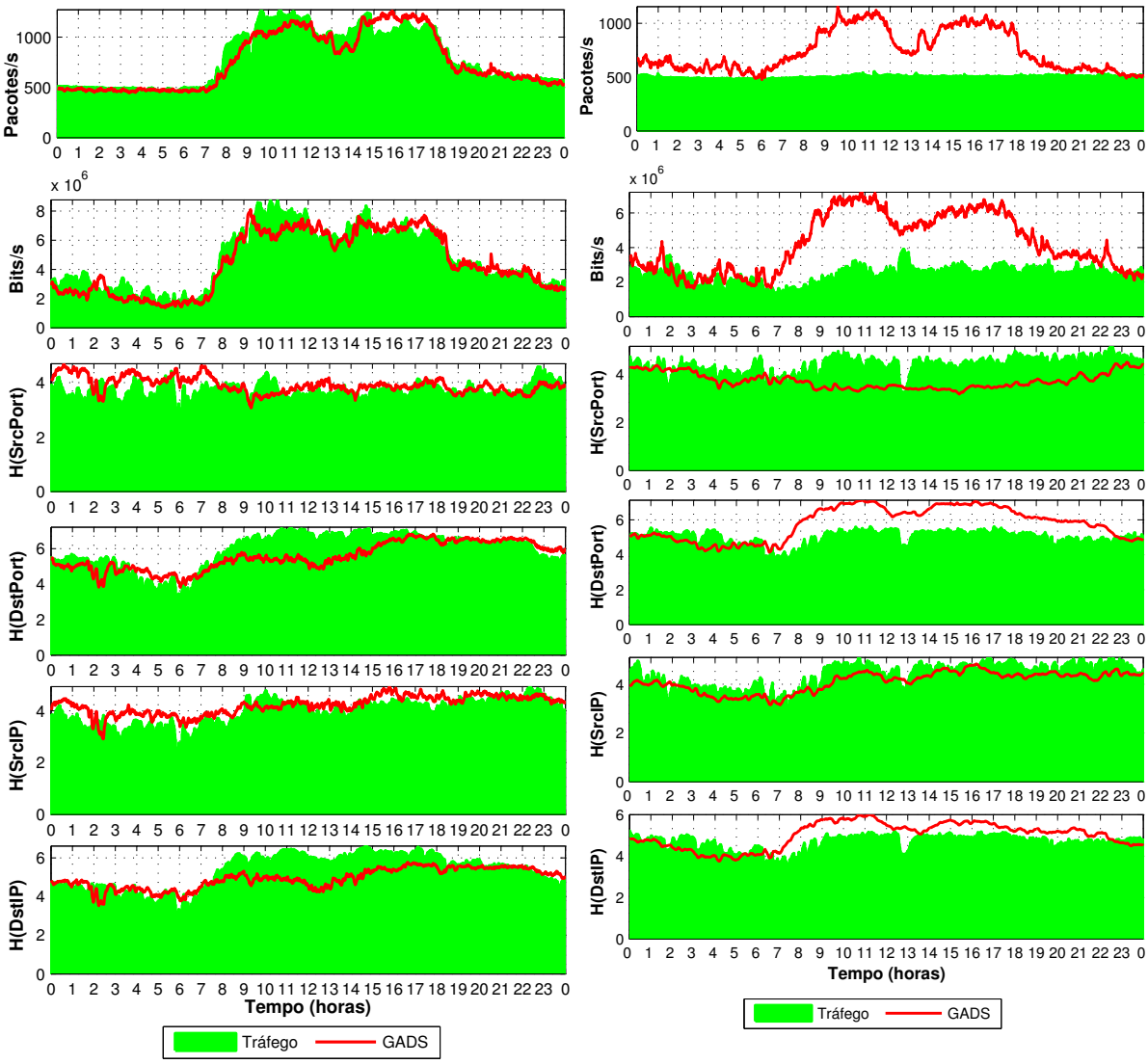

Figura 13 – DSNSF e movimentos do tráfego para os dias 24 e 01 de maio de 2013

4.1.2 Resultados para NMSE

O erro quadrático médio normalizado (NMSE) avalia a habilidade de previsão do comportamento da rede pelo sistema GADS, apresentado neste trabalho. Os valores possíveis estão compreendidos no intervalo $0 \leq \text{NMSE} \leq 1$ e valores próximos de zero são os mais adequados.

A Figura 14 apresenta os valores de NMSE para cada um dos dias analisados para os atributos quantitativos bits por segundo. Percebem-se valores muito próximos de 0 para quase todos os dias analisados. A exceção fica pelos dias 1, 30 e 31 de maio, os feriados e o ponto facultativo, onde a falta de tráfego gerou um valor mais próximo de 1, indicando que o tráfego previsto e o tráfego real não estão em acordo.

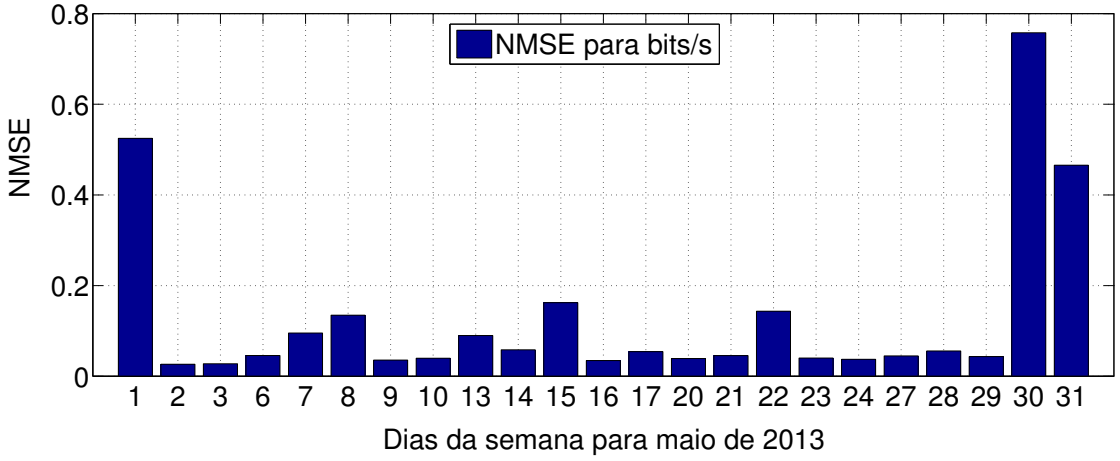

Figura 14 – NMSE para bits no mês de maio de 2013

Na Figura 15 observam-se os valores de NMSE para os atributos pacotes por segundo para todos os dias úteis do mês de maio de 2013 também. Novamente as exceções de valores considerados ótimos ficaram com os dias 1, 30 e 31.

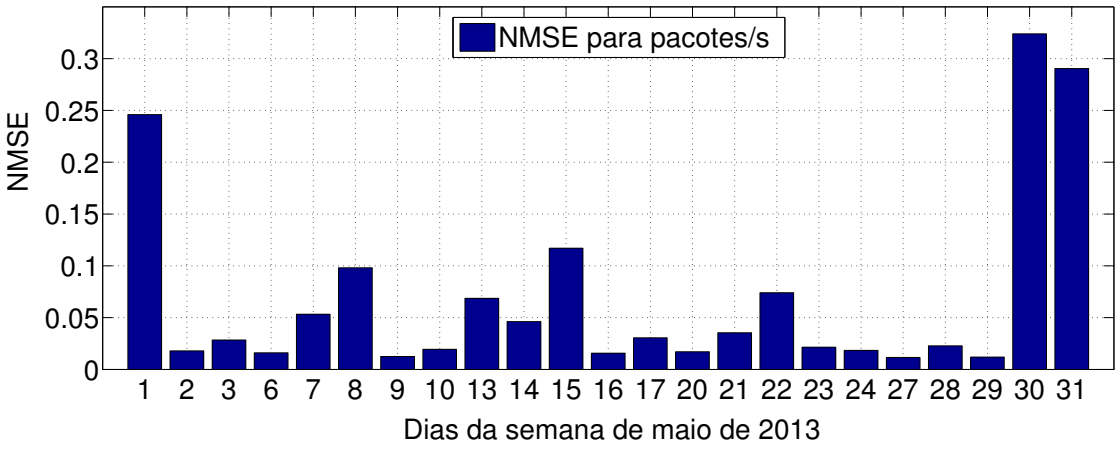

Figura 15 – NMSE para pacotes no mês de maio de 2013

4.1.3 Resultados para Coeficiente de Correlação (CC)

O coeficiente de correlação (CC) mensura o quanto os dados estão ou não correlacionados, indicando em valores, se o tráfego previsto está correlacionado com o tráfego real. Valores próximos de 1 indicam que há correlação entre os dados, ou seja, que a previsão está correta. Da mesma forma que a análise de NMSE, os atributos quantitativos foram analisados separadamente dos atributos qualitativos.

A Figura 16 apresenta os valores de CC para os atributos quantitativos bits por segundo para os dias analisados. Observam-se valores muito próximos de 1 para todos os dias, exceto nos feriados e no dia 31, o que indica que houve correlação entre o DSNSF gerado pelo GADS e os dados reais de cada dia.

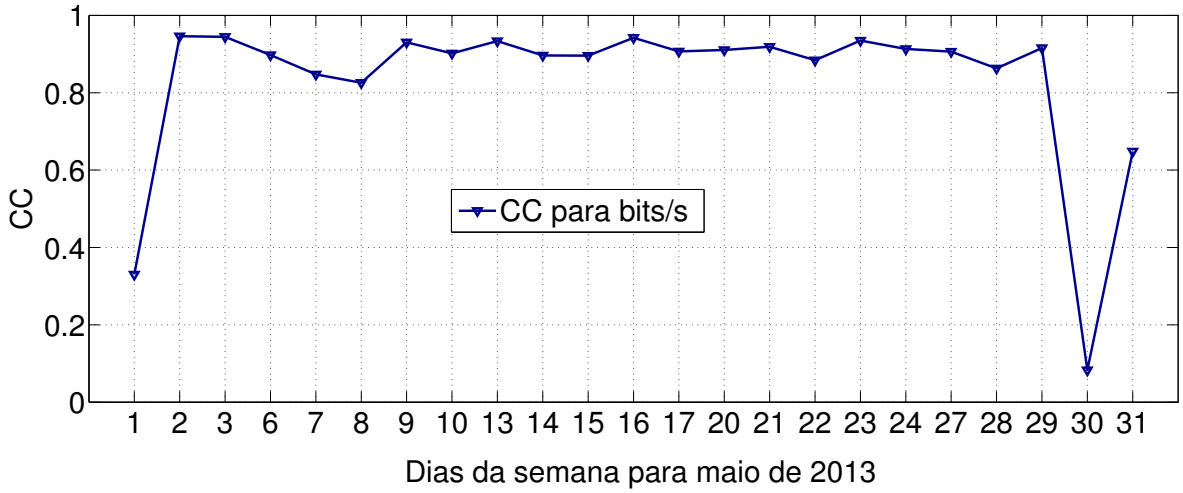

Figura 16 – CC para bits no mês de maio de 2013

Na Figura 17 observam-se os valores de CC para pacotes por segundo nos dias analisados. Da mesma forma que os valores de bits, os números indicam uma alta correlação entre o tráfego previsto e o real para todos os dias, exceto nos dias que não houve atividades na universidade.

4.1.4 Resultados dos testes de desempenho para o ADTW

Como descrito anteriormente, um dos intuitos dos vários testes que serão executados é avaliar os coeficientes de significância de limiares propostos, ou seja, qual a margem de tráfego que será usada para que aquele intervalo seja considerado normal ou anômalo. Para o ADTW estão sendo propostos sete valores possíveis, de acordo com testes feitos previamente. Esses valores serão chamados de coeficientes de significância σ e foram submetidos aos testes de acurácia, *recall*, precisão, especificidade, *f-measure* e à curva ROC e serão apresentados a seguir.

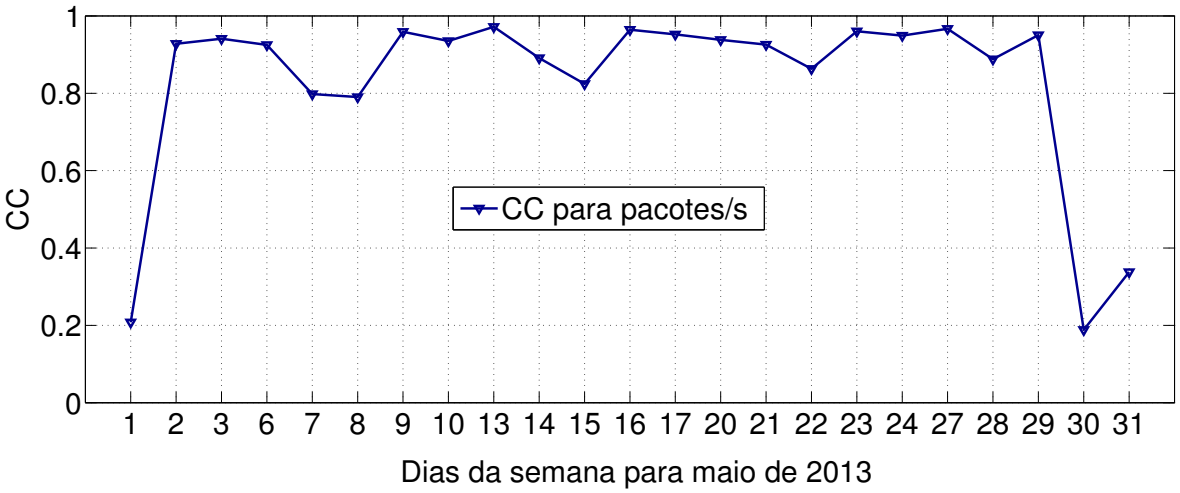

Figura 17 – CC para pacotes no mês de maio de 2013

O primeiro teste de desempenho foi a acurácia, que pode ser analisado na Figura 18. Percebe-se no gráfico, que foram encontrados valores muito próximos nas seis primeiras métricas, todos entre 97% e 98%, ficando apenas o valor de σ em 10%, com baixa acurácia.

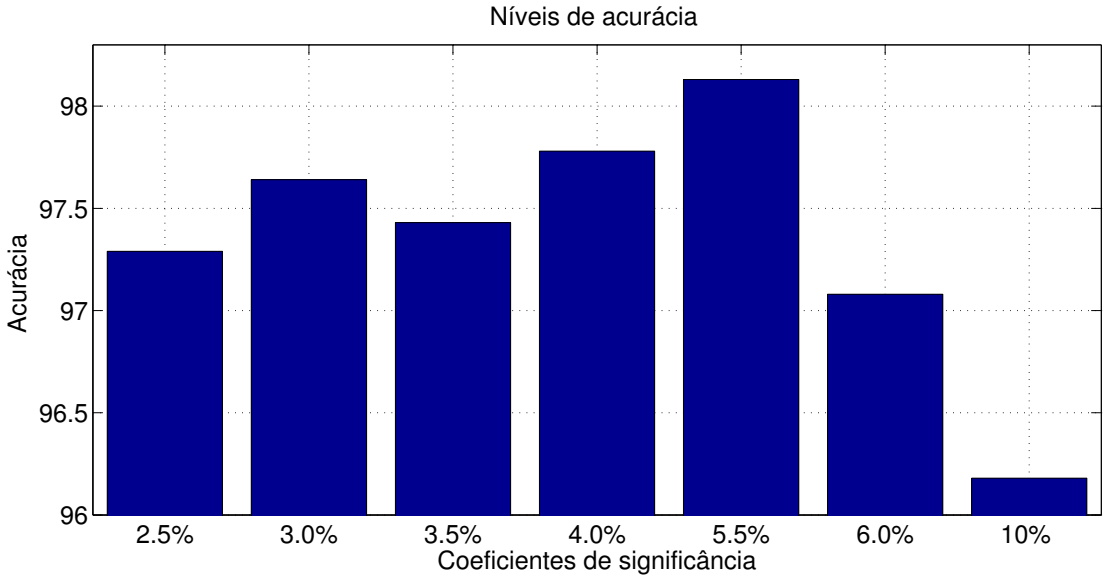

Figura 18 – Acurácia para os diversos valores de σ

As taxas de *recall* encontradas são apresentadas na Figura 19, onde é possível perceber boas taxas para os cinco primeiros valores de σ. Os dois últimos valores que são as taxas de σ em 6% e 10% foram os únicos que apresentaram taxas abaixo de 87%.

