Universidade Estadual de Londrina

EDUARDO HENRIQUE MONTEIRO PENA

UM SISTEMA PARA DETECÇÃO DE ANOMALIAS QUE UTILIZA ASSINATURA DIGITAL DE SEGMENTO DE REDE, ARIMA ADAPTATIVO E LÓGICA PARACONSISTENTE

Londrina
2014

EDUARDO HENRIQUE MONTEIRO PENA

UM SISTEMA PARA DETECÇÃO DE ANOMALIAS QUE UTILIZA ASSINATURA DIGITAL DE SEGMENTO DE REDE, ARIMA ADAPTATIVO E LÓGICA PARACONSISTENTE

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

Orientador: Prof. Dr. Mario Lemes Proença Jr.

Londrina
2014

Catalogação elaborada pela Divisão de Processos Técnicos da Biblioteca Central da Universidade Estadual de Londrina

Dados Internacionais de Catalogação-na-Publicação (CIP)

P397s Pena, Eduardo Henrique Monteiro.
Um sistema para detecção de anomalias que utiliza assinatura digital de segmento de rede, ARIMA adaptativo e lógica paraconsistente / Eduardo Henrique Monteiro Pena. – Londrina, 2014.
128 f. : il.

Orientador: Mário Lemes Proença Junior.
Dissertação (Mestrado em Ciência da Computação) – Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Ciência da Computação, 2014.
Inclui bibliografia.

1. Redes de computadores – Medidas de segurança – Teses. 2. Linguagem de programação lógica – Teses. 3. Sistemas de transmissão de dados – Tráfego – Teses. 4. Assinaturas digitais – Teses. 5. Redes de computadores – Anomalias – Teses. I. Proença Junior, Mário Lemes. II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Ciência da Computação. III. Título.

CDU 519.68.022

EDUARDO HENRIQUE MONTEIRO PENA

UM SISTEMA PARA DETECÇÃO DE ANOMALIAS QUE UTILIZA ASSINATURA DIGITAL DE SEGMENTO DE REDE, ARIMA ADAPTATIVO E LÓGICA PARACONSISTENTE

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

BANCA EXAMINADORA

_____

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

_____

Prof. Dr. Sylvio Barbon Jr.
Universidade Estadual de Londrina – UEL

_____

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

_____

Prof. Dr. José Palazzo Moreira de Oliveira
Universidade Federal do Rio Grande do Sul – UFRGS

Londrina, 13 de agosto de 2014.

AGRADECIMENTOS

A Deus e Seu Filho, pelo imenso amor e cuidado com a minha vida. As palavras são insuficientes para expressar minha gratidão.

Aos meus pais, José Pena e Isabel, pelo apoio incondicional, compreensão e por todo carinho a mim dedicado. A vocês dois e a minhã irmã, Débora, por sempre torcerem por mim.

Ao meu professor e orientador Dr. Mario Lemes Proença Jr., por todo apoio e compreensão, pela confiança em meu trabalho, e por todas as conversas e conselhos que tanto contribuíram para meu crescimento profissional e pessoal.

Aos professores Dr. Sylvio Barbon Jr., Dr. Bruno B. Zarpelão e Dr. José Palazzo Moreira de Oliveira, pela dedicação de tempo e por todo o esforço direcionado à avaliação deste trabalho.

Aos amigos do grupo de redes: Luiz Fernando, Marcos Vinicius e Gilberto, pela amizade, trabalho, dedicação e motivação em todos os momentos. Ao amigo veterano, Alexandro Zacaron, pelo conhecimento e exemplo de superação deixado.

À minha avó, Florinda, pelo grande amor, por seu coração de ouro e por todas suas orações.

À minha amiga e namorada, Carolina, pela paciência, ajuda e cumplicidade ao longo desse processo. Obrigado por me apoiar, principalmente nos momentos mais difíceis.

Ao meu grande amigo e irmão Rodrigo, parceiro nos bons e maus momentos. Agradeço pela ajuda prestada quando mais precisei, pelo apoio e incentivo constante.

À CAPES pelo apoio financeiro concedido.

PENA, Eduardo Henrique Monteiro. **Um sistema para detecção de anomalias que utiliza assinatura digital de segmento de rede, arima adaptativo e lógica paraconsistente**. 2014. 128 f. Dissertação de Mestrado em Ciência da Computação – Universidade Estadual de Londrina, Londrina, 2014.

RESUMO

A identificação de comportamentos anômalos no tráfego de rede é uma tarefa fundamental para a segurança das redes de computadores e previne que situações adversas como falhas em equipamentos, uso demasiado de recursos e ataques comprometam o funcionamento adequado das redes de computadores. Este trabalho apresenta um sistema para detecção de anomalias baseado na caracterização do comportamento normal do tráfego e na aplicação da Lógica Paraconsitente. Atributos de tráfego de rede são utilizados com o modelo Auto-Regressivo Integrado de Médias Móveis para construir as Assinaturas Digitais de Segmento de Rede utilizando análise de Fluxos. Tais assinaturas descrevem perfis de comportamento normal para os atributos analisados e aferem as discrepâncias relativas as leituras reais de tráfego. A avaliação dessas discrepâncias é realizada por meio da implementação da Máquina Paraconsistente Correlacional, a qual utiliza os perfis gerados e leituras reais como fontes de informação evidencial na fundamentação lógica dos níveis de certeza e contradição referentes ao comprometimento da rede. O sistema foi avaliado nas redes da Universidade Estadual de Londrina e da Universidade Tecnológica Federal do Paraná. Os testes e resultados apresentados demonstram a efetividade do sistema na caracterização de tráfego e detecção de anomalias e sugerem sua viabilidade na implantação em ambientes reais.

Palavras-chave: ARIMA. Caracterização de tráfego. Detecção de anomalias. DSNSF. Lógica paraconsistente.

PENA, Eduardo Henrique Monteiro. **A system for anomaly detection that uses digital signature of network segment, adaptive arima and paraconsistent logic**. 2014. 128 p. Masters Dissertation in Computer Science – State University of Londrina, Londrina, 2014.

ABSTRACT

The identification of abnormal behaviors at the network traffic is a fundamental task for network security and prevents adverse situations such as equipment failures, overuse of resources and attacks to jeopardize the proper functioning of computer networks. This work presents a system for anomaly detection based on the characterization of normal traffic behavior and in application of Paraconsistent Logic. Network traffic features are used with the AutoRegressive Integrated Moving Average model to build the Digital Signatures of Network Segment using Flow analysis. Such signatures describe normal behavior profiles for the analyzed features and measure the discrepancies related to real traffic measurements. The evaluation of these discrepancies is accomplished by implementing the Correlational Paraconsistent Machine, which uses the generated profiles and real measurements as sources of evidential information on the logical reasoning of certainties and contradictions levels regarding the network compromise. The system was evaluated at State University of Londrina and Federal University of Technology. The presented tests and results demonstrate the system effectiveness in traffic characterization and anomaly detection and suggests the feasibility of deployment in real environments.

Key words: ARIMA. Traffic characterization. Anomaly detection. DSNSF. Paraconsistent logic.

LISTA DE ILUSTRAÇÕES

Figura 2.1 – Visão geral do funcionamento do Netflow. Adaptado de [22]. 23
Figura 2.2 – Verificação dos atributos de fluxos usando o protocolo NetFlow. Retirada de [22]. 24
Figura 2.3 – Esquema básico do sFlow, adaptado de [26]. 25
Figura 3.1 – Visão geral do sistema para detecção de anomalias proposto. 35
Figura 3.2 – Escolha do número de semanas para ruído branco baseado na média. 49
Figura 3.3 – Escolha do número de semanas para ruído branco baseado na mediana. 49
Figura 3.4 – Diagrama da utilização de modelos ARIMA. 50
Figura 4.1 – O espaço paraconsistente representado pelo reticulado de Hasse, adaptado de [75]. As marcações (F, QF, IC, QV, V, ⊥, T) são relativas ao estado lógico paraconsistente (Tabela 4.1), em função do nível de certeza (eixo horizontal) e contradição (eixo vertical). 54
Figura 4.2 – Interpolação do ponto (0,78,0,04) no reticulado. 63
Figura 5.1 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 22/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 73
Figura 5.2 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 23/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 74
Figura 5.3 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 24/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 74
Figura 5.4 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 25/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 75
Figura 5.5 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 26/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 75
Figura 5.6 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 29/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 76

Figura 5.7 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 30/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 76
Figura 5.8 – Média dos coeficientes de correlação, calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 29 de abril a 17 de maio de 2013. 77
Figura 5.9 – Coeficientes de correlação em relação a cada atributo analisado para o período de 29 de abril a 17 de maio de 2013. 78
Figura 5.10 – Média dos erros NMSE calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 29 de abril a 17 de maio de 2013. 78
Figura 5.11 – NMSE em relação a cada atributo analisado para o período de 29 de abril a 17 de maio de 2013. 79
Figura 5.12 – Nível de evidência real da MPC configurada para anomalia do tipo DoS, segunda-feira - 29/10/2012. 80
Figura 5.13 – Nível de evidência real da MPC configurada para anomalia do tipo DDoS, segunda-feira - 29/10/2012. 81
Figura 5.14 – Nível de evidência real da MPC configurada para anomalia do tipo Flash Crowd, segunda-feira - 29/10/2012. 81
Figura 5.15 – Reticulado de Hasse de uma MPC configurada para anomalia do tipo DoS, segunda-feira - 29/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (**nc, nct**). 82
Figura 5.16 – Reticulado de Hasse de uma MPC configurada para anomalia do tipo DDoS, segunda-feira - 29/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (**nc, nct**). 83
Figura 5.17 – Reticulado de Hasse de uma MPC configurada para anomalia do tipo Flash Crowd, segunda-feira - 29/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (**nc, nct**). 83
Figura 5.18 – Nível de evidência real da MPC configurada para anomalia do tipo DoS, terça-feira - 30/10/2012. 84
Figura 5.19 – Nível de evidência real da MPC configurada para anomalia do tipo DDoS, terça-feira - 30/10/2012. 84

Figura 5.20 – Nível de evidência real da MPC configurada para anomalia do tipo Flash Crowd, terça-feira - 30/10/2012. 84
Figura 5.21 – Reticulado de Hasse de uma MPC configurada para anomalia do tipo DoS, terça-feira - 30/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (**nc, nct**). 85
Figura 5.22 – Reticulado de Hasse de uma MPC configurada para anomalia do tipo DDoS, terça-feira - 30/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (**nc, nct**). 86
Figura 5.23 – Reticulado de Hasse de uma MPC configurada para anomalia do tipo Flash Crowd, terça-feira - 30/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (**nc, nct**). Quando a interpolação (**nc, nct**) atinge o estado Verdadeiro um alarme é disparado. 86
Figura 5.24 – Acurácia do sistema proposto utilizando a MPC e comparação com abordagem baseada em limites de confiança. Resultados para o período entre 22 de outubro e 9 de novembro. 88
Figura 5.25 – Curvas ROCs do sistema proposto e abordagem baseada em limites de confiança. Resultado para o período entre 22 de outubro e 9 de novembro. 88
Figura 5.26 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 22/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS. 90
Figura 5.27 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 23/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS. 90
Figura 5.28 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 24/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS. 91
Figura 5.29 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 25/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS. 91
Figura 5.30 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 26/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS. 92

Figura 5.31 – Média dos coeficientes de correlação, calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 22 de outubro a 9 de novembro de 2012. As medidas ilustradas referem-se a média dos seis atributos analisados, para M1, M2 e ACODS. 92
Figura 5.32 – Coeficientes de correlação em relação a cada atributo analisado para o período de 22 de outubro a 9 de novembro de 2012, para M1, M2 e ACODS. 93
Figura 5.33 – Média dos erros NMSE calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 22 de outubro a 9 de novembro de 2012. As médias ilustradas referem-se aos seis atributos analisados, para M1, M2 e ACODS. 93
Figura 5.34 – Erros NMSE em relação a cada atributo analisado para o período de 22 de outubro a 9 de novembro de 2012, para M1, M2 e ACODS. 93
Figura 5.35 – Acurácia do sistema proposto utilizando a MPC e comparação com ACODS. Resultados para o período entre 22 de outubro e 9 de novembro. 94
Figura 5.36 – Curva ROC do sistema proposto e comparação com abordagem ACODS. Resultado para o período entre 22 de outubro e 9 de novembro. 95
Figura 5.37 – Nível de evidência real da MPC adaptada com o ACODS configurada para anomalia do tipo Flash Crowd, terça-feira - 30/10/2012. 96
Figura 5.38 – Reticulado de Hasse de uma MPC adaptada com o ACODS configurada para anomalia do tipo Flash Crowd, terça-feira - 30/10/2012. 97
Figura 5.39 – Acurácia do sistema proposto utilizando duas MPCs. 98
Figura 5.40 – Curva ROC do sistema proposto comparado com adaptação da MPC com ACODS. Resultado para o período entre 22 de outubro e 9 de novembro. 98
Figura 5.41 – Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 13/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 100
Figura 5.42 – Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 14/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 100