Na Figura 20 são apresentados as taxas de precisão obtidas para cada um dos valores dos coeficientes σ. Nesta figura percebe-se que quanto mais se aumentam os níveis de *thresholds*, mais o sistema apresenta alta taxa de falsos positivos. Este é um gráfico que não pode ser analisado individualmente, pois aumentando os níveis de *thresholds*, aumenta-se também o número de alarmes falso positivos, que são apresentados nos gráficos

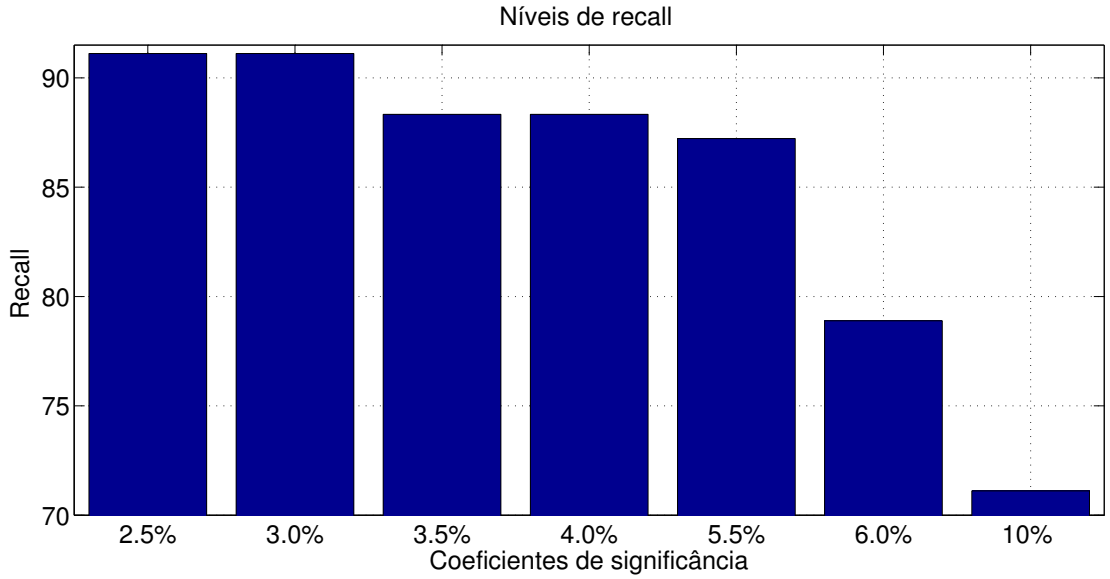

Figura 19 – Recall para os diversos valores de σ

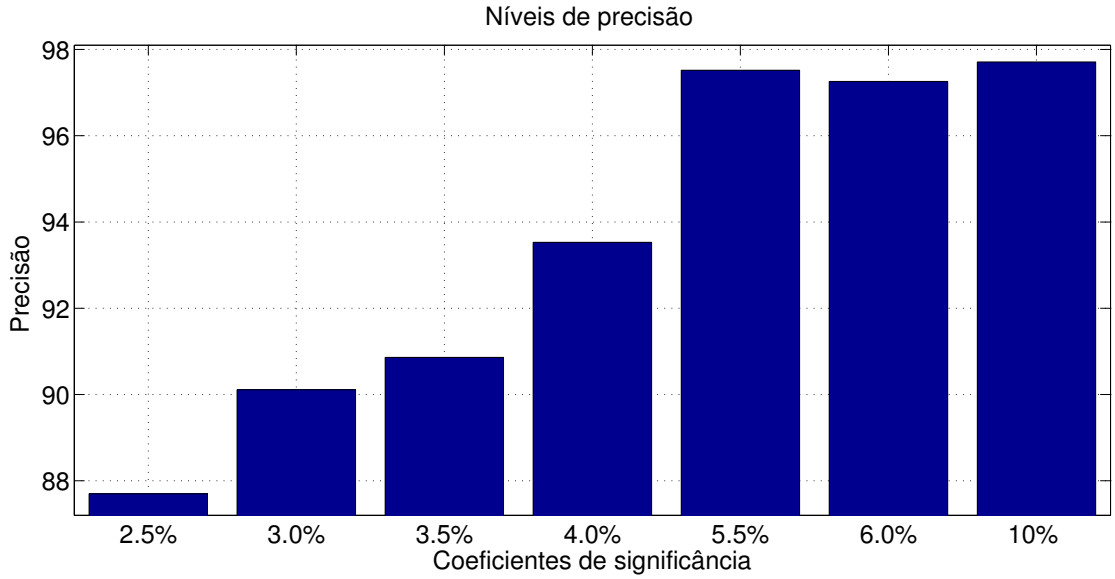

Figura 20 – Precisão para os diversos valores de σ

de especificidade. É importante notar taxas acima de 90% a partir de valores σ de 3%.

As taxas de especificidade são apresentadas na Figura 21. Pelo gráfico notam-se que os valores de σ entre 2.5% e 5.5% apresentam taxas acima de 98%.

Os resultados para o teste de *f-measure* apresentados na Figura 22, têm nos valores de σ entre 3.0% e 5.5% taxas acima de 90%.

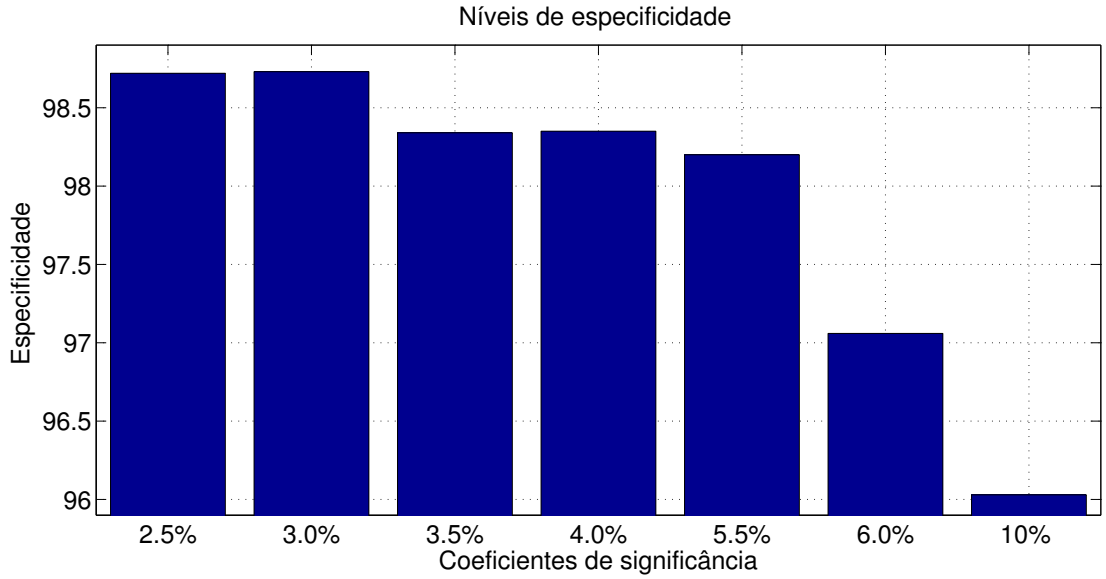

Figura 21 – Especificidade para os diversos valores de σ

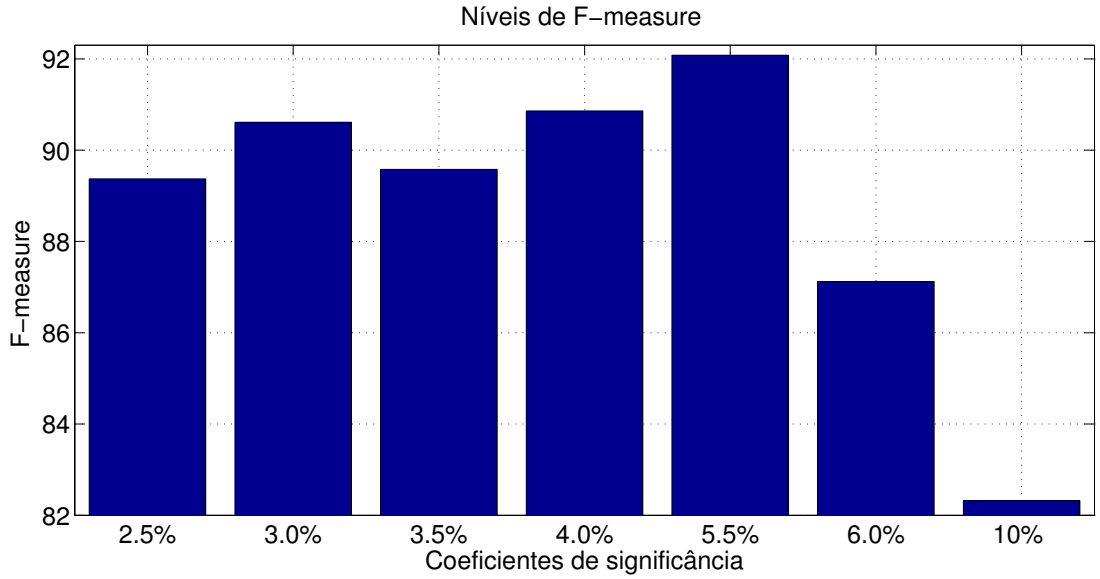

Figura 22 – *F-measure* para os diversos valores de σ

Para que pudesse ser escolhido o melhor dentre os testes realizados para os valores de σ, a Tabela 7 apresenta quais métricas obtiveram valores ideais, ou seja, um aproveitamento acima de 90%. As métricas que obtiveram estes valores estão marcadas com o X na tabela. Na última linha é possível acompanhar o total de métricas cujo valor de σ obteve aproveitamento acima de 90%, onde notam-se os coeficientes σ de 3.0% e 3.5% com um ótimo aproveitamento em todos os 5 testes.

Tabela 7 – Valores de σ e métricas de desempenho

Métrica / σ	2.5	3.0	3.5	4.0	5.5	6.0	10
Acurácia	X	X	X	X	X	X	X
Recall	X	X	X				
Precisão		X	X	X	X	X	X
Especificidade	X	X	X	X	X	X	X
F-measure		X	X	X	X		
Total	3	5	5	4	4	3	3

Ainda para que pudesse ser aferido o melhor dos valores para o coeficiente σ, foi traçada a curva ROC para todos os valores testados. A Figura 23 mostra uma curva ROC para os diversos valores de σ. Como o ADTW busca padrões de comportamentos anômalos na comparação do tráfego real com o DSNSF, o valor de σ representa percentualmente a margem de diferença, para mais ou para menos, entre os dois tráfegos, real e previsto. É possível perceber pelo zoom da imagem que a linha que mais se aproxima do ideal, que seria 100% de verdadeiros positivos e 0% de falsos positivos é a linha na cor preta, que representa um coeficiente σ de 3.5%.

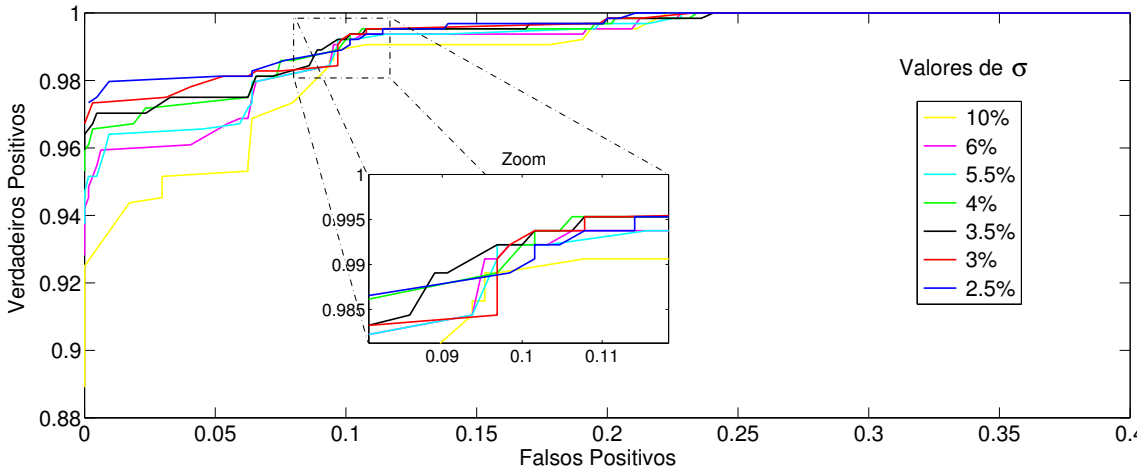

Figura 23 – Curva ROC para os diversos valores de σ do ADTW

Pelos testes realizados é possível inferir que o melhor valor para o coeficiente de significância, ou seja, a margem para que o intervalo de tráfego seja considerado anômalo, é o valor de 3.5%.

4.1.4.1 Alarmes gerados pelo ADTW

Conforme descrito foram realizados três ataques pelo software Scorpius, todos no dia 23 de maio. Um ataque de DoS, um de DDoS e a outra anomalia um *flash crowd*.

A partir dos ataques realizados pelo Scorpius, é possível avaliar a capacidade do GADS em conjunto com o ADTW de gerar alarmes corretamente. Os limiares definidos

para o ADTW foram os de coeficiente σ de 3.5%. Um alarme só será disparado ao administrador se ao menos três atributos forem considerados fora do *threshold*, tanto acima quanto abaixo do valor de σ no mesmo intervalo de tempo. A Figura 24 mostra todos alarmes gerados para cada um dos atributos.

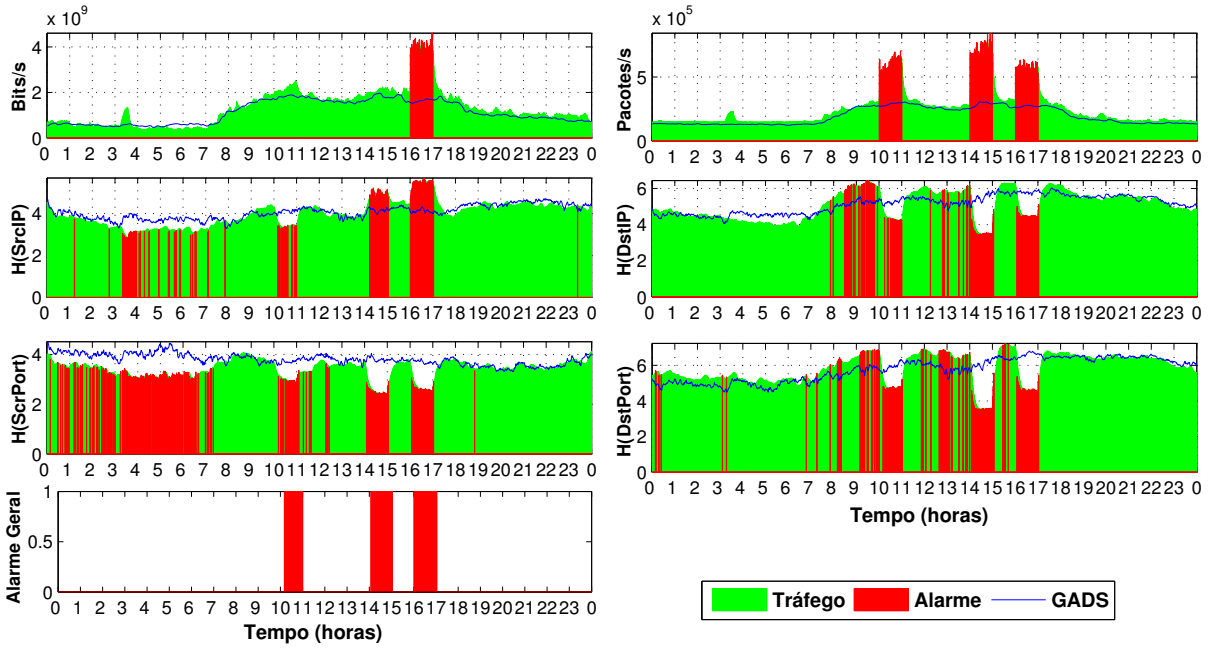

Figura 24 – Alarmes gerados pelo ADTW

O ADTW identificou corretamente o comportamento do tráfego, gerando alarmes apenas durante os períodos de tempo que ocorreram as anomalias sintéticas injetadas pelo Scorpius. É possível verificar para as taxas de coeficiente σ de 3.5%, valores de acurácia de quase 98%, indicando um alto índice de acertos e taxas de falsos positivos de apenas 1%, muito baixas mesmo para os padrões mais exigentes.

4.1.5 Resultados dos testes de desempenho para os limiares de confiança (*thresholds*)

Para o uso dos limiares de confiança também foram propostos sete valores de coeficiente de significância σ, os quais são 1,5%, 1,8%, 2.0%, 2.2%, 2,5%, 2,8% e 3.0%, pois foram os valores que apresentaram os melhores resultados nos testes de desempenho. Todos os valores para o coeficiente de significância σ foram submetidos aos testes de acurácia, *recall*, precisão, especificidade, *f-measure* e à curva ROC, cujos resultados serão apresentados a seguir. É importante salientar que índices muito baixos de coeficiente de significância σ tendem a apresentar muitos alarmes de tráfego anômalo, o que pode gerar altos valores de alarmes falsos positivos, ou seja, intervalos de tráfego que foram classificados como anômalos, porém representavam um comportamento normal do tráfego. Ao mesmo tempo, valores muito altos para o mesmo coeficiente tendem a não gerar muitos

alarmes, o que também não é um bom comportamento, ocasionando os chamados falsos negativos, ou seja, intervalos de tráfego que podiam apresentar anomalias, mas foram classificados como normais.

O teste da acurácia pode ser observado na Figura 25, onde são encontrados índices acima de 90% para os valores de σ a partir de 1.8%.

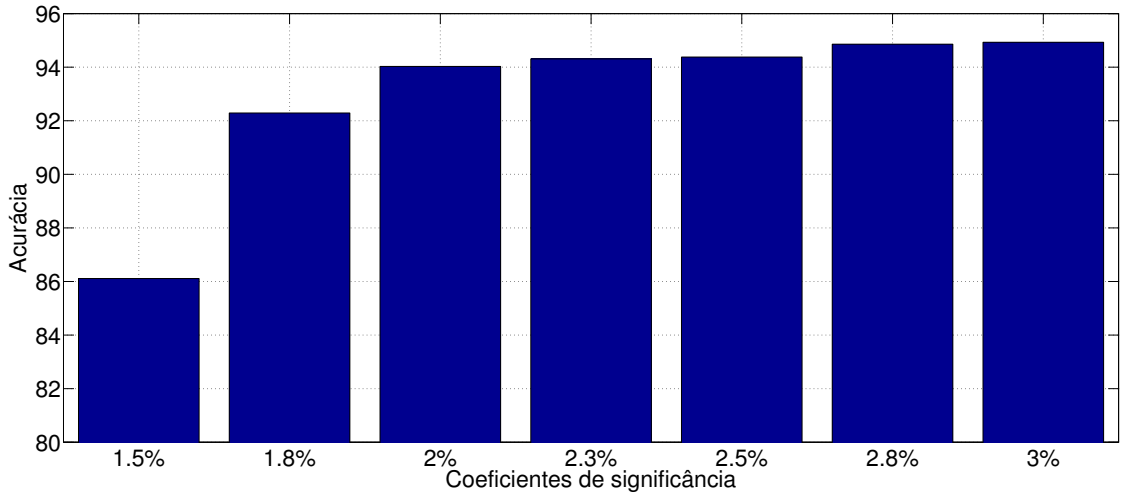

Figura 25 – Acurácia para limiares de confiança

O teste de *recall* apresentado na Figura 26, avalia a porcentagem de amostras positivas classificadas corretamente, sobre o total de amostras realmente positivas, por esta razão apresentou índices bastante altos para os menores valores de coeficiente de significância, e índices baixos para os valores maiores. Foram encontrados índices acima de 75% para os valores de σ de 1.5%, 1.8% e 2%.