Figura 5.43 – Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 15/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 101
Figura 5.44 – Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 16/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 101
Figura 5.45 – Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 17/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana. 102
Figura 5.46 – Média dos coeficientes de correlação, calculados entre os DSNSFs e as leituras reais do tráfego da rede da UTFPR no período de 29 de abril a 17 de maio de 2013. 103
Figura 5.47 – Coeficientes de correlação em relação a cada atributo analisado para o período de 29 de abril a 17 de maio de 2013, UTFPR. 103
Figura 5.48 – Média do erro NMSE calculados entre os DSNSFs e as leituras reais do tráfego da rede da UTFPR no período de 29 de abril a 17 de maio de 2013. 104
Figura 5.49 – NMSE em relação a cada atributo analisado. 104
Figura 5.50 – Acurácia do sistema proposto utilizando a MPC e comparação com abordagem baseada em limites de confiança. 105
Figura 5.51 – Curva ROC do sistema proposto e comparação com abordagem baseada em limites de confiança. Resultado para o período entre 29 de abril e 17 de maio. 105

LISTA DE TABELAS

Tabela 3.1 –Conjunto de atributos da DSNSF. 36
Tabela 3.2 –Comportamento teórico dos modelos **ARp, MA(q)** e **ARMA (p,q)** 45
Tabela 3.3 –Comportamento dos modelos mais recorrentes. 46
Tabela 4.1 –Estados lógicos paraconsistentes. Adaptados de [75] e [58]. 55
Tabela 4.2 –Relação entre anomalias e atributos de tráfego afetados. Adaptado de [39] e [76]. 57
Tabela 5.1 –Cenários de teste. 71
Tabela 5.2 –Informações sobre a simulação de anomalias (UEL). 72
Tabela 5.3 –Informações sobre a simulação de anomalias (UTFPR). 99

LISTA DE ABREVIATURAS E SIGLAS

ACODS	Ant Colony Optimization for Digital Signature
ARIMA	AutoRegresive Integrated Moving Average
BGP	Border Gateway Protocol
BSS	Blind Source Separation
COI	Communities of interest
CPU	Central Processing Unit
DDoS	Distributed Denial of Service
DoS	Denial of Service
DPM	Discriminative Paraconsistent Machine
DSNSF	Digital Signature of Network Segment using Flow Analysis
FAC	Função de AutoCorrelação
FACP	Função de AutoCorrelação Parcial
IETF	Internet Engeneering Task Force
IMAPIT	Integrated Measurement Analysis Platform for Internet
IP	Internet Protocol
IPFIX	IP Flow Information Export
IPv6	Internet Protocol version 6
ISO	International Organization for Standardization
ISP	Internet Service Providers
LP	Lógica Paraconsistente
LPA	Lógica Paraconsistente Anotada
LPA2v	Lógica Paraconsistente Anotada de anotação com dois valores
MIB	Management information base
MPC	Máquina Paraconsistente Correlacional
MPLS	Multiprotocol Label Switching
NMSE	Normalized Mean Square Error
PCA	Principal Component Analysis
PSAMP	Packet Sampling
RFC	Request for Comments
ROC	Receiver Operating Characteristics
SCTP	Stream Control Transport Protocol
sMAPE	Symmetric Mean Absolute Percent Error

SNMP	Simple Network Management Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UEL	Universidade Estadual de Londrina
UTFPR	Universidade Tecnológica Federal do Paraná

LISTA DE SÍMBOLOS

d_t	Distâncias entre DSNSF e leituras de tráfego real
$\varphi_{t,f}$	Distâncias entre DSNSF e leituras de tráfego real limitadas a 1
$H(A)$	Entropia para o intervalo de tempo analisado.
ϕ_{kk}	Função de autocorrelação parcial
L	Função de máxima verossimilhança
f	Índice para atributos
t	Intervalo de tempo t
T	Intervalos analisados durante o dia
$r_{t,f}$	Leitura real do tráfego para um atributo f , no intervalo t
P	Matriz de correlação
R	Matriz de leituras reais de tráfego
X	Matriz de previsão dos atributos do DSNSF
η_c	Nível de certeza
η_{cr}	Nível de certeza real
η_{ct}	Nível de contradição
λ	Nível de evidência desfavorável
μ	Nível de evidência favorável
ω_{er}	Nível de evidência real
F	Número de atributos analisados
n_i	Número de ocorrência de elementos i no intervalo de tempo analisado.
B	Operador de defasagem
Δ	Operador de diferença
p	Ordem do modelo auto-regressivo
q	Ordem do modelo de médias móveis
ϕ	Parâmetro do modelo auto-regressivo
θ	Parâmetro do modelo de médias móveis
k	Passos para defasagem
e_t	Relação anomalias e discrepâncias
κ	Relação entre essas anomalias e atributos afetados

c_k	Resultado da covariância
ρ_k	Resultado da função de auto-correlação
a_t	Ruído branco
Y	Série temporal
Z	Série temporal diferenciada
s	Soma das ocorrências presentes no histograma
ψ_t	Termo evidêncial no intervalo t
N	Total de ocorrência de elementos no intervalo de tempo analisado
σ^2	Variância do conjunto
VN	Verdadeiros Negativos
VP	Verdadeiros Positivos
ξ	Vetor de parâmetros do modelo ARIMA

SUMÁRIO

1 INTRODUÇÃO 17

2 TRABALHOS RELACIONADOS 21
2.1 GERÊNCIA DE REDES 21
2.2 FLUXOS IP 22
2.2.1 Netflow 23
2.2.2 Internet Protocol Flow Information Export (IPFIX) 24
2.2.3 sFlow 25
2.2.4 Impacto da Amostragem na Detecção de Anomalias 26
2.2.5 Análise de Fluxos 26
2.3 DETECÇÃO DE ANOMALIAS 27
2.4 LÓGICA PARACONSISTENTE 31

3 ASSINATURA DIGITAL DE SEGMENTO DE REDE UTILIZANDO ANÁLISE DE FLUXOS 33
3.1 DEFINIÇÃO, APLICAÇÃO E FUNCIONAMENTO 33
3.2 PREPARAÇÃO DOS DADOS 35
3.3 ANÁLISE DE SÉRIES TEMPORAIS 37
3.3.1 Séries Temporais na Caracterização do Tráfego 38
3.4 MODELO AUTO-REGRESSIVO INTEGRADO DE MÉDIAS MÓVEIS 39
3.4.1 O modelo Auto-Regressivo 40
3.4.2 O Modelo de Médias Móveis 40
3.4.3 O Modelo Auto-Regressivo de Médias Móveis 40
3.4.4 O Modelo Auto-Regressivo Integrado de Médias Móveis 41
3.4.5 Propriedades Estatísticas para os Modelos ARIMA 41
3.4.6 Identificação dos Modelos 44
3.4.7 Estimação dos Parâmetros ϕ e θ 46
3.4.8 Padrões para o Ruído Branco 47
3.4.9 Visão Geral da Metodologia ARIMA 49

4 DETECÇÃO DE ANOMALIAS COM A LÓGICA PARACONSISTENTE 51

4.1 Tratamento de Incertezas 51
4.2 Lógica Paraconsistente Anotada de Anotação com Dois Valores (LPA2v) 52
4.2.1 O Nível de Evidência Real 55
4.3 Máquina Paraconsistente Correlacional (MPC) 56
4.3.1 Representações 56
4.3.2 Extração de Informações Evidenciais 57
4.3.3 Funcionamento da MPC 65
4.3.4 Complexidade Computacional 66

5 TESTES E RESULTADOS 67
5.1 AMBIENTES DE REDE E MONITORAMENTO 67
5.2 MÉTRICAS DE AVALIAÇÃO 67
5.2.1 Coeficiente de Correlação 68
5.2.2 Erro Quadrático Médio Normalizado (NMSE) 68
5.2.3 Acurácia 68
5.2.4 Curvas ROC (Receiver Operating Characteristic) 69
5.3 CENÁRIOS DE TESTE 70
5.3.1 Cenário 1 – Caracterização de Tráfego e Detecção de Anomalias (UEL) 72
5.3.2 Cenário 2 – Detecção de Anomalias com a MPC e com o ACODS (UEL) 89
5.3.3 Cenário 3 –MPC Baseada no ARIMA e ACODS (UEL) 95
5.3.4 Cenário 4 – Caracterização e Detecção de Anomalias (UTFPR) 99

6 CONCLUSÃO 106

REFERÊNCIAS 109

APÊNDICES 117
APÊNDICE A – Exemplos de modelos utilizados 118
APÊNDICE B – Ant Colony Optimization for Digital Signature 124
APÊNDICE C – Limites de confiança 127

1 INTRODUÇÃO

Os processos de inovação que ocorrem em ambientes tecnológicos fazem com que as redes de computadores sejam parte integrante da sociedade atual. O extenso número de usuários e dispositivos eletrônicos conectados à Internet, as velocidades de banda larga, a maior quantidade e qualidade de vídeos e a disseminação do *Wi-Fi* são alguns fatores que dependem diretamente de um ambiente de comunicação confiável. Dessa forma, as redes devem ser devidamente gerenciadas, garantindo o cumprimento de diversos requisitos em termos de desempenho, disponibilidade e segurança [1].

Algumas circunstâncias podem dificultar o controle adequado das redes, como falhas em equipamentos ou aplicações, congestionamentos e atividades maliciosas. De modo geral, esses inconvenientes são relacionados às anomalias de tráfego de rede. Uma anomalia pode ser definida como um desvio súbito e acentuado no comportamento normal dos níveis de tráfego [2]. Caso não seja tratada adequadamente, ela pode comprometer a segurança e o desempenho da rede, causando assim, prejuízo ao usuário final. Dessa forma, a gerência de redes deve estar preparada para detectar qualquer tipo de anomalia, implementando medidas proativas que garantam a operabilidade, disponibilidade e segurança das informações.

Diversas técnicas para a detecção de anomalias têm sido desenvolvidas e, embora sejam consideradas importantes componentes em sistemas de segurança de redes, ainda representam um objeto de estudo desafiador [3]. De maneira geral, as técnicas para detecção de anomalias são divididas em duas categorias: detecção baseada nas assinaturas das anomalias e detecção baseada na caracterização do comportamento normal da rede.

A detecção de anomalias baseada em assinatura monitora os pacotes que trafegam pela rede e compara-os com as especificações armazenadas em uma base de informação. Notificações ou alarmes são gerados quando conteúdos semelhantes são identificados. Embora apresente poucos alarmes falsos, a incapacidade dessas técnicas em detectar anomalias que estejam fora de uma base de informação representa uma indesejada fragilidade [4].

A detecção baseada na caracterização do comportamento normal da rede, mais especificamente, do comportamento normal do tráfego da rede, detecta anomalias com base em desvios entre medições em tempo real e um perfil de comportamento normal estabelecido através de uma análise histórica de dados. Tal abordagem apresenta vantagens como a capacidade de detectar anomalias desconhecidas e a facilidade de implantação e adaptação em diferentes cenários.

Apesar de seus benefícios, o uso da caracterização do comportamento normal do tráfego da rede na detecção de anomalias ainda encontra questões em aberto, como a definição do comportamento normal, a disponibilidade de dados e a eficiência na detecção [5], [6] e [7]:

- *Definição do comportamento normal.* O tráfego de rede pode apresentar características não estacionárias ou irregulares, sofrendo alterações ao longo do tempo [8]. Assim, estabelecer uma região que abranja todos os possíveis padrões normais se torna um processo desafiador.
- *Disponibilidade de dados.* Os dados utilizados para treinamento e validação das técnicas de detecção frequentemente contêm ruídos que podem ser anomalias. Distingui-los, e removê-los, sem perder as características inerentes ao tráfego normal, não é uma tarefa trivial.
- *Eficiência na detecção*. As técnicas presentes na literatura ainda apresentam problemas relacionados a eficiência na detecção de anomalias. Em particular, devido às taxas de falsos positivos, ou seja, situações normais que são, erroneamente, classificadas como anomalias.

Esse trabalho apresenta um sistema para detecção de anomalias baseado na caracterização do comportamento normal de tráfego, com o objetivo de objetivo de auxiliar na gerência de redes.

O sistema é composto por duas partes. Primeiramente, uma ferramenta denominada Assinatura Digital de Segmento de Rede utilizando análise de Fluxos, ou *Digital Signature of Network Segment using Flow Analysis* (DSNSF), é utilizada na geração de perfis de comportamento normal para atributos extraídos do tráfego de rede. Para o cálculo das especificações geradas pela DSNSF foi utilizado o modelo Auto-Regressivo Integrado de Médias Móveis [9], ou *AutoRegressive Integrated Moving Average* (ARIMA).

A DSNSF é utilizada na avaliação do comportamento real do tráfego em busca de possíveis anomalias, e como ela caracteriza um padrão de comportamento esperado para o tráfego, desvios em relação a mesma representam indícios sobre eventos anômalos. A segunda parte do sistema é responsável pela avaliação desses desvios.