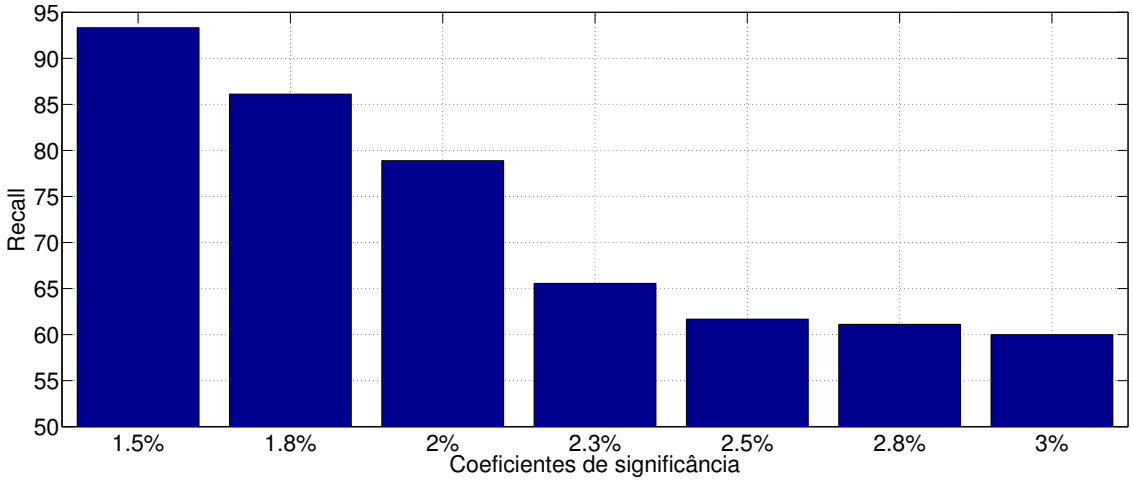

Figura 26 – Recall para limiares de confiança

Já o teste de precisão apresentado na Figura 27 mostra a porcentagem de amostras positivas classificadas de forma correta, sobre todas as amostras classificadas como positivas, e por este motivo apresenta índices elevados para os valores de coeficiente de significância mais altos, com taxas acima de 75% para os índices maiores que 2%.

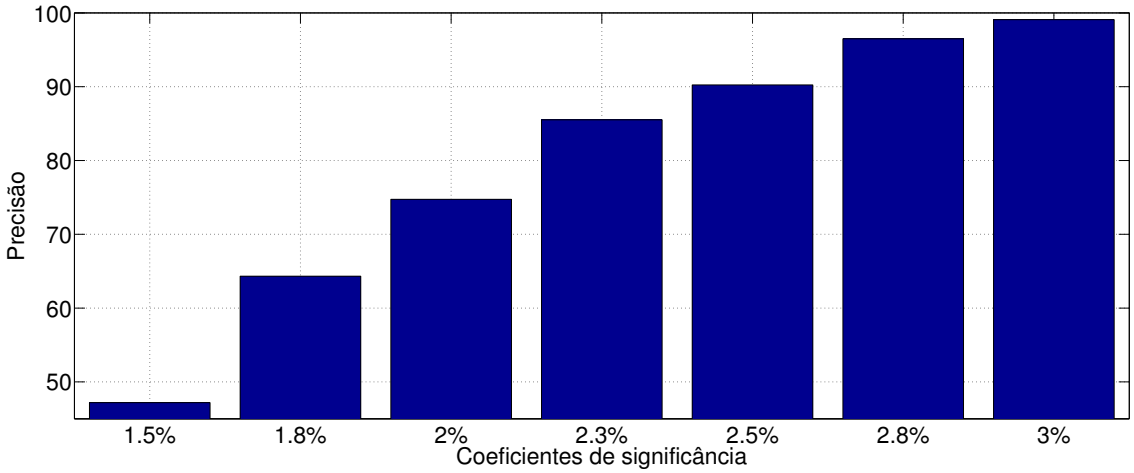

Figura 27 – Precisão para limiares de confiança

O próximo teste é o de especificidade, apresentado na Figura 28, e que ilustra a porcentagem de amostras negativas identificadas de forma correta, ou seja, que realmente não eram uma anomalia. Percebem-se valores muito próximos, todos acima de 90% para todos os índices testados.

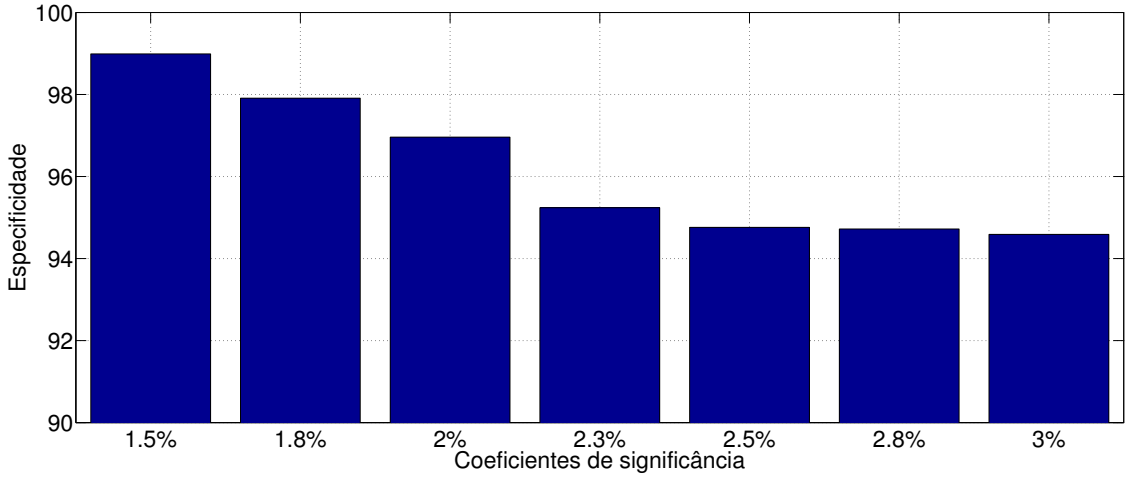

Figura 28 – Especificidade para limiares de confiança

O teste de *f-measure* é ilustrado na Figura 29, e apresenta a média entre a precisão e o *recall*, mostrando valores próximos de 75% para os coeficientes 2.8% e 3% e acima de 75% para o índice de 2%.%

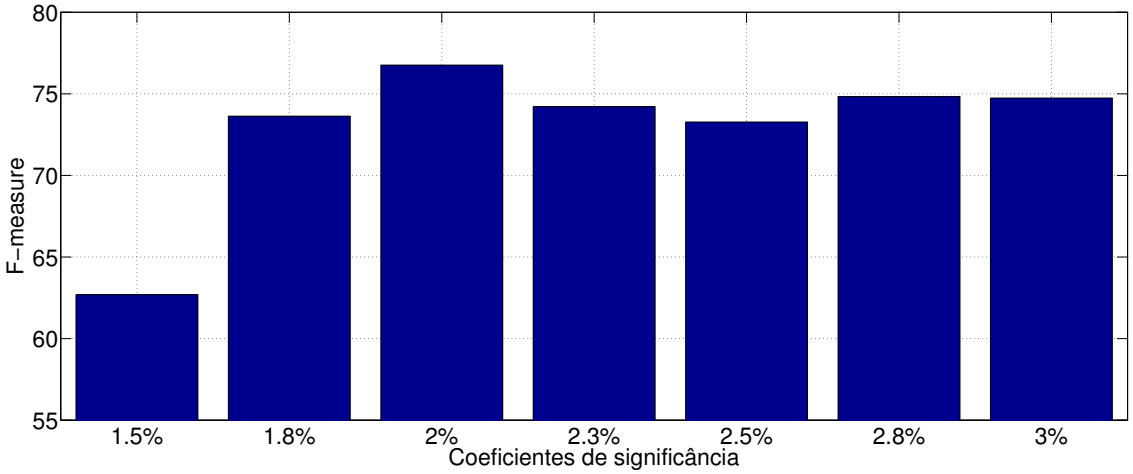

Figura 29 – F-measure para limiares de confiança

Mais uma vez, todos os valores considerados bons foram compilados na Tabela 8 para serem melhor apresentados. Devido aos índices dos testes serem mais baixos em relação ao ADTW foram considerados bons valores acima de 75%. As métricas que obtiveram esses valores estão marcadas com o X na tabela, sendo que na última linha encontram-se os totais de métricas cujo valor de σ obteve aproveitamento acims de 75% nos testes realizados. Percebe-se que o valor de σ para 2% apresentou melhor desempenho em relação aos seus pares.

Tabela 8 – Valores de σ e métricas de desempenho para os limiares de confiança

Métrica / σ	1.5	1.8	2.0	2.3	2.5	2.8	3
Acurácia	X	X	X	X	X	X	X
Recall	X	X	X				
Precisão			X	X	X	X	X
Especificidade	X	X	X	X	X	X	X
F-measure			X				
Total	3	3	5	3	3	3	3

Ainda no intuito de aferir o melhor valor para o coeficiente de significância σ, é apresentada na Figura 30 a curva ROC representando todos os valores testados para σ, que representa percentualmente a margem de tráfego superior e inferior, calculada com base no DSNSF e que será comparada com o tráfego real na busca de anomalias. Na curva ROC, os valores ideais são aqueles que se aproximarem de 100% de verdadeiros positivos e 0% de falsos positivos. No zoom observa-se uma linha pontilhada, que representa uma linha imaginária entre os valores de 100% dos dois eixos (X e Y). Sendo assim, é possível observar na figura, pelo zoom da imagem, que por uma margem muito pequena, o melhor valor é o apresentado na linha vermelha, que representa o coeficiente σ para 2%. Esse resultado ratifica os valores encontrados nos outros testes de desempenho.

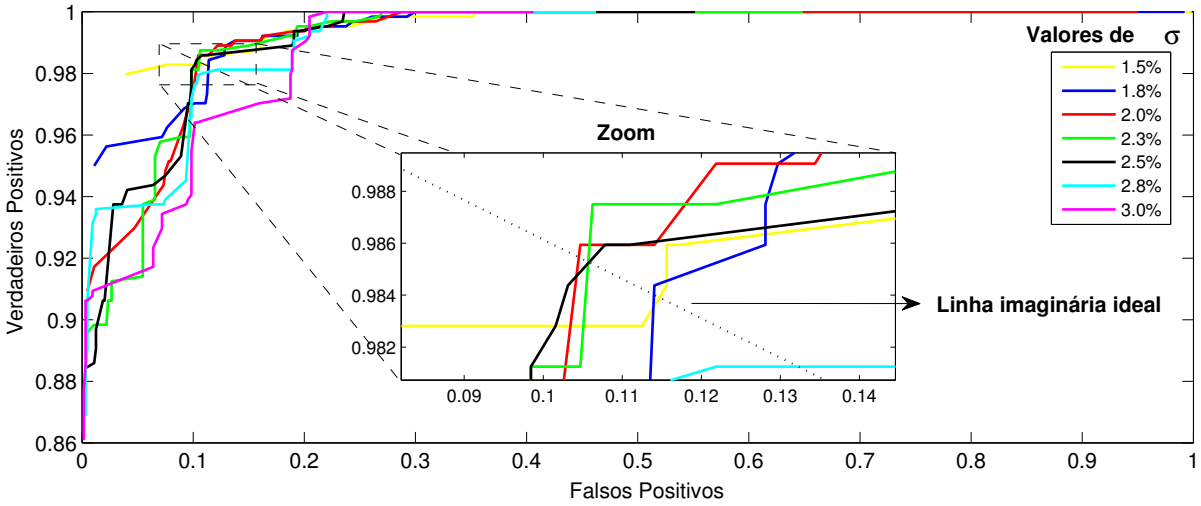

Figura 30 – Curva ROC para os diversos valores de σ dos limiares de confiança

4.1.5.1 Alarmes gerados pelos limiares de confiança

A Figura 31 ilustra os alarmes gerados pelo GADS para o coeficiente de significância σ ajustado para os valores de 2% superior e inferior. O tráfego real pode ser observado na cor verde, o DSNSF está representado pela linha azul e as linhas pontilhadas na cor preta são os limiares de confiança superior (*threshold up*) e inferior (*threshold down*). Os limiares são calculados usando como base o DSNSF e quando o tráfego real ultrapassa quaisquer dos limites inferior ou superior é gerado um alarme, que na figura é representado pela cor vermelha. Quando três ou mais atributos geram um alarme para o mesmo intervalo de tempo, o administrador da rede é notificado e essas notificações podem ser observadas no quadro alarme geral.

É interessante notar no quadro alarme geral da Figura 31 a quantidade de alarmes considerados falsos positivos, especialmente quando há pouco tráfego sendo analisado. O alto índice de alarmes do tipo falso positivos é uma das desvantagens das técnicas de detecção de anomalias baseadas na caracterização do tráfego. Estes alarmes ocorrem principalmente nos horários em que quase não há atividade na UEL, compreendidos entre as 23 horas e as 7 horas. Se esses horários não fossem levados em consideração ou durante esses horários houvesse um ajuste nos valores dos limiares, haveria apenas um alarme de falso positivo. Porém, estes ajustes mais precisos ficarão como proposta para trabalhos futuros.

Foram observados para os limiares de confiança boas taxas de acurácia com 94% e taxas de falso positivo de apenas 3,8% apesar da quantidade de alarmes gerados. As taxas de verdadeiros positivos ficaram em 78,9%. Com isso foram identificadas corretamente as anomalias injetadas no tráfego de rede, porém com um desempenho menos efetivo que da técnica anterior, que usava o ADTW.

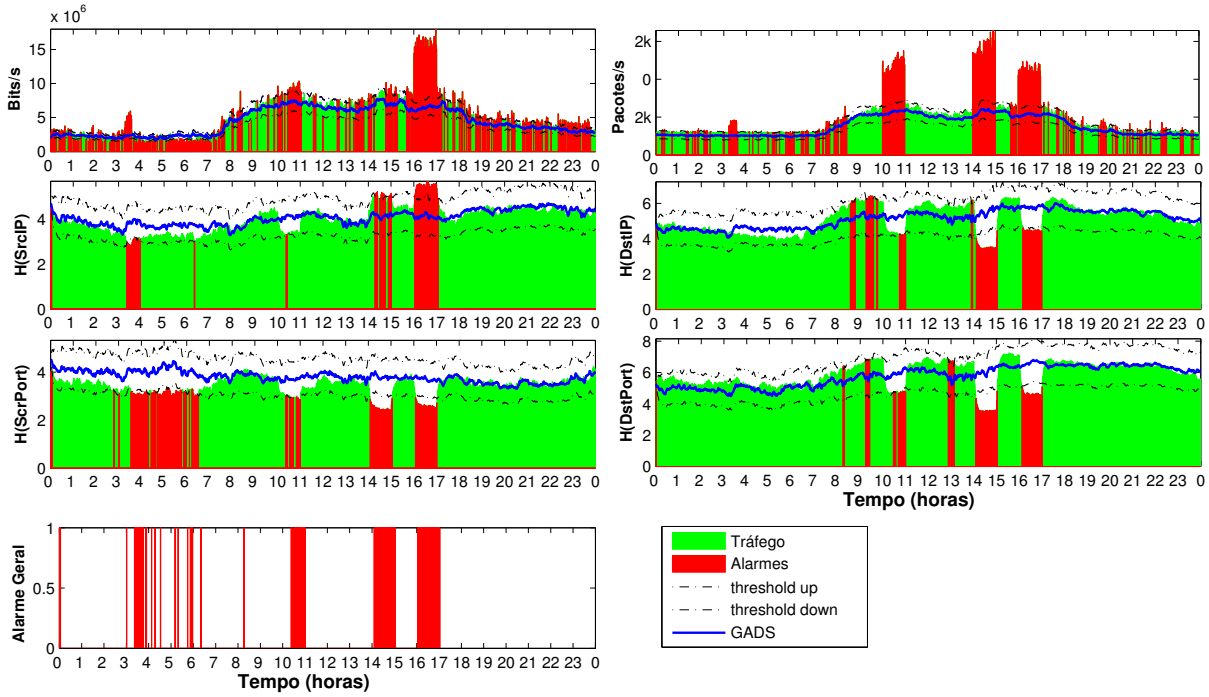

Figura 31 – Alarmes gerais para os limiares de confiança

4.1.6 Resultados dos testes de desempenho para a Lógica Paraconsistente (LP)

A detecção de anomalias usando a LP foi proposta por Pena *et al.* [49], onde o DSNSF era gerado usando o modelo Auto-Regressivo Integrado de Médias Móveis, o *AutoRegressive Integrated Moving Average* (ARIMA). A LP usa sua extensão, a *Lógica Paraconsistente Anotada* de anotação com dois valores (LPA2v), que considera as evidências favoráveis e contrárias a uma determinada proposição. Foi desenvolvida pelo autor, a ferramenta denominada Máquina Paraconsistente Correlacional (MPC), que recebe como entrada os dois DSNSFs e as leituras de tráfego reiais da rede. Depois, a ferramenta utiliza uma sequência de procedimentos para extrair informações evidenciais derivadas do contraste entre cada DSNSF e as leituras reais.

Como são necessárias duas proposições, ou seja dois DSNSFs para que a detecção de anomalias possa ser feita, foi usada a abordagem utilizada no trabalho de [8] que gera uma assinatura de rede denominada ACODS, usando a metaheurística *Ant Colony Optimization* (ACO) para a geração da assinatura digital. O outro DSNSF usado foi o proposto neste trabalho, gerado pelo GADS.

Pelas particularidades do funcionamento da MPC, foram usados os valores padrões de limiares e cálculos internos propostos em [49] e os resultados são os ilustrados na Figura 32.

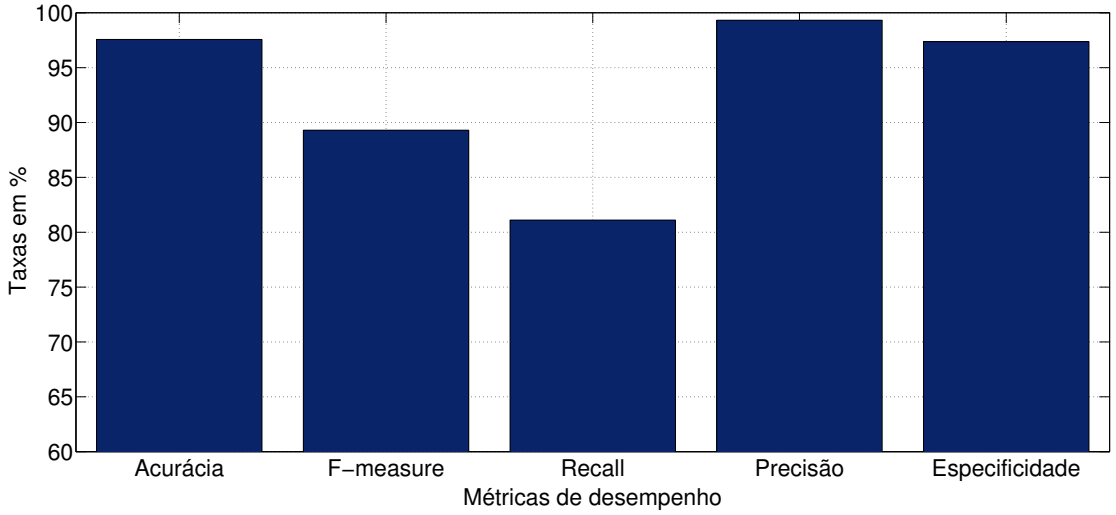

Figura 32 – Métricas de desempenho para a detecção de anomalias através da LP

Os alarmes são apresentados na Figura 33, onde podem ser identificados na cor vermelha os alarmes gerados pela MPC. Pode-se notar que os alarmes foram gerados nos intervalos de tempo que ocorreram as injeções sintéticas das anomalias, com pouco índice de falsos positivos.