Para tanto esse trabalho propõe uma solução baseada na Lógica Paraconsistente (LP) [10], mais especificamente em sua extensão, a Lógica Paraconsistente

Anotada de anotação com dois valores (LPA2v), que considera evidências favoráveis e contrárias a uma determinada proposição.

Sistemas de informação inteligentes englobam certo grau de imprecisão [11], em particular nas tarefas de gerenciamento de rede, onde as inconsistências podem surgir em diferentes níveis de análise, por exemplo, na aquisição de dados, caracterização e análise de fluxos. Logo, a LPA2v é utilizada nesse trabalho como uma alternativa à lógica binária (também conhecida como clássica) [12], identificando não só comportamentos extremos ("Anomalia", "Normal"), mas também indefinições ou contradições que possam surgir na busca da solução do problema proposto.

A metodologia empregada nesse trabalho utiliza a Lógica Paraconsistente 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 as DSNSFs e leituras reais do tráfego da rede e utiliza uma sequência de procedimentos para extrair informações evidenciais derivadas do contraste entre cada DSNSF e as leituras reais. Essas informações são utilizadas em uma formalização das indefinições e contradições relativas ao comprometimento da rede. Anomalias são identificadas quando todas as informações evidenciais apontam para a mesma resposta (presença de uma anomalia no tráfego), com o mínimo de contradição.

O objetivo principal desse trabalho foi o estudo e desenvolvimento de um sistema para detecção de anomalias que contribua para o gerenciamento de segurança e desempenho de redes. De maneira geral, o trabalho apresenta as seguintes contribuições:

- *Análise de diferentes atributos e especificação da detecção*. O sistema proposto analisa seis atributos de tráfego, os quais especificam características de volume e distribuição do tráfego. A detecção é feita com base no distúrbio causado por três tipos de anomalias (*DoS, DDoS* e *Flash Crowd*) no comportamento normal de determinados atributos.
- *Uso do modelo ARIMA*: Adaptação do mesmo com o objetivo de explorar sua flexibilidade para geração das Assinaturas Digitais de Segmento de Rede utilizando análise de Fluxos.
- *Composição de modelos*. Assinaturas digitais e Lógica Paraconsistente são combinadas com o objetivo de reduzir a quantidade de alarmes falsos.

- *Métricas paraconsistentes*. As métricas extraídas da análise paraconsistente descrevem níveis de intensidade relativos ao nível de certeza sobre um evento anômalo, auxiliando o administrador ou gerente de rede na tomada de decisão.

A organização do presente trabalho está feita da seguinte forma: o Capítulo 2 apresenta alguns fundamentos concernentes à área de gerência de redes e sumariza algumas publicações relacionadas à caracterização de tráfego e detecção de anomalias; o Capítulo 3 motiva o uso das assinaturas digitais de segmento de rede e apresenta o modelo ARIMA; o Capítulo 4 discute alguns princípios sobre paraconsistência e descreve a arquitetura da Máquina Paraconsistente Correlacional, a qual foi utilizada na detecção de anomalias; o Capítulo 5 apresenta os testes e resultados conduzidos em ambientes de redes reais; o Capítulo 6 apresenta as conclusões e sugestões para trabalhos futuros.

2 TRABALHOS RELACIONADOS

Este capítulo apresenta alguns conceitos e artigos relacionados ao presente trabalho. Visando a compreensão do estado da arte das soluções para detecção de anomalias em redes de computadores, é realizada uma exposição abrangendo os elementos tipicamente encontrados na gerência de redes, na caracterização e medição de tráfego e nas técnicas para detecção de anomalias.

2.1 GERÊNCIA DE REDES

As redes de computadores se tornaram infraestruturas indispensáveis para a sociedade atual com seu apelo por troca de informações em altas velocidades. Por meio delas, diversas ferramentas e serviços fundamentais para o desenvolvimento empresarial, governamental e organizacional foram desenvolvidas. No entanto, todo esse desenvolvimento implica na necessidade de melhorias no processo de gerenciamento. De maneira geral, é papel central do gerente ou administrador da rede, detectar problemas que possam implicar na queda ou diminuição da qualidade de um determinado serviço [13].

A gerência de redes pode ser definida como um processo de controle abrangendo ações, metodologias, procedimentos e ferramentas para administração, operação e manutenção dos sistemas de rede de forma segura [14]. Segundo o padrão da Organização Internacional de Normalização, ISO, (*International Organization for Standardization*), a arquitetura para o gerenciamento de redes pode ser dividida em cinco áreas de atuação [15].

A primeira área lida com a gerência de falhas, detectando, alertando e corrigindo operações incomuns no funcionamento de uma rede. Uma falha é uma situação anormal tipicamente criada por operações incorretas ou um grande número de erros. Sua recuperação exige esforços de gerenciamento visando a identificação e a reparação do equipamento ou sistema onde a falha ocorreu.

A segunda área cuida do gerenciamento de configuração, e está relacionada à inicialização e configuração da rede. Ela também é responsável pelas tarefas de manutenção e atualização das ligações entre os equipamentos durante a operação da rede.

A terceira área está relacionada ao gerenciamento de contabilização, responsável pelo controle de acesso aos dispositivos da rede pelos seus usuários. Sua meta é evitar que um determinado usuário abuse de seus privilégios de acesso.

A quarta área refere-se ao gerenciamento do desempenho, que consiste no monitoramento e controle dos recursos e atividades presentes na rede. O monitoramento de determinados aspectos da rede como grau de utilização, perfil de tráfego, vazão e latência geram estatísticas que auxiliam no planejamento e administração da rede.

A quinta e última área de atuação é o gerenciamento da segurança, o qual deve proteger todos os recursos da rede. Uma política de segurança efetiva deve ser estabelecida, tratando de assuntos como criptografia, monitoração e controle do acesso à rede e das informações obtidas nela. Isso pode ser feito com o auxílio de coleta, armazenamento e exame de relatórios de segurança.

Ao levar em conta o pleno domínio das cinco áreas de gerenciamento, se pressupõe que o gerente de redes tenha informações suficientes dos eventos que acontecem na rede. No entanto, o gerenciamento manual dessas informações seria inviável. Devido à grande quantidade de dados, cada vez mais recursos humanos especializados seriam necessários. Dessa forma, soluções mais automatizadas entram em cena para ajudar as atividades de gerência [13].

O monitoramento da rede pode ser realizado de forma passiva, analisando apenas o conteúdo do tráfego da rede sem interferir em seu fluxo de dados. Essa abordagem consegue analisar grande quantidade de informações sem prejudicar o desempenho da rede. Um sistema de detecção de anomalias eficaz deve conter uma fonte de dados que forneça informações adequadas para o monitoramento das atividades da rede, e que auxiliem no reconhecimento de comportamentos anômalos no tráfego.

Uma solução que visa o monitoramento das redes de alta velocidade é a análise de fluxos IP. Ela utiliza agrupamento de pacotes com sessões de transmissões unidirecionais e que compartilham características como protocolo de transporte, endereços IP de origem e destino e portas [16]. Os fluxos podem ser relacionados à uma aplicação, equipamento de rede ou usuário. Assim, protocolos de gerenciamento de redes e estatísticas baseadas em análise de fluxos vêm sendo cada dia mais empregados na detecção de anomalias [17].

2.2 FLUXOS IP

Um fluxo IP pode ser definido como uma sequência de pacotes unidirecionais ou bidirecionais transmitida entre um dispositivo de origem e outro de destino que tenham atributos em comum em um determinado período de tempo [18]. Introduzida pela empresa *Cisco Systems* por meio de um protocolo proprietário chamado *NetFlow* [19], os fluxos IP

armazenam informações de alto nível sobre as conexões entre dispositivos, sendo amplamente utilizados no monitoramento, planejamento e segurança de redes [18]. Atualmente, diversos equipamentos suportam extração e exportação de fluxos IP, sendo os protocolos mais comuns o *NetFlow*, da Cisco [20]; o IPFIX, da IETF (*Internet Engineering Task Force*) [18]; e o sFlow, da InMon [21].

2.2.1 Netflow

O *NetFlow* é uma tecnologia de monitoramento de tráfego desenvolvida por Darren Kerr e Barry Bruins na empresa *Cisco Systems* em 1996. Ela define um padrão para exportação de informações e estatísticas sobre o tráfego de rede. Os dispositivos de rede analisam pacotes sendo transmitidos em suas interfaces e capturam estatísticas em nível de fluxo, com base em uma configuração de amostragem ou filtragem. Essas informações são armazenadas em *caches*, que são agregados e exportados através de um protocolo de transporte, tipicamente o UDP (User Datagram Protocol) ou o SCTP (Stream Control Transport Protocol). A Figura 2.1 ilustra o funcionamento do *Netflow* de maneira geral.

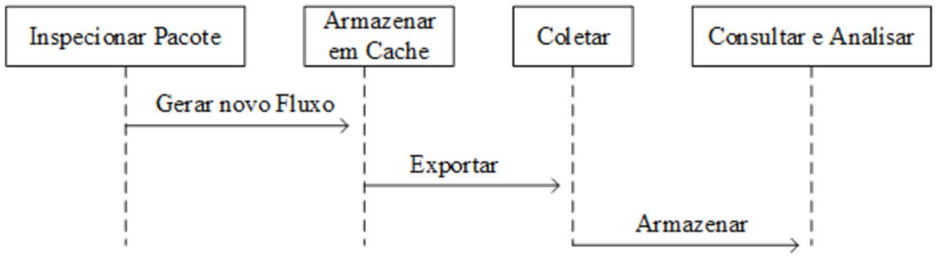


Figura 2.1 – Visão geral do funcionamento do *Netflow*. Adaptado de [20].

Os pacotes encaminhados a partir de um roteador ou *switch* têm os seguintes atributos do cabeçalho IP verificados: endereço IP de origem, endereço IP de destino, porta de origem, porta de destino, protocolo da camada de rede, classe de serviço e interface do roteador ou *switch*. Todos os pacotes contendo esses mesmos atributos são agrupados em um fluxo. A Figura 2.2 ilustra a inspeção de atributos de pacote.

Qualquer diferença entre os atributos de um novo pacote e um fluxo é suficiente para a criação de um novo fluxo. Um gerenciamento do *cache* de fluxos é feito, e as informações nele contidas são periodicamente exportadas a um coletor externo. O tamanho do

cache varia de 1024 a 524 mil entradas de registros de fluxos, atribuído via *software*, respeitando a capacidade física do sensor de gerenciamento.

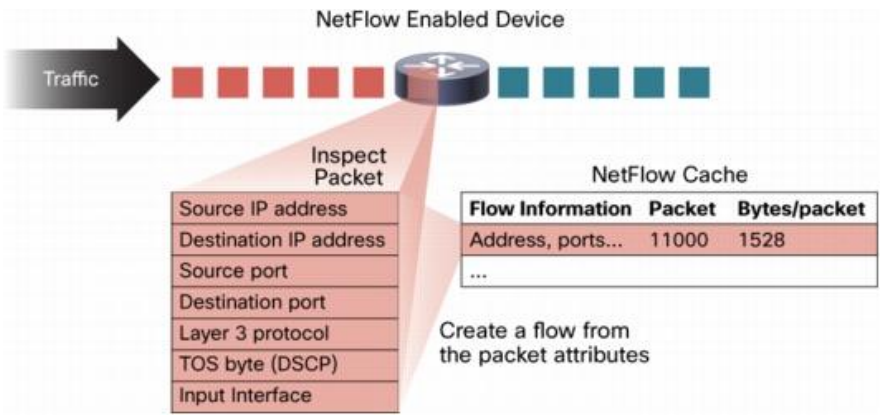


Figura 2.2 - Verificação dos atributos de fluxos usando o protocolo NetFlow. Retirada de [20].

Algumas regras são empregadas para tornar o gerenciamento escalável: fluxos inativos durante certo tempo são descartados; algumas políticas de exclusão e substituição de fluxos são empregadas após o preenchimento completo do *cache*; o fim de uma conexão TCP.

Em relação ao formato de exportação, o *NetFlow* fornece várias versões sendo a 5 e a 9 as mais populares. A partir da versão 9, extensibilidade e flexibilidade foram adicionadas para a integração com MPLS, IPv6 e BGP. Uma variação chamada *Sampled NetFlow* foi desenvolvida para reduzir a granularidade computacional do número de fluxos exportados. Ele é implementado com base em amostragem feita por meio da seleção do n-éssimo pacote ou randomicamente selecionando um pacote em uma janela de tempo determinada.

2.2.2 *Internet Protocol Flow Information Export* (IPFIX)

O *IP Flow Information Export protocol* (IPFIX) é um protocolo padronizado pela IETF baseado no *Netflow* versão 9 [22] que foi concebido como uma alternativa a protocolos proprietários para o monitoramento de redes por meio da exportação de fluxos. O protocolo de transmissão de informação está definido na RFC 5101, o modelo de informação na RFC 5102; já a RFC 5103 define a exportação bidirecional de fluxos.