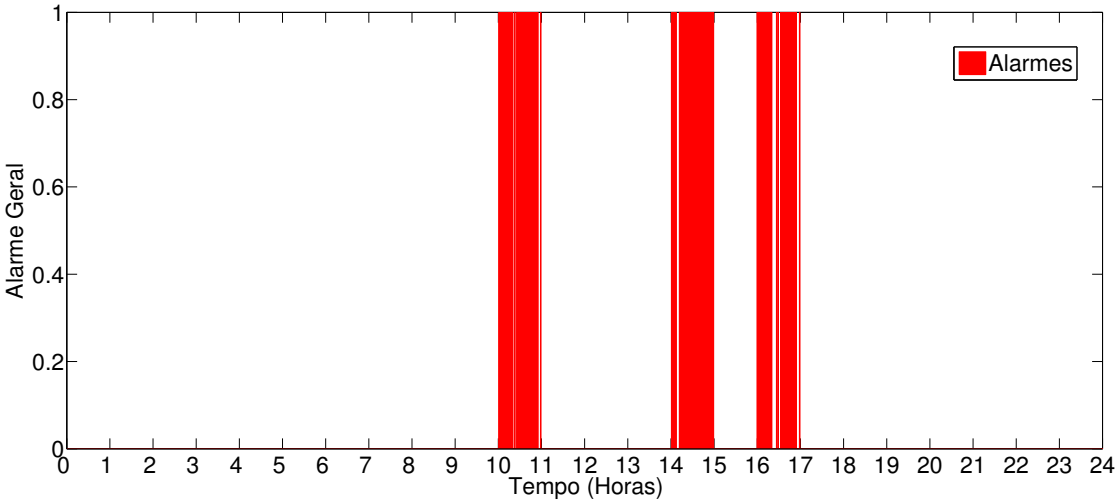

Figura 33 – Alarme geral para a detecção de anomalias através da LP

Os alarmes e as taxas para as métricas de desempenho encontrados na detecção de anomalias pela MPC mostram uma boa eficácia deste método. A MPC obteve uma alta taxa de acurácia de 97,5%, taxas de verdadeiro positivo de 81.1% e uma baixíssima taxa de falsos positivos de 0.08%.

4.1.7 Comparação dos três métodos

Foram avaliados três métodos de detecção de anomalias em redes de computadores, cujos resultados foram derivados de avaliações usando cinco métricas de

desempenho distintas, todas baseadas na matriz de confusão. Os três métodos, na ordem que foram apresentados, são o método baseado na técnica ADTW (*Adaptive Dynamic Time Warping*), o método baseado nos limites de confiança simples (*thresholds*) e usando a técnica denominada lógica paraconsistente, implementada pela Máquina Paraconsistente Correlacional (MPC).

Estas três abordagens foram avaliadas sob os mesmos aspectos e usaram a mesma assinatura digital de segmento de rede, chamada de DSNSF, gerada através do GADS *(Genetic Algorithm for Digital Signature)*, que utilizou o algoritmo genético para a geração desta assinatura, cujo objetivo é fazer a caracterização do tráfego de rede baseado no histórico desta rede.

Para que os alarmes pudessem ser gerados para os três métodos de detecção de anomalias, foram injetadas através do software Scorpius anomalias sintéticas, as quais foram um ataque de DoS, um ataque de DDoS e uma anomalia do tipo *flash crowd*. Os três métodos foram calibrados para que pudessem obter as melhores taxas nas cinco métricas de desempenho analisadas.

A Figura 34 ilustra o desempenho de cada método de detecção de anomalias para cada uma das cinco métricas avaliadas. Pode-se observar na figura, o ADTW na cor azul, o método usando *thresholds* na cor vermelha e a MPC na cor verde. No eixo y do gráfico estão as taxas obtidas em % para cada teste, que são apresentados no eixo x. Como já dito anteriormente, nenhuma das cinco métricas pode ser avaliada individualmente, sendo a análise do conjunto delas a melhor opção quando procura-se comparar seus resultados. É notável na figura que o método ADTW obteve um melhor aproveitamento em relação aos seus pares na maior parte das avaliações, obtendo o maior número de métricas próximas de 100%.

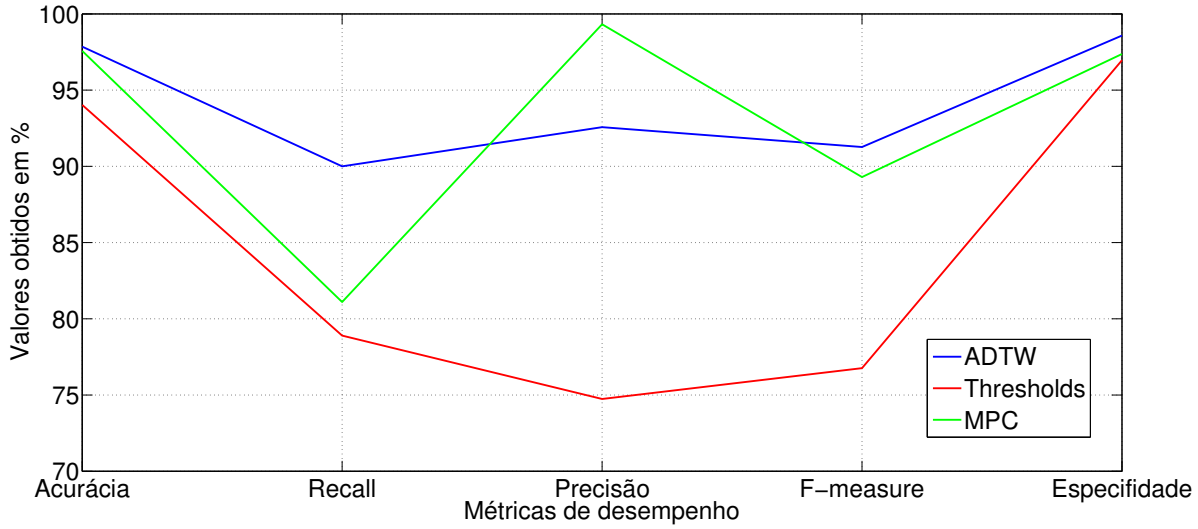

Figura 34 – Comparação do desempenho dos três métodos

Além dos valores encontrados nas métricas de desempenho, devem também ser analisadas as taxas de verdadeiro positivos (TPR) e as taxas de falso positivos, que são calculadas pelas Equações 4.1 e 4.2.

$$TPR = \frac{TP}{TP + FN} \tag{4.1}$$

$$FPR = \frac{FP}{FP + TN} \tag{4.2}$$

As taxas TPR são calculadas pela quantidade total de amostras verdadeiro positivas (TP), ou seja, todas as vezes que o tráfego foi classificado como anômalo e realmente apresentou uma anomalia, sobre a soma destas amostras com as que foram classificadas como normais e eram anômalos (FN). Já as FPR são calculadas pela quantidade total de amostras que foram classificadas como anômalas e eram normais (FP), sobre a soma destas com as amostras que foram classificadas corretamente como tráfego normal (TN).

As Figuras 35 e 36 apresentam ambas as taxas de TPR e FPR para os três métodos avaliados. Na Figura 35 observam-se taxas acima de 90% para o ADTW, apresentado na cor azul e taxas próximas a 80% para os outros dois métodos, o que caracteriza um excelente aproveitamento para o ADTW, visto que em mais de 90% das vezes que uma amostra foi classificada como anomalia ou normal, a predição foi correta.

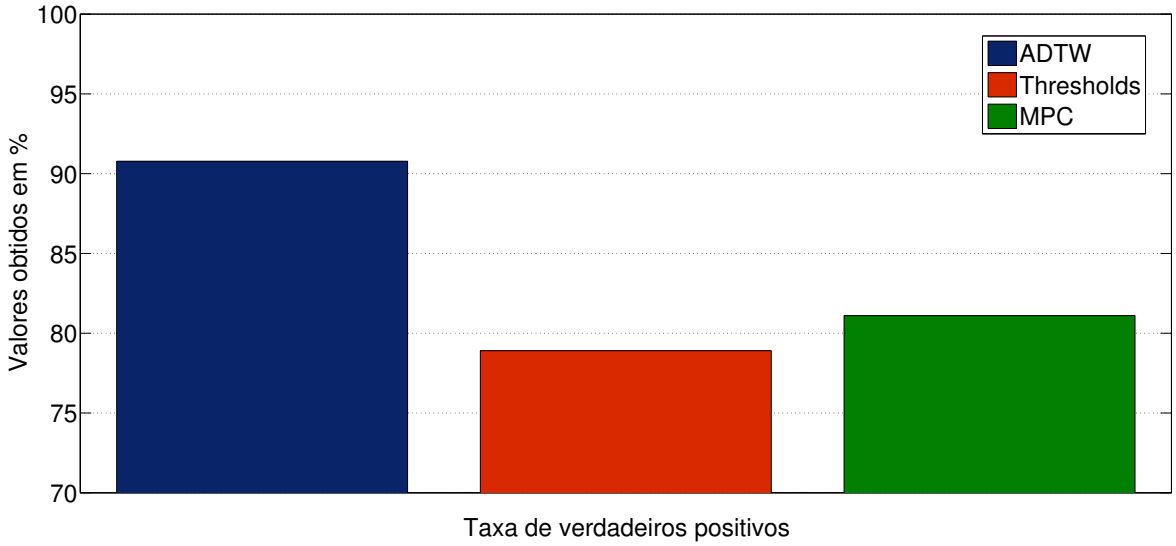

Figura 35 – Taxas de verdadeiro positivos (TPR)

Já na Figura 36 as taxas mais baixas são as ideais, pois apontariam as vezes que um tráfego foi identificado de forma incorreta. Nesta imagem observam-se valores muito baixos para a MPC, representada na cor verde, e valores de quase 1% para o ADTW, o que leva a inferir um melhor aproveitamento para a MPC. É importante levar

em consideração para este resultado a quantidade de alarmes gerado para cada um dos métodos, onde quanto menor a quantidade de alarmes gerados, menor a probabilidade de erros nesta predição.

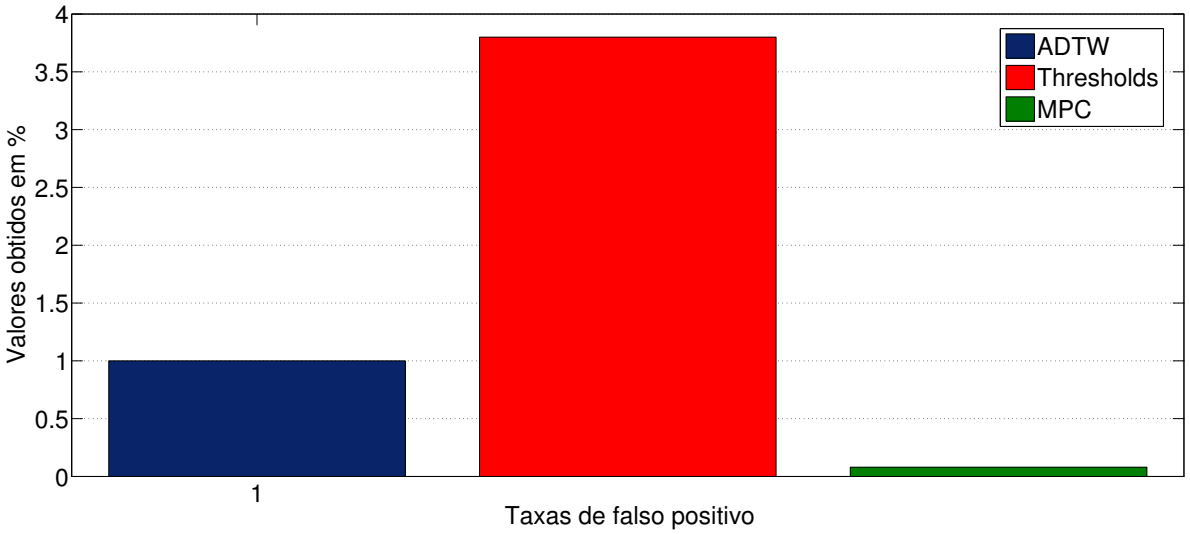

Figura 36 – Taxas de falso positivos (FPR)

A Tabela 9 apresenta um resumo dos resultados obtidos pelas métricas de desempenho, mostrando qual dos métodos se saiu melhor em relação aos outros em cada métrica. Para cada método de detecção de anomalia com resultado superior aos demais, foi colocado um X na célula correspondente e ao final é apresentada a soma de quantas métricas este método foi superior. É possível ver que o método ADTW foi superior em cinco dos sete testes efetuados e a MPC nos outros dois, sendo possível inferir uma vantagem do ADTW em relação aos demais para o cenário proposto.

Tabela 9 – Resumo das métricas de desempenho

Métrica / Método	**ADTW**	**Thresholds**	**MPC**
Acurácia	X		
Recall	X		
Precisão			X
Especificidade	X		
F-measure	X		
TPR	X		
FPR			X
Total	**5**	**0**	**2**

Dadas as particularidades de cada método de detecção de anomalias, a vantagem do ADTW pode ser justificada pela sua complexidade em avaliar as tendências de crescimento e de diminuição do tráfego de rede, e não apenas avaliar se os valores absolutos de tráfego e limiares são maiores ou menores entre si.

4.2 Cenário 2 - UTFPR

O segundo cenário mostra o comportamento do GADS na rede da Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Toledo. Os dados desta rede foram adquiridos usando o formato Netflow. A coleta do tráfego da UTFPR é realizada no gateway principal da rede, que comporta aproximadamente quatrocentos dispositivos e é exportada no formato Netflow v9. Os dados foram coletados entre os meses de abril e maio do ano e 2013, sendo submetidos aos mesmos testes do cenário anterior.

Os resultados serão apresentados em dois grupos novamente, conforme descrito na seção 3.2 , sendo um contendo os atributos quantitativos bits e pacotes e outro com os atributos qualitativos, sendo estes as entropias de IPs e portas de origem e destino.

A apresentação deste cenário tem o objetivo de ilustrar um ambiente diferente do anterior no que tange tanto a dimensão deste, quanto a forma de aquisição dos dados. O ambiente da UTFPR campus Toledo tem uma quantidade de máquinas e usuários menor em relação à UEL, comportando aproximadamente quatrocentos dispositivos frente aos quase dez mil da UEL. A forma como os dados foram adquiridos também difere em ambas, uma vez que a UEL usou o padrão sFlow, que usa amostragem dos fluxos, e a UTFPR usou o padrão Netflow v9 que não usa amostragem.

A transformação dos dados binários em texto segue o mesmo padrão do primeiro cenário. Com isso o algoritmo usado para a geração das assinaturas digitais segue sendo exatamente o mesmo, bem como as métricas usadas para aferição do modelo GADS. Apenas os dias úteis foram considerados e foi usada a janela das 24h de cada dia.

Para a detecção de anomalias foram injetadas anomalias sintéticas com o software Scorpius durante uma semana. Após esse procedimento, os dados foram submetidos às análises relativas aos atributos de dados qualitativos, ou seja, a acurácia, a *f-measure*, o *recall*, a precisão e a especificidade. Para os dados quantitativos serão apresentados os cálculos de Coeficiente de Correlação (CC).

4.2.1 DSNSFs gerados para UTFPR

Para este cenário, o GADS gerou as assinaturas digitais para três semanas, compreendidas entre os dias 29 de abril e 17 de maio de 2013. Para a geração das assinaturas, foram novamente usados os mesmos dias da semana nas três semanas anteriores a este dia. Na Figura 37 é possível analisar o comportamento dos dias 6 e 7 de maio de 2013 na UTFPR campus Toledo, onde para a geração da assinatura do dia 6, foram usados os dias 15, 22 e 29 de abril daquele ano, e para o dia 7 de maio, foram usados os dias 16, 23 e 30 de abril. Na figura o tráfego real do dia é representado pela cor verde e o DSNSF pela cor vermelha.

(a) Dia 6 de maio de 2013 (b) Dia 7 de maio de 2013

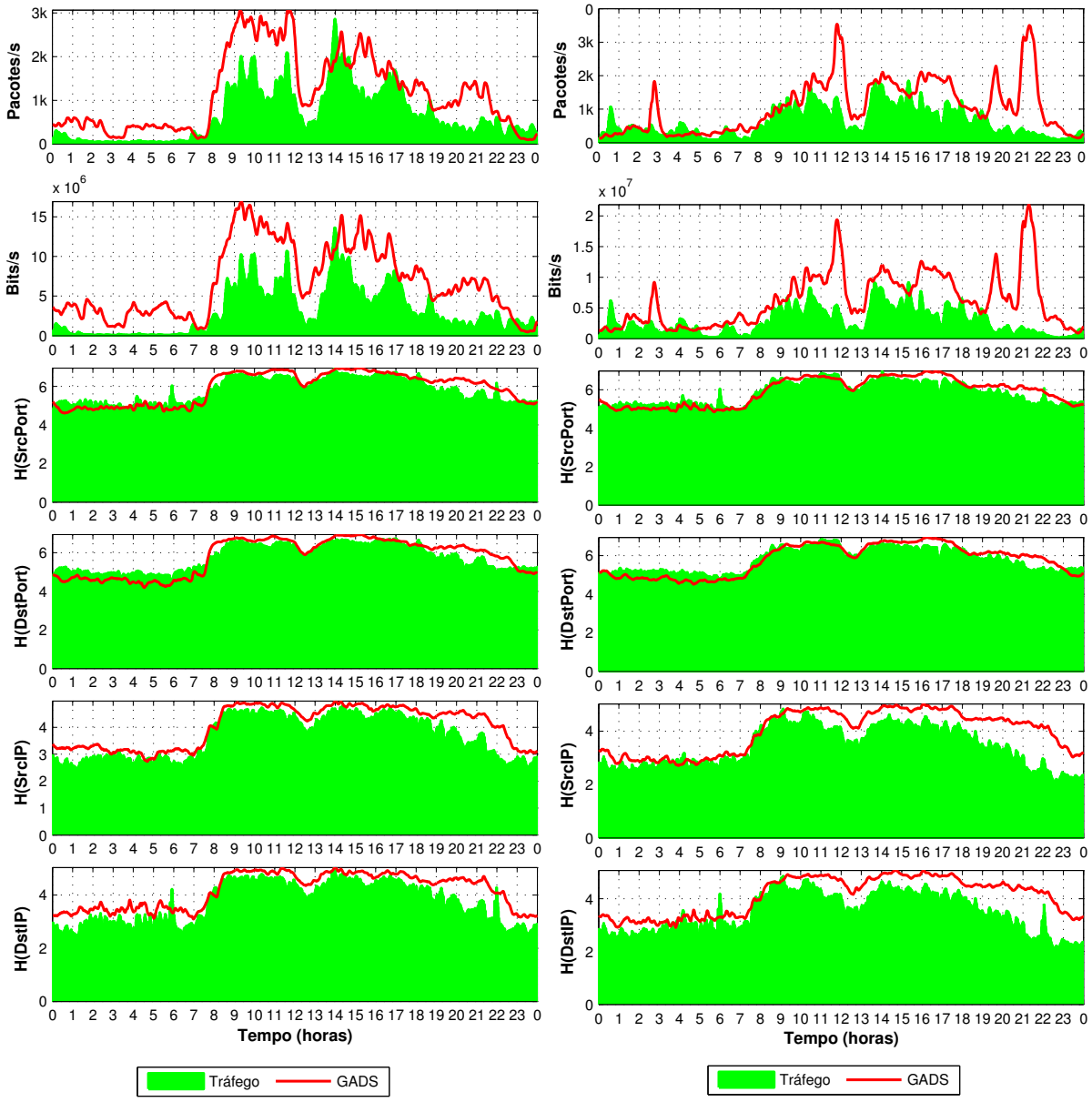

Figura 37 – DSNSF e movimentos do tráfego da UTFPR para os dias 6 e 7 de maio de 2013

Como a quantidade de dados trafegada é consideravelmente menor se comparada à UEL, quaisquer oscilações no tráfego são facilmente notadas. Além disso, se houve uma anormalidade, como um dia em que poucos usuários geraram um tráfego acima do normal em apenas algum momento, essa alteração irá interferir no aprendizado do GADS. Da mesma forma, apenas um dia que por qualquer motivo os acessos foram abaixo do esperado, também pode fazer com que a assinatura digital, nas semanas subsequentes, seja alterada de forma mais contundente.

As Figuras 38 e 39, apresentam as assinaturas geradas pelo GADS entre os dias 7 a 10 de maio de 2013 para o tráfego da UTFPR campus Toledo. É possível observar nas figuras que o DSNSF acompanha na maior parte do dia o movimento do tráfego real, mas novamente pela pouca quantidade de dados trafegada, quaisquer oscilações no tráfego são sentidas para os atributos de tráfego bits e pacotes por segundo. Já as entropias seguem um movimento mais uniforme.

(a) Dia 8 de maio de 2013 (b) Dia 9 de maio de 2013

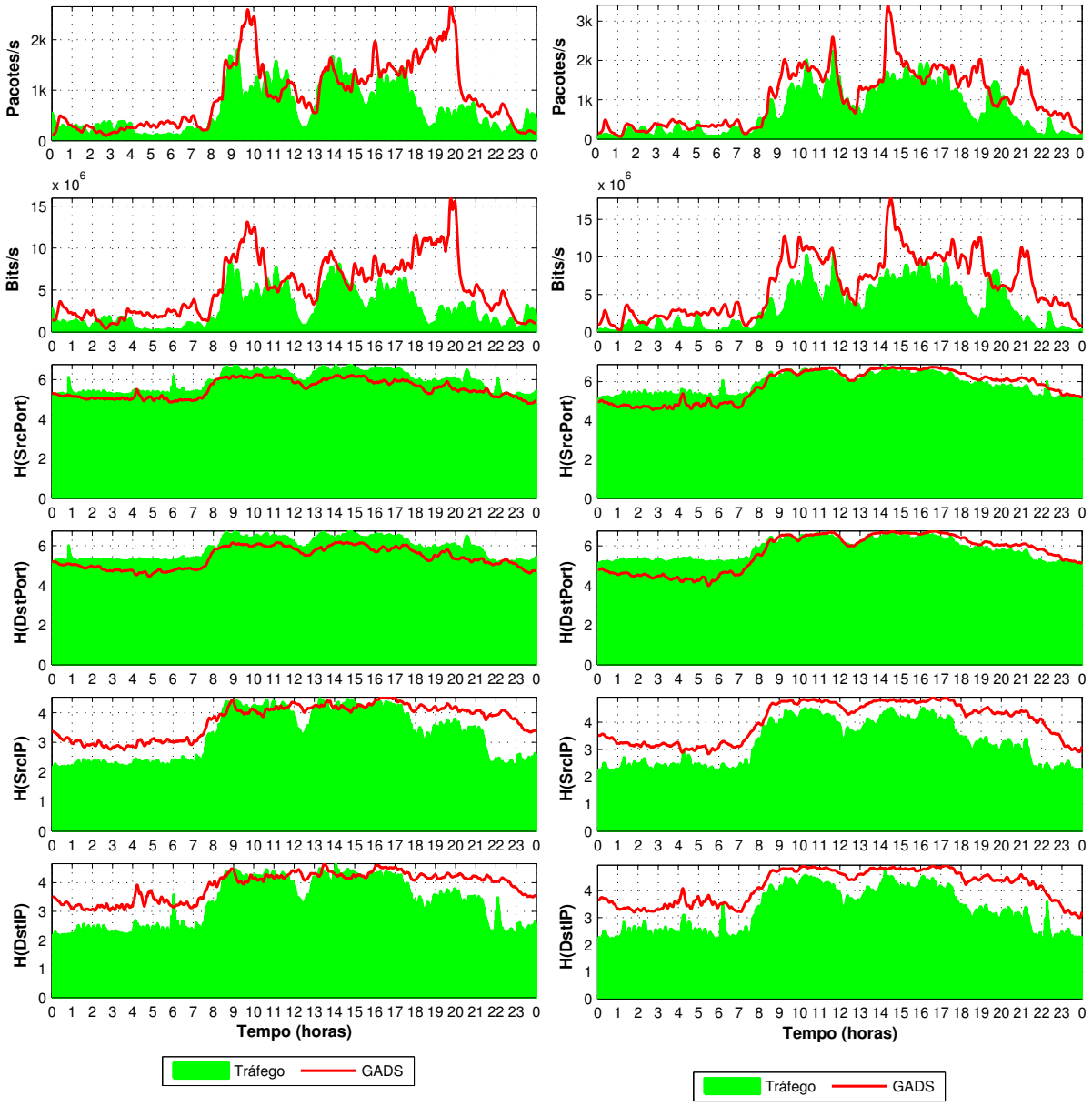

Figura 38 – DSNSF e movimentos do tráfego da UTFPR para os dias 6 e 7 de maio de 2013

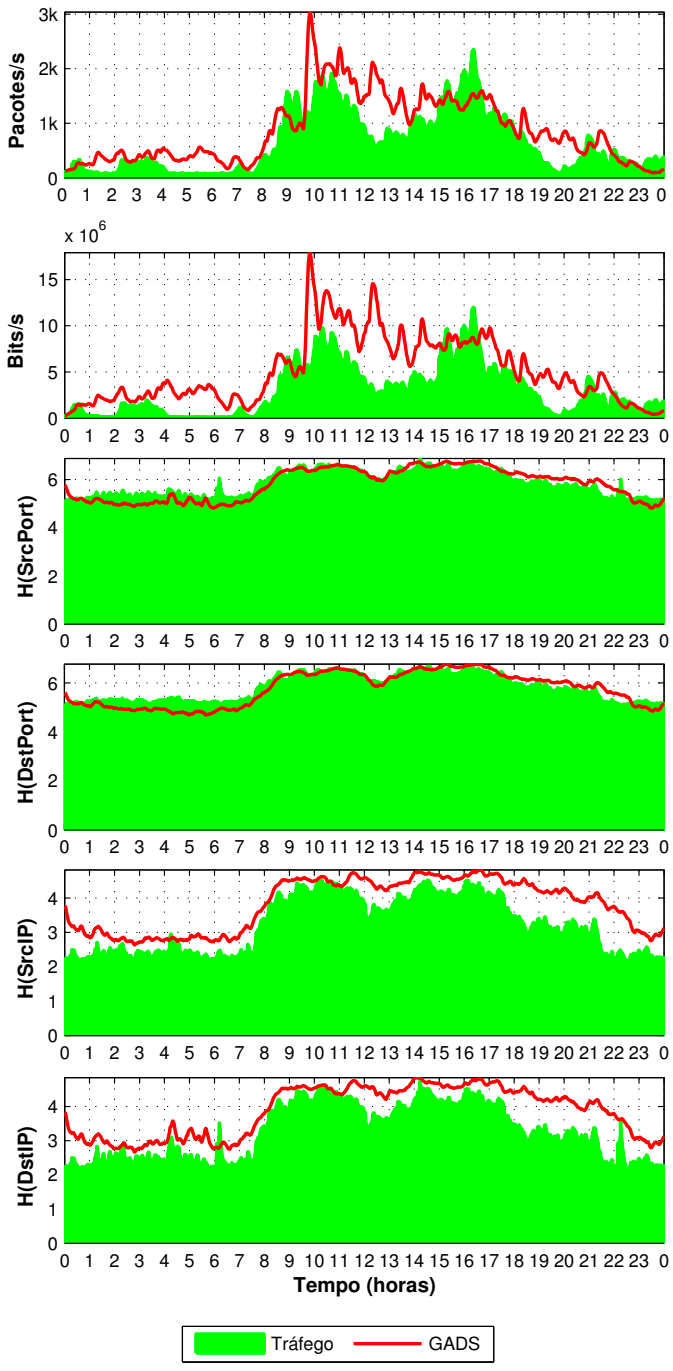

Figura 39 – DSNSF para o dia 10 de maio de 2013 para a UTFPR Toledo

4.2.2 Resultados para CC na UTFPR

Os atributos quantitativos, bits e pacotes foram submetidos aos cálculos do coeficiente de correlação (CC) para mensurar quanto os dados estão ou não correlacionados. Os resultados são apresentados nas Figuras 40 e 41, onde é possível notar valores não tão ideiais quanto os encontrados no primeiro cenário, com uma média de 0.45 para bits e 0.5 para pacotes incluindo o feriado de primeiro de maio (dia do trabalhador) e de 0.48 para bits e 0.54 excluindo o feriado.

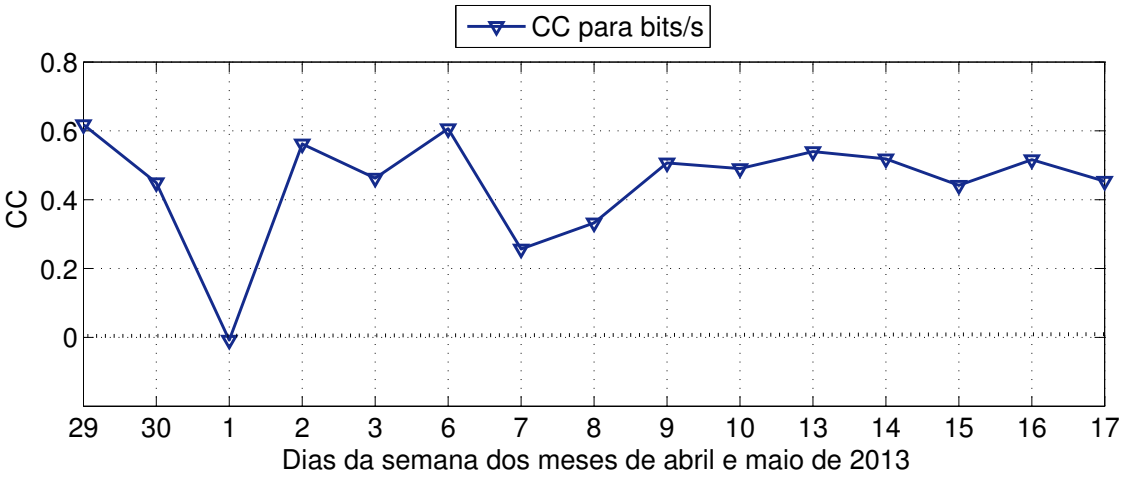

Figura 40 – CC para bits para os dias 29 e 30 de abril e 1 a 17 de maio de 2013 para a UTFPR Toledo

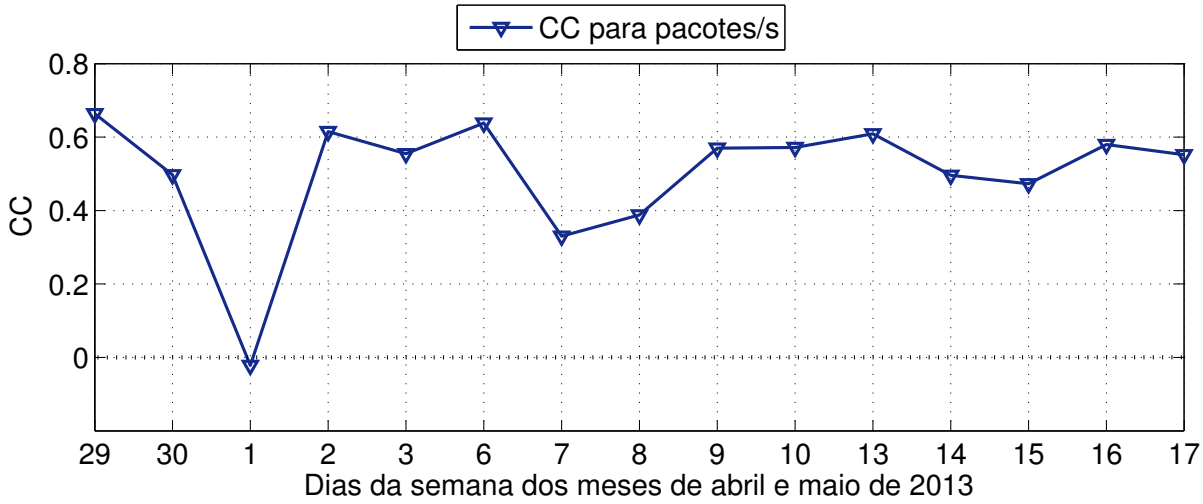

Figura 41 – CC para pacotes para os dias 29 e 30 de abril e 1 a 17 de maio de 2013 para a UTFPR Toledo

Os valores de correlação médios encontrados nas assinaturas para os atributos bits e pacotes não irão interferir na detecção de anomalias, onde como já exposto, o alarme só é gerado caso haja uma anomalia detectada em três ou mais atributos e serão usados os seis atributos. A detecção de anomalias será abordada na próxima seção.

4.2.3 Detecção de anomalias na UTFPR Toledo com ADTW

Para aferir a capacidade de detecção de anomalias do GADS para o cenário da UTFPR no campus de Toledo, as assinaturas serão usadas com a técnica de detecção que se saiu melhor na comparação do primeiro cenário, o *Adaptive Dynamic Time Warping* (ADTW), que identifica padrões de comportamento alinhando o tráfego real e o DSNSF, na busca de divergências entre os padrões para a identificação de anomalias no tráfego.

Serão submetidos ao ADTW as assinaturas digitais e o tráfego real dos seis atributos, coletados no tráfego de rede da UTFPR durante uma semana, compreendida entre os dias 6 a 10 de maio de 2013. Foram injetadas anomalias sintéticas nos arquivos de fluxos de dados sistematicamente durante essa semana, para que os resultados pudessem ser submetidos às métricas de desempenho usadas neste trabalho que são a acurácia, a *f-measure*, o *recall*, a precisão e a especificidade obtidas através do cálculo da matriz de confusão.

A Figura 42 ilustra todas as cinco métricas calculadas para cada um dos cinco dias da semana analisados.

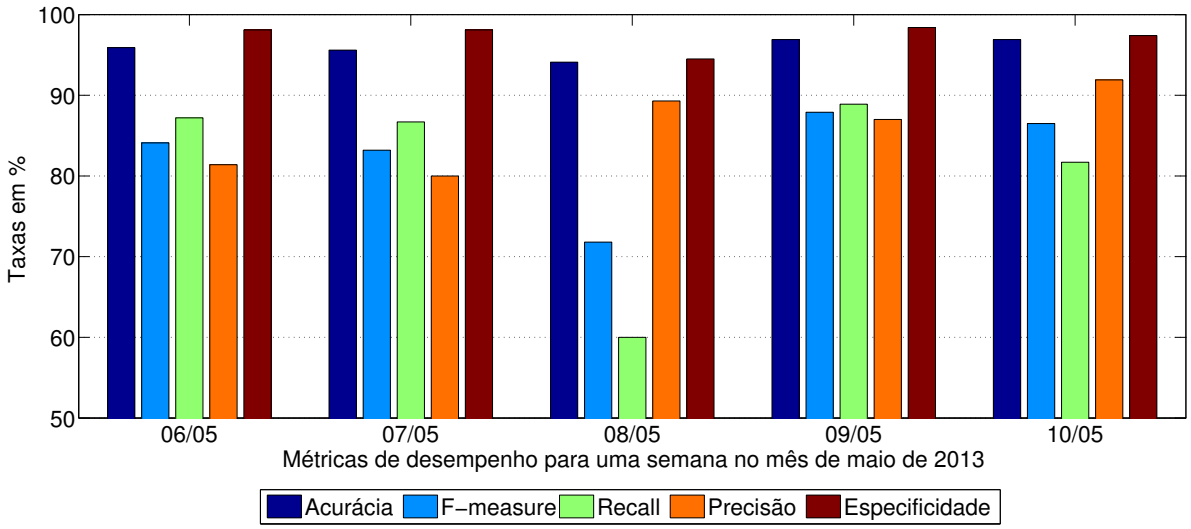

Figura 42 – Métricas de desempenho para o ADTW na UTFPR Toledo

É possível notar valores acima de 94% para a acurácia em todos os dias, o que pode ser considerado um ótimo desempenho. Além da acurácia, são encontrados bons valores no *recall*, indicando que a taxa de acerto de anomalias, quando estão acontecendo, está acima de 85% na maior parte dos dias, obtendo uma média de 80.1% na semana. As taxas de precisão e especificidade também tem bons valores, obtendo médias de 85.9% e 97.3% respectivamente. Já a taxa de *f-measure* obteve média de 82.7% nesta semana.