O funcionamento do IPFIX é muito similar ao *Netflow*. Os pacotes que chegam aos dispositivos, roteadores ou *switches*, são analisados sequencialmente ou por meio de amostragem. A padronização do IPFIX se preocupou com o tratamento de sobrecarga dos dispositivos, utilizando análise amostral pelo protocolo PSAMP (*Packet Sampling*). O esquema de amostragem presente no IPFIX fornece diversos tipos de amostragem: baseada em contagem, temporização, regras de filtros, aleatória, e probabilística.

O IPFIX foi desenvolvido para atender requisitos cada vez mais exigentes no âmbito de análise de tráfego de rede, fornecendo um modelo extensível e flexível que pode ser customizado. A aceitação e padronização deste está em desenvolvimento [23]; a tendência é que cada vez mais fabricantes providenciem suporte ao IPFIX em seus dispositivos.

2.2.3 sFlow

O *sFlow,* acrônimo de *Sampling Flow,* foi desenvolvido pela *InMon* [21], e se tornou um padrão na indústria definido na RFC 3176. Esse protocolo utiliza amostragem para atingir escabilidade, tendo sua aplicação voltada para redes de altas velocidades. A amostragem de fluxo pode ser feita por estatística baseada em pacotes ou baseada no tempo. O *sFlow* é uma tecnologia que vem sendo integrada com roteadores e *switches* de alguns fabricantes de equipamentos de rede, como Alcatel, *Extreme*, HP e *Hitachi*. Essa integração é feita por meio de um agente denominado *sFlow*, o qual pode ser visto como um *software* que combina contadores de interface e amostras de fluxos em um datagrama *sFlow*, enviando-as a coletores *sFlow*. A Figura 2.3 apresenta o um esquema básico do *sFlow.*

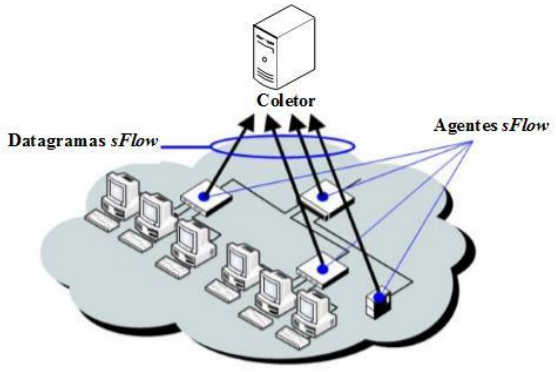


Figura 2.3 - Esquema básico do *sFlow, a*daptado de [24].

2.2.4 Impacto da amostragem na detecção de anomalias

Como descrito anteriormente, as tecnologias de fluxos IP oferecem diversos métodos para amostragem de pacotes. Tais métodos têm sido amplamente empregados em redes de grande porte, como forma de reduzir a quantidade de dados que os fluxos geram, diminuindo a complexidade de análise e gerência.

De maneira geral, as técnicas de amostragem foram introduzidas no monitoramento de redes como forma de aliviar o grande consumo de CPU, de memória nos roteadores e a alta taxa de transmissão de dados utilizadas pelos mesmos para coletar e transportar informações sobre fluxos. A implementação é feita de maneira simples, com baixa utilização de CPU e memória. Porém, alguns estudos tem mostrado sua ineficiência em descrever com exatidão algumas métricas estatísticas sobre a rede [25].

No que diz respeito à utilização de fluxos com amostragem na detecção de anomalias, uma série de fatores deve ser considerada, como técnicas de amostragem utilizadas, atributos sendo levados em consideração, métodos para detecção utilizados, e assim por diante.

O trabalho de Mai et al. [26] investiga a capacidade de dados amostrados na detecção de anomalias. As técnicas de amostragem mais comuns são analisadas, e os atributos de fluxos mais degradados são identificados. Semelhantemente, Brauckhoff et al. [27] avaliam os impactos da amostragem nas métricas de detecção. Um modelo cujo objetivo é tornar o processo de amostragem mais eficiente para a detecção de anomalias é apresentado em [28]. Um fator em comum apontado por esses trabalhos é de que atributos de volume e distribuição são aqueles que menos são degradados pelo processo de amostragem.

2.2.5 Análise de fluxos

A análise de fluxos IP é um processo que inclui coleta, agregação e armazenamento de fluxos, para posterior pesquisa e análise na busca de informações relevantes [16]. Essas informações são, basicamente, relacionadas à mensuração, gerenciamento e segurança de redes, usualmente direcionadas a administradores de rede e ISPs (*Internet Service Providers*). O monitoramento de fluxos em um ambiente de rede pode ter como objetivo o monitoramento da rede, das aplicações ou de usuários.

O monitoramento da rede fornece informações sobre todos os dispositivos da rede, como roteadores e *switches*, e é tipicamente utilizada para detecção de problemas. Muitos

aspectos da rede têm sido investigados com as informações sobre fluxos: medidas de latência (*network delays*) [29], caracterização de tráfego[30], qualidade de serviço [31], etc.

O monitoramento das aplicações é tipicamente utilizado para planejamento e alocação de recursos, uma vez que descreve informações sobre o uso da rede em nível de aplicação. Liu and Huebner [32] investigam características estocásticas de atributos derivados de fluxos para aplicações populares e que consomem maiores recursos. Um conceito chamado Comunidade de Interesse, *Communities of interest* (COI), é investigado sobre a perspectiva de fluxos em [33]. Este trabalho apresenta uma heurística para a classificação de tráfego. Essa classificação é útil na priorização de tráfego, e constantemente utilizada por ISPs.

O monitoramento de usuário fornece informações sobre a utilização de recursos e aplicações pelo usuário, sendo essas informações utilizadas para planejamento, controle de acesso e definição de políticas de segurança. Um trabalho que relaciona as dependências entre sistemas computacionais, componentes de *software* e usuários pode ser visto em [34]. Algumas heurísticas e métodos são aplicados para definição de comportamento de usuários comuns da rede em [35].

A análise de fluxos não está limitada apenas às aplicações mencionadas acima. De modo geral, todas essas perspectivas podem ser utilizadas em conjunto para uma determinada aplicação. Dessa forma, a análise de fluxos IP fornece uma fonte rica de recursos para a detecção de anomalias.

2.3 DETECÇÃO DE ANOMALIAS

Anomalias são alterações inesperadas e significativas, subitamente encontradas nos níveis do tráfego de rede [7]. Independentemente da natureza de uma anomalia, maliciosa ou não intencional, é de suma importância sua análise e diagnóstico uma vez que a mesma pode congestionar o trafego da rede, e até mesmo esgotar os recursos disponíveis para o usuário final [36].