A Figura 43 representa os alarmes gerados pelo ADTW para o dia 9 de maio de 2013.

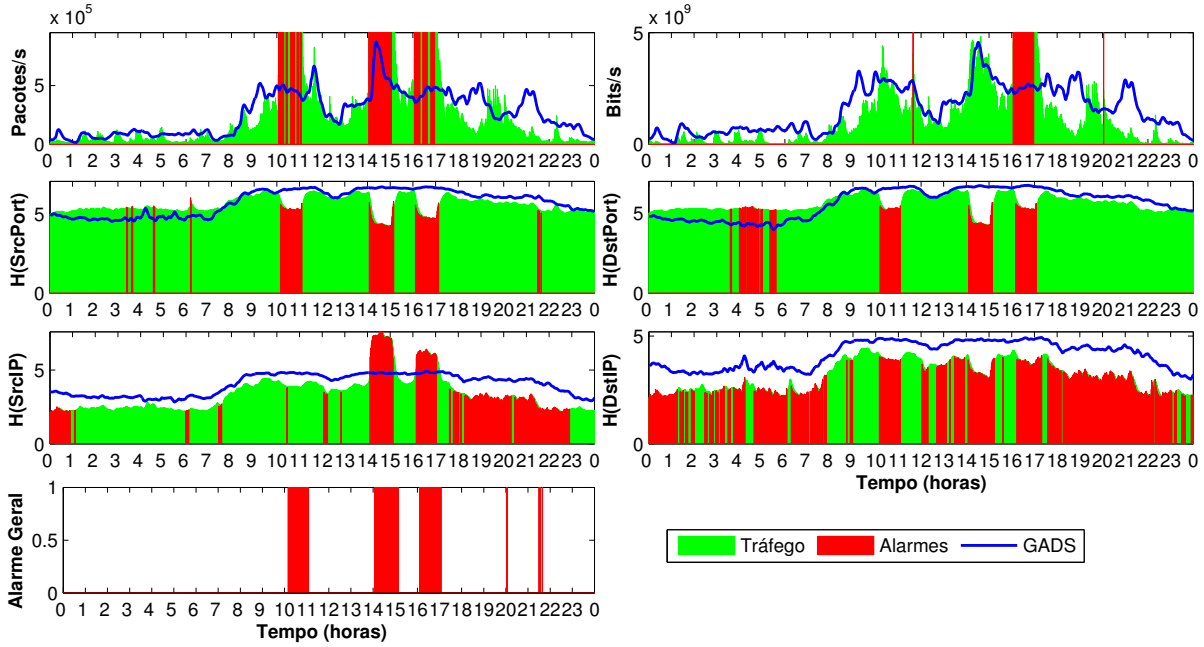

Figura 43 – Alarmes gerados para o ADTW para o dia 9 de maio de 2013 na UTFPR campus Toledo

Nela é possível notar como os ataques simulados puderam ser notados graças à assinatura digital gerada pelo GADS. Notam-se nas figuras na cor verde o tráfego real, a linha na cor azul representa o DSNSF e na cor vermelha os alarmes gerados naquele intervalo de tempo, quando o ADTW identificou diferenças nos padrões de comportamento entre o tráfego real e a assinatura digital. É possível notar que apesar de identificar como anômalo um grande período de tempo nas entropias de IP e porta de origem, no quadro "Alarme geral" apenas são classificados como intervalos anômalos os períodos que houve ao menos três alarmes, ou seja, em três atributos diferentes. Neste quadro são identificados os três ataques realizados e são gerados mais três alarmes, às 20h e aproximadamente às 21:30h, que foram considerados como falsos positivos.

Assim como no cenário anterior, o ADTW foi capaz de identificar de forma eficaz o tráfego anômalo injetado através do Scorpius, obtendo boas taxas de verdadeiro positivo (TPR) com 80.1% e uma ótima taxa de falso positivo (FPR) com 2%.

4.3 Cenário 3 - Comparação com modelo ACODS usando dados adquiridos em 2015

O terceiro cenário utiliza dados mais recentes obtidos na rede da Universidade Estadual de Londrina no ano de 2015 e compara a abordagem proposta nesse trabalho, usando o Algoritmo Genético, com o modelo ACODS (*Ant Colony Optimization for Digital Signature*), que também usa assinaturas digitais para fazer a caracterização do tráfego.

As assinaturas geradas pelos modelos ACODS e GADS foram utilizadas na detecção de anomalias usando o método ADTW, empregando a mesma metodologia usada neste trabalho, onde os seis atributos dos fluxos de dados são analisados separadamente e se uma anomalia for detectada no mesmo intervalo de tempo em ao menos três atributos, então é gerado um alarme para o administrador da rede.

Os registros de fluxos IP utilizados neste cenário foram obtidos no mês de outubro do ano de 2015 e serão apresentados os resultados da semana compreendida entre os dias 26 a 30 de outubro do mesmo ano. A inclusão das anomalias sintéticas seguiu a mesma abordagem do primeiro cenário, sendo usado o dia 28 de outubro para a injeção do tráfego anômalo e posterior análise.

4.3.1 DSNSF de GADS e ACODS para UEL no ano de 2015

É possível notar na Figura 44 o movimento do tráfego real e os DSNSF's gerados pelo GADS e pelo ACODS para os dias 26 e 27 de outubro de 2015.

Na figura observa-se o tráfego representando a assinatura gerada pelo GADS na cor azul, enquanto a assinatura gerada pelo ACODS é visualizada na cor vermelha. Notam-se que ambas as assinaturas têm comportamentos muito parecidos, acompanhando os movimentos de tráfego real durante os períodos de maior e menor tráfego em cada dia, sem que haja um distanciamento entre ambas assinaturas em relação uma a outra e também em relação ao tráfego real.

Já a Figura 45 apresenta os DSNSF's gerados pelo GADS e ACODS para os dias 28 e 29 de outubro de 2015. E na Figura 46 os mesmos resultados para o dia 30 de outubro de 2015. Mais uma vez são observados valores bastante semelhantes para ambas as assinaturas digitais, com ambas acompanhando o movimento do tráfego real para cada um dos dias apresentados.

(a) Dia 26 de outubro de 2015 (b) Dia 27 de outubro de 2015

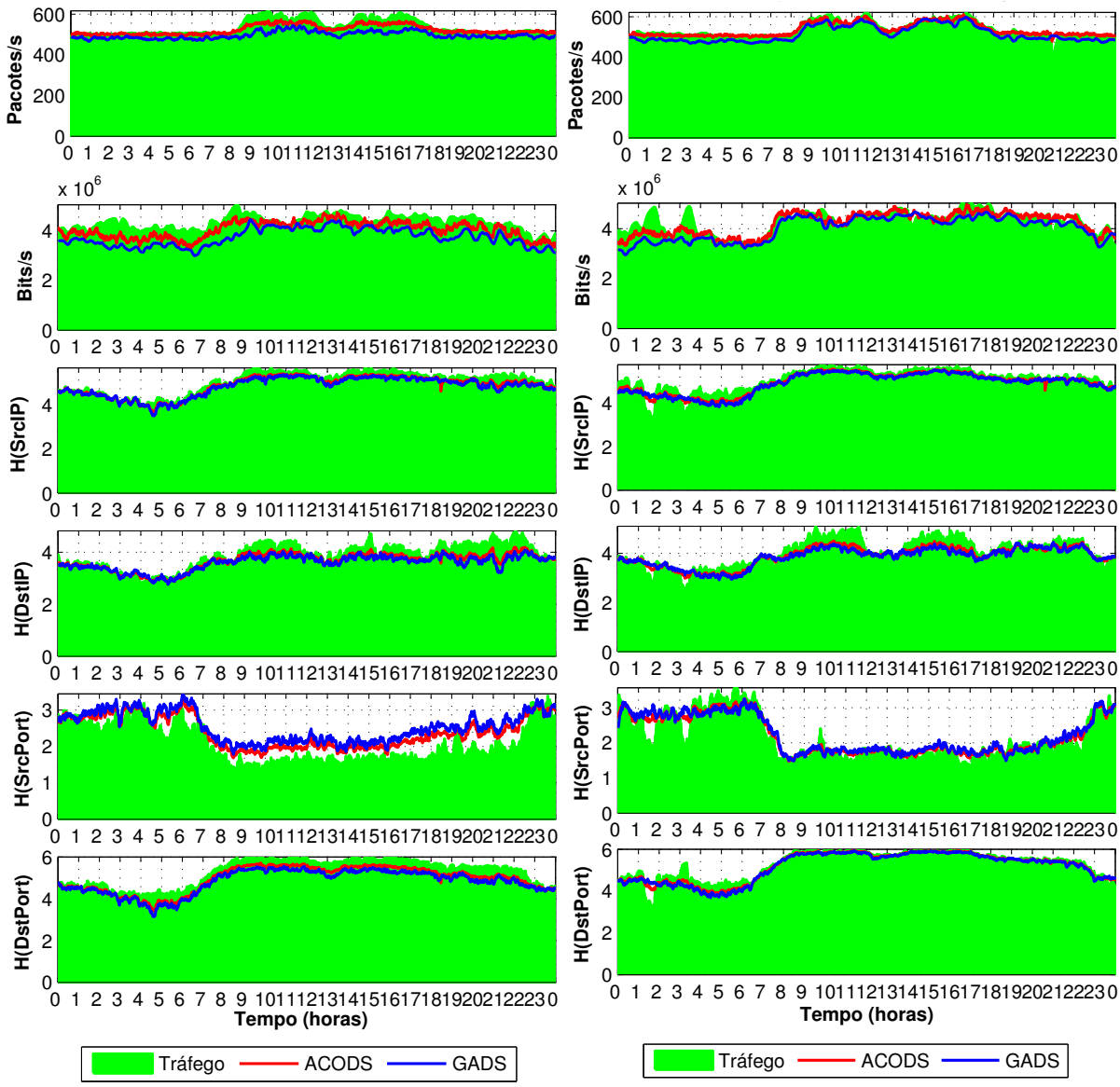

Figura 44 – DSNSF e movimento do tráfego dos dias 26 e 27 de outubro de 2015 para UEL

(a) Dia 28 de outubro de 2015 (b) Dia 29 de outubro de 2015

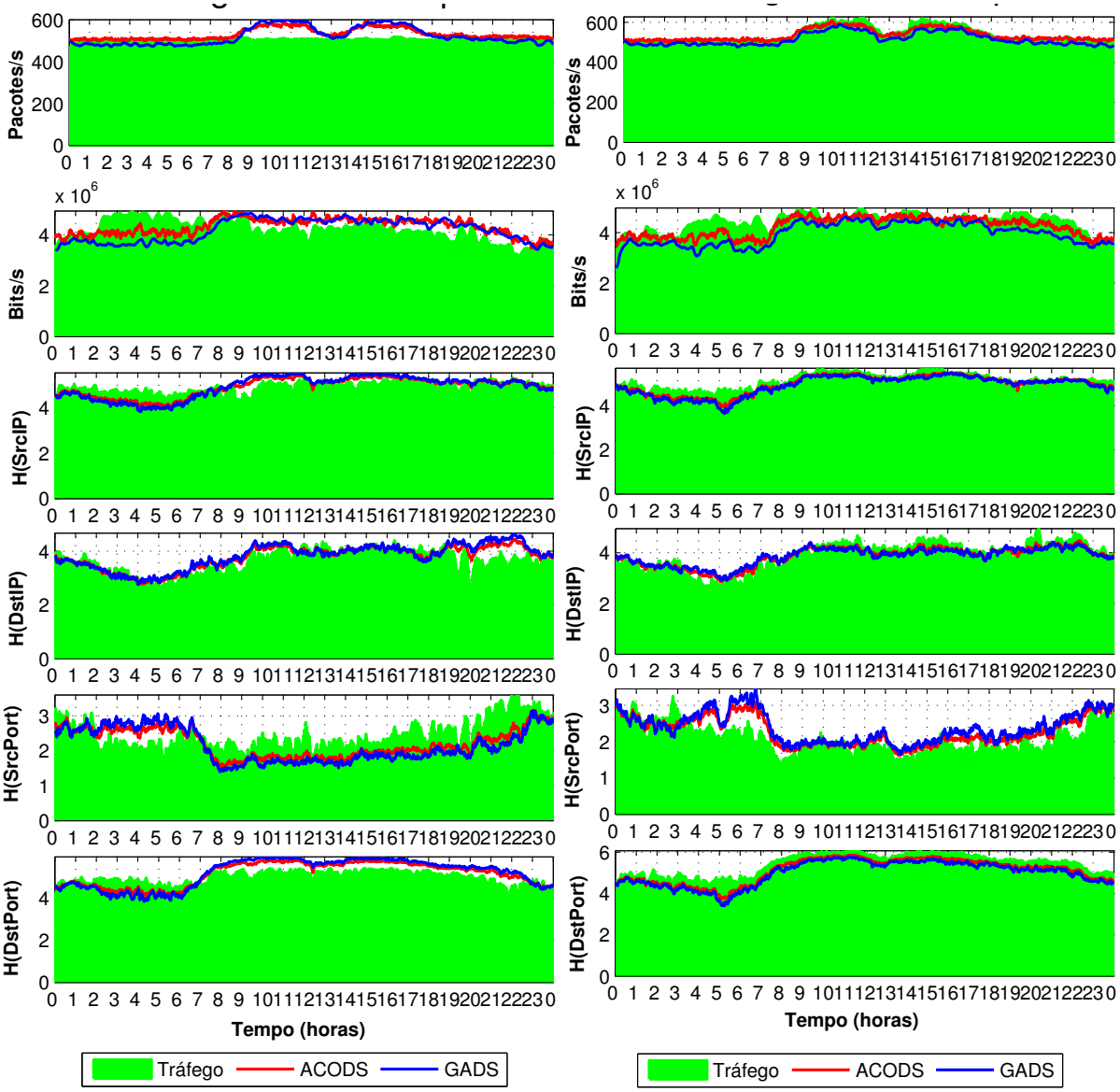

Figura 45 – DSNSF e movimento do tráfego dos dias 28 e 29 de outubro de 2015 para UEL

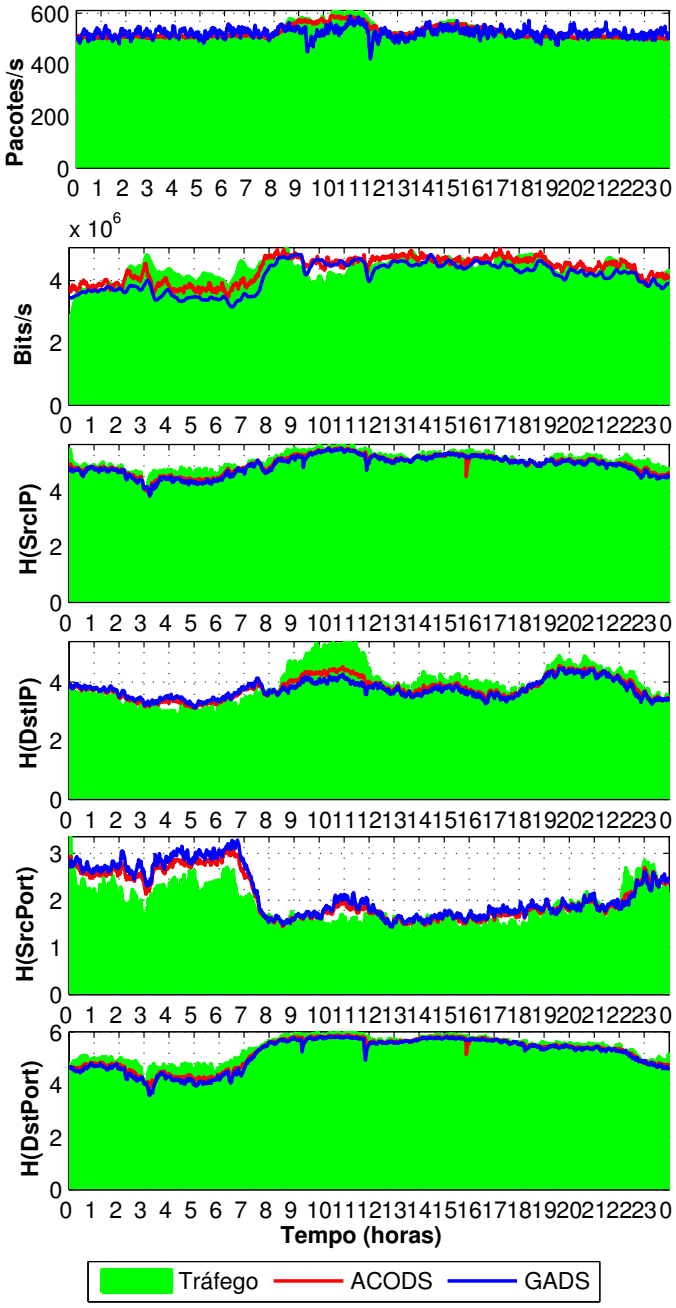

Figura 46 – DSNSF e movimento do tráfego do dia 30 de outubro de 2015

4.3.2 NMSE e CC para ACODS e GADS

Para verificar a habilidade de ambos os modelos na predição do tráfego, as duas assinaturas digitais foram submetidas aos testes de coeficiente de correlação (CC) e erro quadrático médio normalizado (NMSE) para os atributos quantitativos, ou seja, bits e pacotes por segundo.

Os resultados para o erro quadrático médio normalizado (NMSE) para a última semana do mês de outubro de 2015 são apresentados na Figura 47, onde é possível observar excelentes valores para ambos os modelos, com taxas para bits/s abaixo de 0.04 para todos

os dias nos dois modelos e taxas abaixo de 0.02 para pacotes/s também em ambos modelos em todos os dias. Quanto mais próximos de zero, mais a predição do tráfego se aproximou do tráfego real naquele dia, e como todos os valores foram próximos de zero é possível inferir que a predição dos dois modelos foi bem realizada.

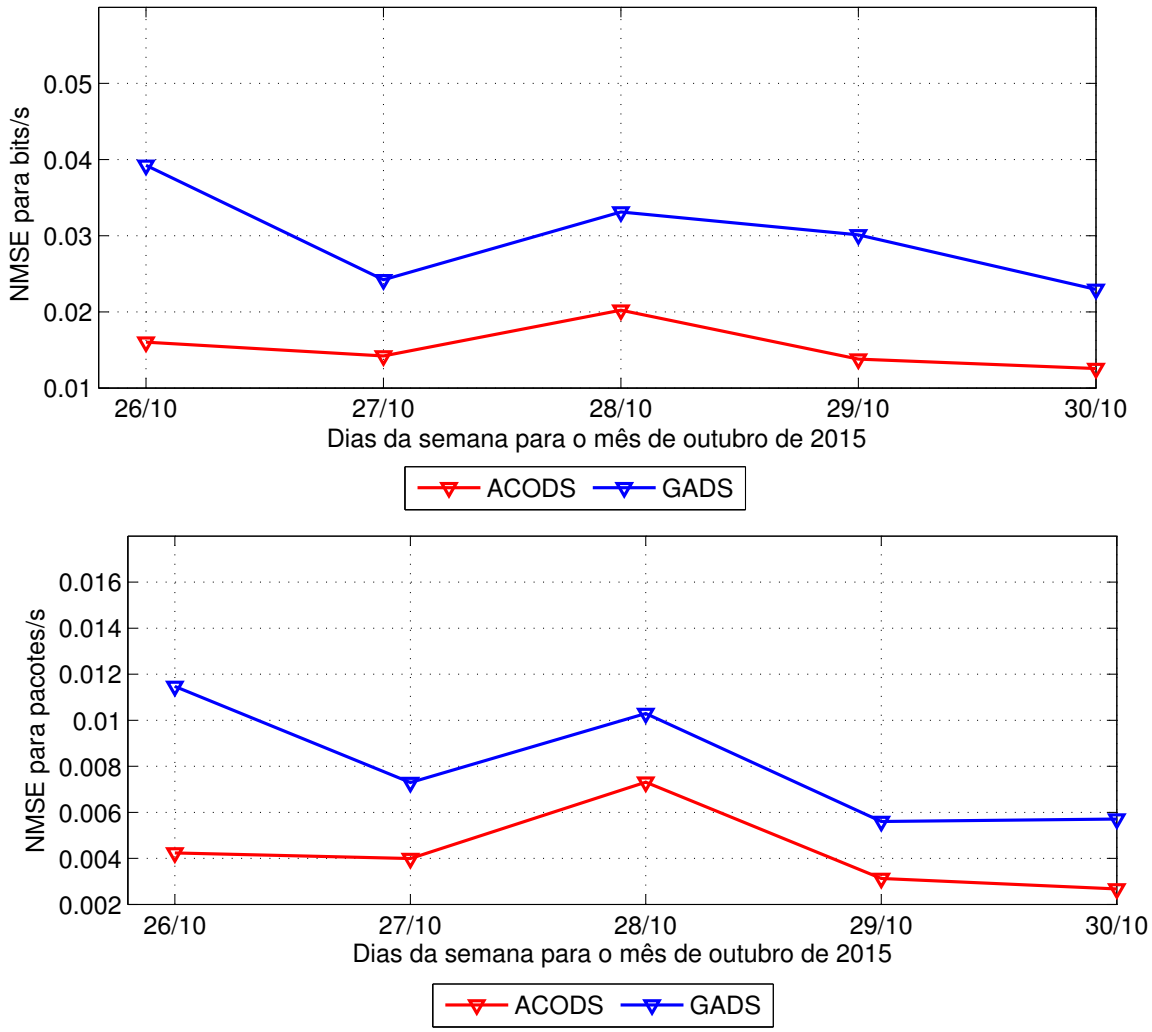

Figura 47 – NMSE para o mês de outubro de 2015 para ACODS e GADS para bits/s e pacotes/s

Os resultados para os coeficientes de correlação (CC) da mesma semana são ilustrados na Figura 48, onde é possível notar novamente valores muito próximos se comparados os dois modelos, tanto na medição para bits quanto para pacotes por segundo. É possível observar que os resultados não foram tão precisos quanto no primeiro cenário, principalmente por atividades que aconteceram ao longo das madrugadas, como backups. Mas apesar dessa queda nos valores de CC os resultados em geral continuam satisfatórios.

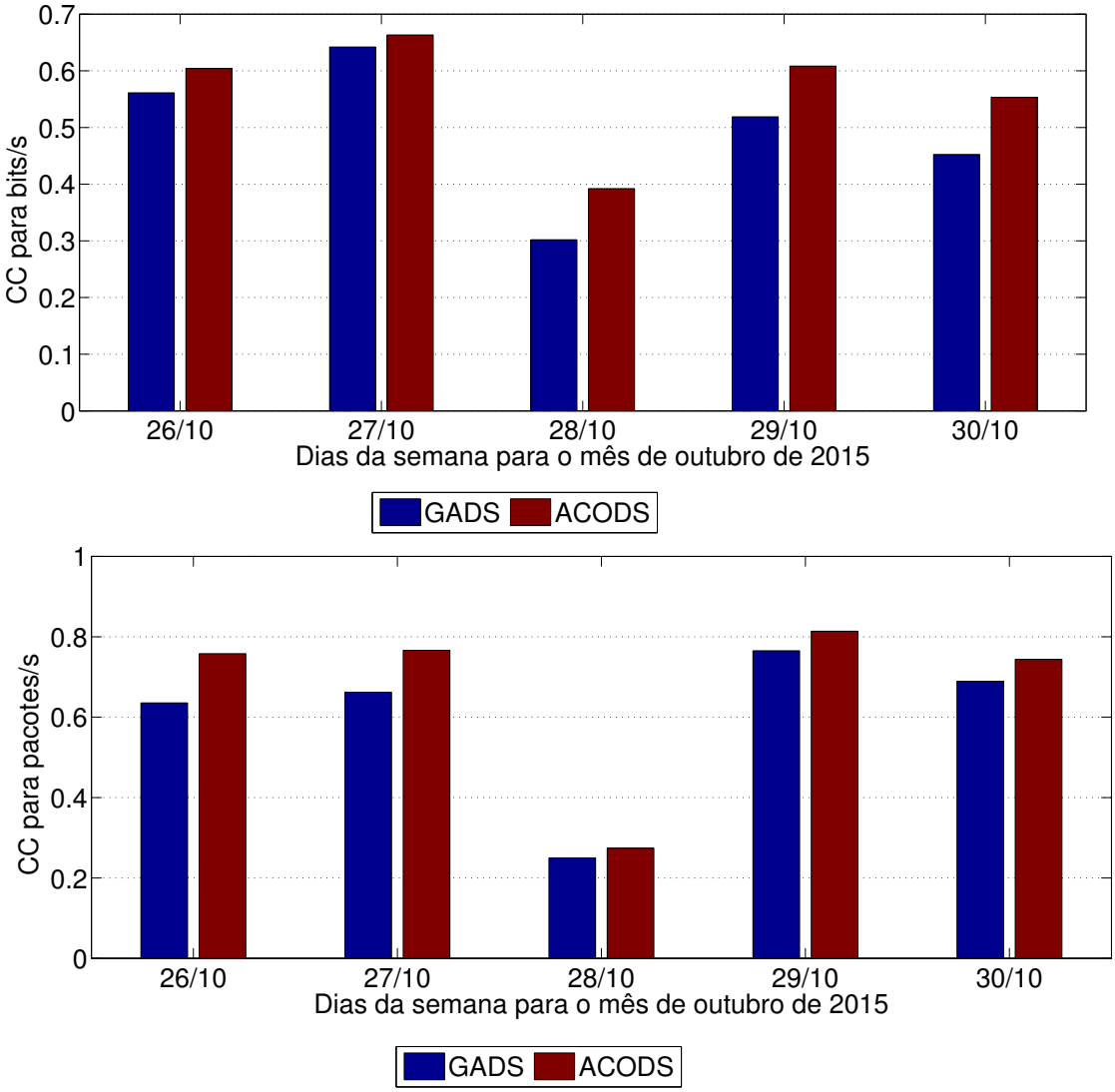

Figura 48 – CC para o mês de outubro de 2015 para ACODS e GADS para bits/s e pacotes/s

4.3.3 Detecção de anomalias para ACODS e GADS

Assim como no primeiro cenário, foi injetado tráfego anômalo em um dia, para que este comportamento pudesse ser analisado. Novamente foram injetados três tipos de anomalias, sendo um ataque de DoS entre as 10h e as 11h, um ataque de DDoS entre as 14h e as 15h e uma anomalia do tipo *flash crowd* entre as 16h e as 17h. As métricas usadas para analisar o desempenho da detecção de anomalias foram as mesmas dos outros cenários. O dia cujo tráfego foi analisado foi o dia 28 de outubro de 2015 e o método empregado foi novamente o ADTW, pela sua boa performance nos outros cenários. Sob essas condições, foram submetidos à análise as assinaturas digitais geradas pelo ACODS e pelo GADS.

Os resultados da detecção de anomalias usando o ADTW para o ACODS são apresentados na Figura 49, onde é possível ver quais alarmes foram gerados para cada

um dos atributos de fluxos de dados empregado na análise, que estão apresentados na cor vermelha. Pode-se notar a pequena quantidade de alarmes fora dos intervalos em que comprovadamente houve uma anomalia, intervalos estes que foram responsáveis pela geração de alarmes ao administrador, como pode-se notar na caixa “Alarme Geral”.

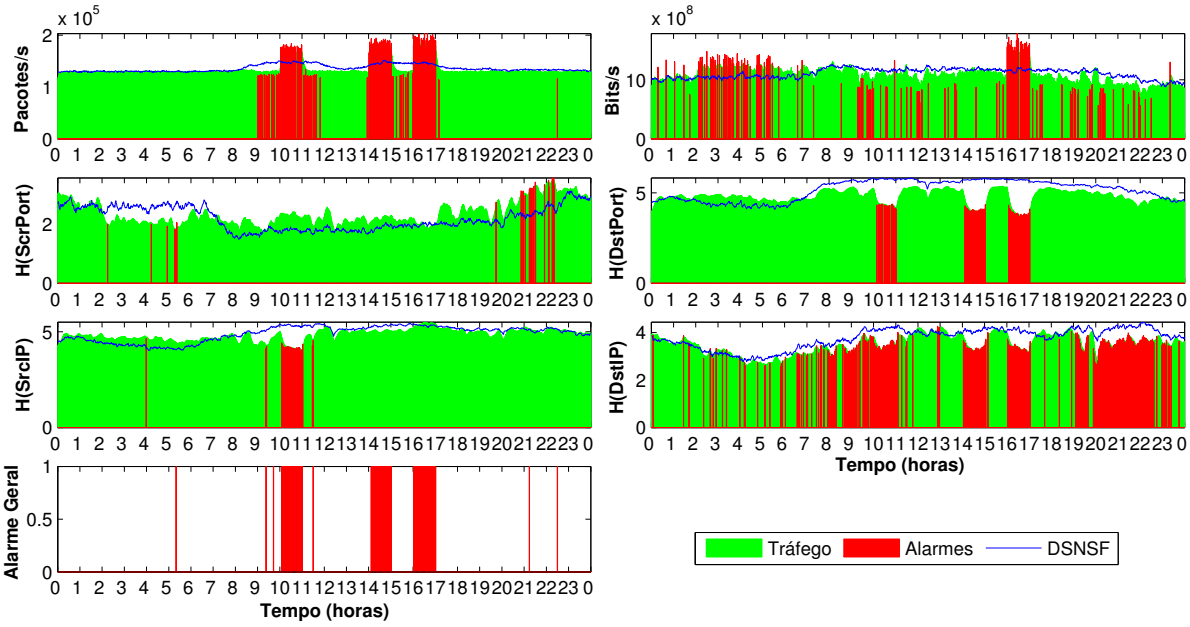

Figura 49 – Alarmes gerados pelo ADTW para ACODS

Na Figura 50 são apresentados os resultados dos alarmes para o GADS. Nesta figura também notam-se alarmes gerados durante os intervalos que sabidamente são anômalos em cada um dos atributos analisados. Na caixa “Alarme Geral” são apresentados os alarmes enviados ao administrador da rede, ou seja, aqueles em que três ou mais atributos tiveram um tráfego anômalo no mesmo intervalo de tempo.

Na Figura 51 é possível analisar cada uma das métricas de desempenho empregadas nesse trabalho, para cada um dos métodos de geração de assinaturas digitais deste cenário. Ambos os métodos tiveram altas taxas de acurácia, acima de 95%, o que os classifica como bastante eficientes. As taxas obtidas em todas as outras métricas também foram satisfatórias, onde os dois métodos obtiveram índices muito próximos, com ligeira vantagem para o ACODS.

Ambos os algoritmos usam a clusterização como método para a classificação dos dados adquiridos. Analisando o algoritmo ACODS e o comparando com o GADS é possível inferir a pequena vantagem de desempenho pela sua característica em descartar os clusters com pouca representatividade, o que faz com que as assinaturas digitais não tenham tanta oscilação, visto que ambos utilizam o valor da média dos centros dos clusters para cada ponto dentro do DSNSF. Outra característica que difere em ambos é o fato de o ACODS

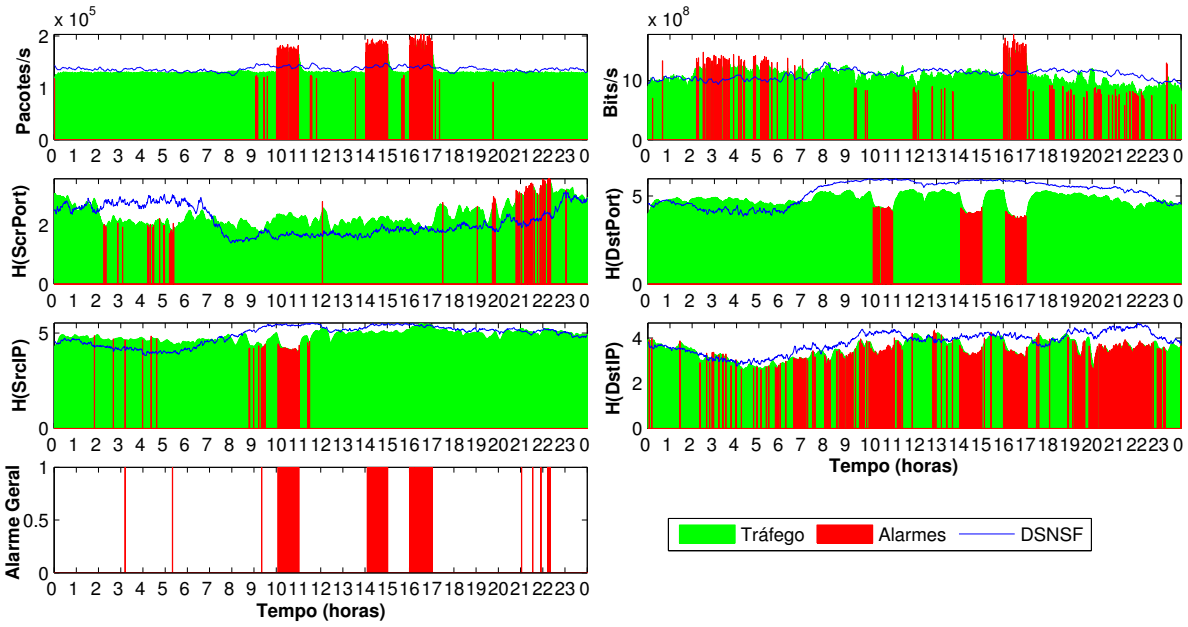

Figura 50 – Alarmes gerados pelo ADTW para GADS

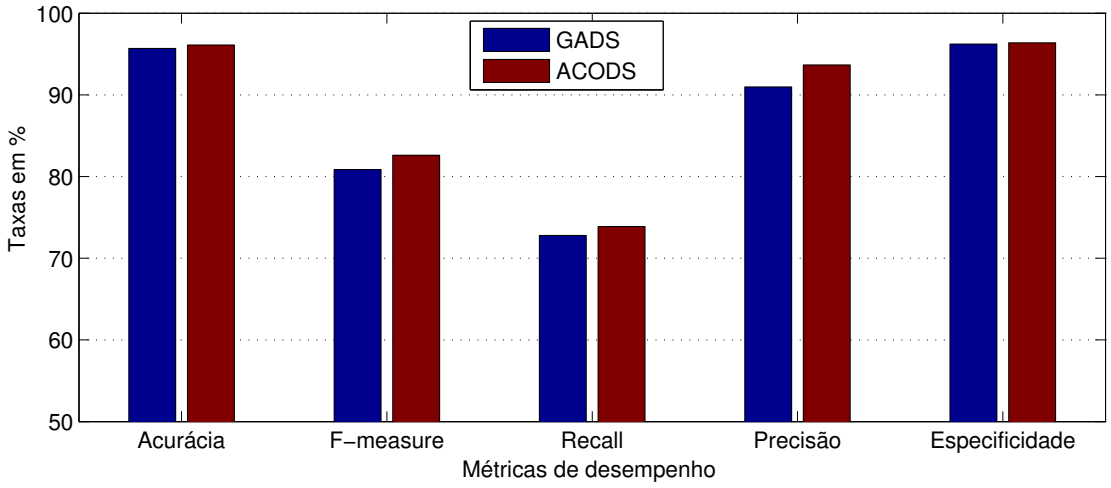

Figura 51 – Métricas de desempenho para ACODS e GADS

utilizar a média ponderada, enquanto o GADS usa a média aritmética dos centróides.

4.4 Conclusão do capítulo

O objetivo dos testes realizados foi provar a capacidade do GADS em fazer de forma eficiente a caracterização do tráfego, além de destacar a eficiência deste na detecção de anomalias em redes.

O primeiro cenário apresentou a rede da Universidade Estadual de Londrina (UEL) utilizando dados adquiridos no ano de 2013, onde foi possível notar que a

caracterização do tráfego foi feita de forma bastante eficiente. Para a detecção de anomalias foram testadas três técnicas, todas já usadas em conjunto com o DSNSF, onde a técnica chamada de ADTW foi a mais eficiente.

No segundo cenário foram utilizados dados extraídos da UTFPR campus Toledo, uma rede cerca de vinte e cinco vezes menor que a rede do primeiro cenário. Mesmo com uma menor quantidade de dados trafegados, a caracterização do tráfego foi bem sucedida. A detecção de anomalias usou a técnica que obteve mais sucesso no cenário anterior, o ADTW, e obteve ótimos valores de acurácia, provando que o sistema acerta de forma eficiente suas previsões de comportamento.

Para o terceiro cenário foram utilizados dados da UEL adquiridos no ano de 2015. Para provar a capacidade do GADS em caracterizar o tráfego e detectar anomalias foi feita uma comparação com outra técnica de sucesso, o ACODS. Os resultados finais de ambos foram muito parecidos, mostrando a qualidade do GADS para auxiliar administradores no monitoramento de redes.

5 CONCLUSÃO

Um dos maiores desafios dos administradores de redes ainda é fazer o gerenciamento e o monitoramento de seus recursos de forma eficaz e eficiente. Com o aperfeiçoamento das tecnologias de transmissão de dados e o consequente aumento da largura de banda usada nas redes atuais, é completamente inviável fazer o monitoramento destes recursos de forma manual. Nesse âmbito, cresce cada vez mais a demanda para ferramentas que auxiliem de forma não supervisionada a gerência da rede como um todo, identificando não só ataques intencionais que comprometam a continuidade dos negócios, como quaisquer anomalias que prejudiquem o funcionamento da rede.

Este trabalho apresentou a ferramenta *Genetic Algorithm for Digital Signature* (GADS), que tem o objetivo de auxiliar estes administradores, fazendo a caracterização do tráfego em redes, apresentando uma assinatura digital deste segmento de rede, que é a previsão de qual deverá ser o comportamento normal para aquela rede naquele intervalo de tempo. Para isso foi usada uma técnica de mineração de dados conhecida como clusterização e neste processo foi usado o algoritmo genético (GA), resultando em um sistema chamado *Genetic Algorithm for Digital Signature* (GADS), que gera a assinatura digital de um segmento de redes, chamada *Digital Signature of Network Segment using Flow analysis* (DSNSF).

O GA é um algoritmo inspirado na teoria da evolução natural das espécies, simulando a evolução de uma população ao longo de suas gerações, onde a cada geração indivíduos vão se tornando mais adaptados ao meio em que vivem, para assim conseguirem sobreviver e passar os seus genes às próximas proles.

Foram realizados testes com dados reais obtidos através das técnicas de análises fluxos de dados. Estes dados foram coletados na Universidade Estadual de Londrina (UEL) e na Universidade Tecnológica Federal do Paraná (UTFPR), campus de Toledo. Foi utilizada a janela de 1 minuto para cada um dos pontos do DSNSF, que é um tempo suficiente para que o impacto de uma anomalia seja reduzido.

Além da caracterização do tráfego, foram usadas três técnicas distintas para a detecção de anomalias nestes cenários, todas já empregadas pelo grupo de pesquisa em redes de computadores e comunicação de dados do departamento de computação da UEL, para que suas habilidades na detecção de anomalias pudessem ser avaliadas. A primeira foi o ADTW, uma adaptação da técnica usada para a comparação de padrões de reconhecimento de voz. Depois foi abordada a técnica de uso de limiares (*thresholds*) simples, para limitar o tráfego normal do anômalo. A última foi a Lógica Paraconsistente

(LP), através de uma ferramenta que fornece um panorama das incertezas e contradições obtidas quando diferentes modelos são utilizados na detecção de anomalias.

Foram utilizados três cenários, que visaram apresentar a eficiência do GADS para detecção de anomalias. A rede da UEL conta com aproximadamente dez mil hosts, enquanto a rede da UTPFR apresenta um tamanho cerca de vinte e cinco vezes menor, tendo em torno de quatrocentos hosts. Além desses dois cenários, foi feita uma comparação com outro algoritmo usado para geração de assinaturas digitais, o *Ant Colony Optimization for Digital Signature* (ACODS), dessa vez utilizando dados mais recentes, obtidos na rede da UEL no ano de 2015.

Para cada um dos cenários foram analisados seis atributos coletados por meio das técnicas de análise de fluxos de dados. Esses atributos são a quantidade de bits e pacotes trafegados por segundo, a distribuição do número de portas de origem e destino, e a distribuição do número de IPs de origem e destino. O uso destes seis atributos em conjunto proporciona uma análise comportamental bastante completa do uso dos recursos da rede em cada intevalo de tempo.

As três técnicas de detecção de anomalias foram usadas com o GADS para aferição de suas habilidades em detectar comportamentos anômalos injetados de forma sintética. Todas as técnicas obtiveram taxas de acurácia acima de 93%, com leve vantagem para a técnica ADTW sobre a LP. Na análise de todas as medidas de desempenho a técnica que obteve o melhor aproveitamento foi o ADTW, motivo pelo qual esta técnica foi usada nos outros dois cenários para novos testes.

No cenário da UTFPR campus Toledo, o GADS obteve mais de 94% de acurácia na detecção de anomalias em conjunto com o ADTW, apesar de as taxas de coeficiente de correlação não indicarem valores ótimos devido à granularidade dos dados, onde por consequência da pouca quantidade de usuários na rede, é possível que apenas um usuário fazendo uso de uma maior quantidade de recursos possa interferir no resultado final e até no aprendizado do GADS. Porém mesmo com essas dificuldades, o GADS foi capaz de apresentar resultados satisfatórios.

O último cenário comparou duas heurísticas que fazem uso da clusterização para a geração de um DSNSF, o GA e o ACO. Neste cenário foram utilizados dados mais recentes, obtidos da mesma forma que os do primeiro cenário e os resultados se mantiveram muito próximos, com taxas de acurácia acima de 95% para a detecção de anomalias nos dois métodos analisados.

As contribuições deste trabalho podem ser destacadas pelo estudo da aplicação do algoritmo genético para a caracterização do tráfego e consequente detecção de anomalias em redes. Para essa finalidade foram utilizadas três técnicas distintas, o que mostra a capacidade do GADS em se adequar a diferentes métodos e a sua confiança na

predição comportamental da rede. Além disso, esse sistema foi comparado com outro similar, o ACODS, onde os resultados dos testes demonstraram sua capacidade e eficiência na detecção de anomalias.

Dessa forma o GADS alcança seus objetivos, mostrando-se capaz de auxiliar administradores de rede no uso de seus recursos, onde através de um monitoramento não supervisionado, alarmes são gerados de forma automática no caso de anormalidades, deixando os gestores da rede ocupados com problemas que mereçam realmente sua dedicação.

Dentre os trabalhos futuros estão a análise da viabilidade do uso de outros atributos dos fluxos de dados, o uso da correlação entre todos os atributos para a obtenção da assinatura digital e verificar a eficiência do sistema usando o descarte de clusters com menor representatividade, para que dados advindos de redes com menor quantidade de dados trafegados, tenham a mesma eficiência das grandes redes.

REFERÊNCIAS

[1] LAKHINA, A.; CROVELLA, M.; DIOT, C. Mining anomalies using traffic feature distributions. *SIGCOMM Comput. Commun. Rev.*, ACM, New York, NY, USA, v. 35, n. 4, p. 217–228, ago. 2005. ISSN 0146-4833. Disponível em: <http://doi.acm.org/10.1145/1090191.1080118>.

[2] CHANG, S. et al. A flow-based anomaly detection method using entropy and multiple traffic features. In: *Broadband Network and Multimedia Technology (IC-BNMT), 2010 3rd IEEE International Conference on*. [S.l.: s.n.], 2010. p. 223–227.

[3] CHANG, S. et al. A flow-based anomaly detection method using sketch and combinations of traffic features. In: *Network and Service Management (CNSM), 2010 International Conference on*. [S.l.: s.n.], 2010. p. 302–305.

[4] AHMED, M.; MAHMOOD, A. N.; HU, J. A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, v. 60, p. 19 – 31, 2016. ISSN 1084-8045. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1084804515002891>.

[5] SALAH, K.; KAHTANI, A. Performance evaluation comparison of snort nids under linux and windows server. *J. Network and Computer Applications*, v. 33, n. 1, p. 6–15, 2010.

[6] MARNERIDES, A.; SCHAEFFER-FILHO, A.; MAUTHE, A. Traffic anomaly diagnosis in internet backbone networks: A survey. *Computer Networks*, v. 73, p. 224 – 243, 2014. ISSN 1389-1286. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1389128614002850>.

[7] SHANNON, C. E. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, ACM, v. 5, n. 1, p. 3–55, 2001.

[8] CARVALHO, L. F. et al. Using ant colony optimization metaheuristic and dynamic time warping for anomaly detection. In: *Software, Telecommunications and Computer Networks (SoftCOM), 2013 21st International Conference on*. [S.l.: s.n.], 2013. p. 1–5.

[9] PENA, E. et al. Anomaly detection using digital signature of network segment with adaptive arima model and paraconsistent logic. In: *Computers and Communication (ISCC), 2014 IEEE Symposium on*. [S.l.: s.n.], 2014. p. 1–6.

[10] ASSIS, M. V. O. de; RODRIGUES, J. J. P. C.; PROENÇA JR., M. L. A seven-dimensional flow analysis to help autonomous network management. *Inf. Sci.*, v. 278, p. 900–913, 2014.

[11] FERNANDES JR., G. et al. Statistical, forecasting and metaheuristic techniques for network anomaly detection. In: *Proceedings of the 30th Annual ACM Symposium on Applied Computing*. New York, NY, USA: ACM, 2015. (SAC '15), p. 701–707. ISBN 978-1-4503-3196-8.

[12] LEE, K.; YUN, S. Hybrid memory-efficient multimatch packet classification for {NIDS}. *Microprocessors and Microsystems*, v. 39, n. 2, p. 113 – 121, 2015. ISSN 0141-9331. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0141933115000113>.

[13] PATCHA, A.; PARK, J.-M. An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, v. 51, n. 12, p. 3448–3470, 2007.

[14] CISCO, S. *NetFlow Services Solutions Guide*. 2008. <http://www.cisco.com/en/US/docs/ios/solutions_docs/netflow/nfwhite.pdf>.

[15] SPEROTTO, A. et al. An overview of ip flow-based intrusion detection. *Communications Surveys Tutorials, IEEE*, v. 12, n. 3, p. 343–356, Third 2010. ISSN 1553-877X.

[16] RAHMANI, H.; SAHLI, N.; KAMOUN, F. Ddos flooding attack detection scheme based on f-divergence. *Computer Communications*, v. 35, n. 11, p. 1380–1391, 2012.

[17] BHUYAN, M. H.; BHATTACHARYYA, D. K.; KALITA, J. K. Network anomaly detection: Methods, systems and tools. *IEEE Communications Surveys and Tutorials*, v. 16, n. 1, p. 303–336, 2014.

[18] DELJAC Željko; RANDIć, M.; KRčELIć, G. Early detection of network element outages based on customer trouble calls. *Decision Support Systems*, v. 73, p. 57 – 73, 2015. ISSN 0167-9236. Disponível em: <http://www.sciencedirect.com/science/article/pii/S016792361500041X>.

[19] AHMED, M.; MAHMOOD, A. N. Network traffic analysis based on collective anomaly detection. In: *Industrial Electronics and Applications (ICIEA), 2014 IEEE 9th Conference on*. [S.l.: s.n.], 2014. p. 1141–1146.

[20] GRUPO de pesquisa em redes de computadores e comunicação de Dados da UEL. 2014. <http://redes.dc.uel.br/>.

[21] PROENÇA JR., M. L.; ZARPELAO, B. B.; MENDES, L. S. Anomaly detection for network servers using digital signature of network segment. In: *Proceedings - Advanced Industrial Conference on Telecommunications/Service Assurance with Partial and Intermittent Resources Conference/E-Learning on Telecommunications Workshop AICT/SAPIR/ELETE 2005*. [S.l.: s.n.], 2005. v. 2005, p. 290–295, doi:10.1109/AICT.2005.26. ISBN 0769523889.

[22] ADANIYA, M. H. A. C.; ABRAO, T.; PROENÇA JR, M. L. . Anomaly detection using metaheuristic firefly harmonic clustering. *Journal of Networks*, v. 8, n. 1, 2013.

[23] ZACARON, A. M. et al. Digital signature of network segment using flow analysis. In: *DCNET/ICE-B/OPTICS'12*. [S.l.: s.n.], 2012. p. 35–40.

[24] BRUTLAG, J. D. Aberrant behavior detection in time series for network monitoring. In: *LISA*. [S.l.: s.n.], 2000. v. 14, p. 139–146.

[25] DUBES, R.; JAIN, A. K. Clustering techniques: The user's dilemma. *Pattern Recognition*, v. 8, n. 4, p. 247 – 260, 1976. ISSN 0031-3203.

[26] MAULIK, U.; BANDYOPADHYAY, S. Genetic algorithm-based clustering technique. *Pattern Recognition*, v. 33, n. 9, p. 1455–1465, 2000.

[27] HOLLAND, J. H. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence.* [S.l.]: U Michigan Press, 1975.

[28] HOLLAND, J. H. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence.* Cambridge, MA, USA: MIT Press, 1992. ISBN 0262082136.

[29] SHEIKH, R. H.; RAGHUWANSHI, M. M.; JAISWAL, A. N. Genetic algorithm based clustering: A survey. In: *First International Conference on Emerging Trends in Engineering and Technology, ICETET '08, Nagpur, Maharashtra, India, July 16-18, 2008.* [S.l.: s.n.], 2008. p. 314–319.

[30] LIN, H.-J.; YANG, F.-W.; KAO, Y.-T. An efficient ga-based clustering technique. *Tamkang Journal of Science and Engineering*, TAMKANG UNIVERSITY, v. 8, n. 2, p. 113, 2005.

[31] CORREIA, M. Algoritmos genéticos. *Dos Algarves: A Multidisciplinary e-Journal*, ESGHT / UALG - Universidade do Algarve, v. 12, p. 36–43, 2003. ISSN 0873-7347.

[32] GOLDBERG, D. E.; HOLLAND, J. H. Genetic algorithms and machine learning. *Machine learning*, Springer, v. 3, n. 2, p. 95–99, 1988.

[33] MITCHELL, M. *An introduction to genetic algorithms.* [S.l.]: MIT Press, 1998. I-VIII, 1-208 p. ISBN 978-0-262-63185-3.

[34] SRINIVAS, M.; PATNAIK, L. M. Genetic algorithms: A survey. *IEEE Computer*, v. 27, n. 6, p. 17–26, 1994.

[35] JONG, K. A. D. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems.* Tese (Doutorado), Ann Arbor, MI, USA, 1975.

[36] SPEARS, W. M.; ANAND, V. A study of crossover operators in genetic programming. In: *ISMIS*. [S.l.: s.n.], 1991. p. 409–418.

[37] SRINIVAS, M.; PATNAIK, L. M. Adaptive probabilities of crossover and mutation in genetic algorithms. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 24, n. 4, p. 656–667, 1994.

[38] BANZHAF, W. et al. *Genetic Programming: An Introduction: on the Automatic Evolution of Computer Programs and Its Applications.* San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1998. ISBN 1-55860-510-X.

[39] HAAG, P. *NFDUMP*. 2015. <http://nfdump.sourceforge.net>.

[40] ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, n. 0, p. 53 – 65, 1987. ISSN 0377-0427.

[41] COWGILL, M.; HARVEY, R.; WATSON, L. A genetic algorithm approach to cluster analysis. *Computers Mathematics with Applications*, v. 37, n. 7, p. 99 – 108, 1999. ISSN 0898-1221.

[42] COLANZI, T. E. et al. Empirical studies on application of genetic algorithms and ant colony optimization for data clustering. In: *SCCC*. [S.l.: s.n.], 2010. p. 1–10.

[43] TSENG, L. Y.; YANG, S. B. A genetic approach to the automatic clustering problem. *Pattern Recognition*, v. 34, n. 2, p. 415 – 424, 2001. ISSN 0031-3203.

[44] RIBEIRO, F. J. L.; TRELEAVEN, P. C.; ALIPPI, C. Genetic-algorithm programming environments. *IEEE Computer*, v. 27, n. 6, p. 28–43, 1994.

[45] WITTEN, I. H.; FRANK, E.; HALL, M. A. *Data Mining: Practical Machine Learning Tools and Techniques*. 3rd. ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011. ISBN 0123748569, 9780123748560.

[46] POLI, A. A.; CIRILLO, M. C. On the use of the normalized mean square error in evaluating dispersion model performance. *Atmospheric Environment. Part A. General Topics*, v. 27, n. 15, p. 2427 – 2434, 1993. ISSN 0960-1686.

[47] SAKOE, H.; CHIBA, S. Dynamic programming algorithm optimization for spoken word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, v. 26, n. 1, p. 43–49, Feb 1978. ISSN 0096-3518.

[48] COSTA, N. C. D.; KRAUSE, D.; BUENO, O. Paraconsistent logics and paraconsistency. *Philosophy of logic*, of, v. 5, p. 655–781, 2007.

[49] PENA, E. et al. Correlational paraconsistent machine for anomaly detection. In: *Global Communications Conference (GLOBECOM), 2014 IEEE*. [S.l.: s.n.], 2014. p. 551–556.

[50] PHAAL, P.; PANCHEN, S.; MCKEE, N. *InMon corporation s sFlow: A method for monitoring traffic in switched and routed networks*. [S.l.], 2001.

[51] CARVALHO, L. F. et al. Unsupervised learning clustering and self-organized agents applied to help network management. *Expert Systems with Applications*, v. 54, p. 29 – 47, 2016. ISSN 0957-4174. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0957417416000555>.

[52] ASSIS, M. V. O. de; PROENÇA JR., M. L. Scorpius: sflow network anomaly simulator. *Journal of Computer Science*, v. 11, Issue 4, p. 662–674, 2015. ISSN 1552-6607.

TRABALHOS PUBLICADOS PELO AUTOR

1. Fadir Salmen; Paulo R. Galego Hernandes Jr.; Luiz Fernando Carvalho; Mario Lemes Proença Jr. **Using Firefly and Genetic Metaheuristics for Anomaly Detection based on Network Flows**, The Eleventh Advanced International Conference on Telecommunications (AICT 2015), Junho/2015, p. 113-118, ISBN: 978-1-61208-411-4. Bruxelas, Bélgica. (Qualis CC 2012, B1)

2. Paulo R. Galego Hernandes Jr.; Luiz Fernando Carvalho; Mario Lemes Proença Jr.. **Digital Signature of Network Segment using Flow Analysis through Genetic Algorithm and ACO metaheuristics**. XXXIII International Conference of the Chilean Computer Science Society (SCCC 2014), Novembro/2014, ISBN 978-956-7576-68-5. Talca, Chile. (Qualis CC 2012, B3)

3. Paulo R. Galego Hernandes Jr.; Luiz Fernando Carvalho; Gilberto Fernandes Jr.; Mario Lemes Proença Jr.. **Digital Signature of Network Segment Using Genetic Algorithm and Ant Colony Optimization Metaheuristics**. The Eighth International Conference on Emerging Security Information Systems and Technologies (SECURWARE 2014), Novembro/2014, ISBN: 978-1-61208-376-6. Lisboa, Portugal. (Qualis CC 2012, B3)