Segundo Campbell [37], entender o comportamento de anomalias ou ataques é um procedimento que permite o desenvolvimento de defesas mais efetivas. O trabalho apresentado por ele descreve algumas características do ataque de negação de serviço (ou *Denial of Service - DoS*). Resumidamente, um *DoS* visa o congestionamento de um serviço ou servidor e é realizado por um único *host.* Outro ataque semelhante ao DoS é o *Distributed Denial of Service* ou *DDoS*, o qual ocorre quando vários *hosts* que, intencionalmente (exceções

em casos de *BotNets*, ou redes zumbis, compostas de *hosts* controlados remotamente por *hackers*), atacam *hosts* de destino [38].

De acordo com Estevez *et al.* [7], a detecção de anomalias está relacionada com a capacidade de reconhecer ações ou eventos que venham a comprometer a confidencialidade, integridade ou disponibilidade de uma rede. Tipicamente, os modelos para detecção de anomalias são divididos em duas categorias: baseado em assinatura ou baseado na caracterização do comportamento normal do tráfego.

A detecção de anomalias baseada em assinatura depende de uma especificação das principais características das anomalias, a qual é armazenada em alguma base de informação. Durante o monitoramento do tráfego de rede, um sistema de detecção baseado em assinatura busca por conteúdos que tenham as mesmas características das assinaturas de sua base, notificando seus administradores quando conteúdos semelhantes são identificados. Uma vez implantado, um sistema baseado nessa abordagem pode proteger a rede sem necessidade de treinamento histórico. Além disso, essa abordagem é bastante eficaz, apresentando taxas reduzidas de alarmes falsos. No entanto, com o aumento da capacidade das redes e as constantes novas ameaças a sua segurança e a utilização de assinaturas se tornou incapaz de adaptar suas especificações aos novos cenários e tipos de anomalias. A base contendo as assinaturas necessita de uma constante atualização manual, o que consume, demasiadamente, recursos da equipe de administração de redes [5].

A caracterização do comportamento normal do trafego é uma abordagem para detecção de anomalias baseada em um perfil, obtido a partir de um estudo do comportamento da rede, o qual caracteriza movimentos e atividades consideradas normais. Caso um evento da rede desvie significativamente do perfil de normalidade, pode ser considerado como anômalo [36]. Uma das principais vantagens de um sistema baseado em comportamento é a capacidade de detectar anomalias desconhecidas, sem a necessidade de constante atualização. Além disso, os perfis de atividade normal são tipicamente personalizados para cada rede, dificultando a exploração de suas vulnerabilidades. A necessidade de um período de treinamento para definição dos perfis de normalidade, juntamente com as variações naturais de comportamento das redes são alguns desafios na implantação de sistema utilizando essa abordagem.

As soluções para detecção de anomalia de sistemas baseados na caracterização do comportamento normal da rede geralmente utilizam abordagens estatística, métodos baseados em mineração de dados, ou aprendizagem de máquina [5].

A seguir há a descrição de alguns trabalhos que discutem técnicas para detecção.

Uma análise estatística do tráfego da rede coletado em registros de fluxos é feita em Lakhina *et al.* [36], em que as anomalias são identificadas e diagnosticadas, uma vez que o método proposto tem a capacidade de localizar a origem dessa anomalia. Os dados são analisados por uma técnica chamada análise de componentes principais, a qual separa o tráfego em dois subespaços – normal e anômalo. Um sinal mais ruidoso contendo picos de tráfego significativos caracteriza o subespaço anômalo. Os autores reportaram resultados promissores com a utilização da técnica, sendo pioneiros no diagnóstico de anomalias. No entanto, um estudo levantado por Ringberg *et al.* [39], o qual relaciona a sensibilidade desse modelo com novos conjuntos de dados e parâmetros, aponta para a dificuldade de implantação da técnica em ambientes reais, devido à dificuldade encontrada no ajuste de parâmetros, o qual pode ocasionar grande variação na taxa de falso positivo.

Alguns trabalhos tiveram considerável importância para o desenvolvimento da área de detecção de anomalias. Hajji [40] desenvolve um *baseline* que caracteriza a operação normal da rede. A técnica estatística empregada utiliza a distribuição assintótica da diferença entre sucessivas estimativas do comportamento da rede. Os picos não previstos na modelagem, são tidos como eventos anômalos. O principal problema da abordagem proposta é a utilização de dados livres de anomalias, o que faz com que a utilização em cenários reais fique comprometida. Thottan e Chuanyi [41] fazem uma análise de alguns trabalhos para detecção de anomalias. Um modelo também é proposto, o qual utiliza processamento estatístico de sinal que aplica de um processo Auto-Regressivo para modelagem dos dados. O processamento dos dados, coletados em variáveis residentes em MIBs SNMP, permite estabelecer uma correlação de objetos, a qual auxilia no processo de identificação de anomalias.

Técnicas de processamento de sinais em séries temporais, oriundas de dados presentes em MIBs e em registros de fluxos de dados, são empregadas na análise de anomalias de tráfego de rede por Plonka e Barford [2]. O sistema proposto, denominado IMAPIT (*Integrated Measurement Analysis Platform for Internet*), decompõe as séries temporais relativas às estatísticas do tráfego de rede por meio de transformadas *wavelet*, o que possibilita uma classificação de anomalias com base no domínio da frequência. A ferramenta proposta utiliza um algoritmo denominado *deviation score* para detectar anomalias, analisando as altas-frequências de eventos de curta duração e baixas frequências de eventos de longa duração. O trabalho demonstra se a análise de fluxos é suficiente para expor diferentes tipos de anomalias de tráfego de rede próximo a tempo real, por uma análise estatística dos dados coletados no roteador de borda de uma Universidade.

Jiang e Papavassiliou [42], assumem que atividades anormais diferem substancialmente das atividades normais de uma rede, ou seja, toda execução normal de uma atividade segue um determinado padrão, o qual pode ser adotado como um perfil de comportamento da rede que sinaliza qualquer desvio observado como uma anomalia. O trabalho utiliza uma base de dados históricos para cálculo de séries temporais unidimensionais, assumindo que a intensidade do tráfego pode ser calculada diretamente com a duração de cada atividade.

O trabalho desenvolvido por Wang *et al.* [43] cria uma discussão sobre o uso de métodos estatísticos para detecções de anomalias em *data centers*. O trabalho destaca a importância do uso de técnicas leves, ou seja, métodos simples que não afetam o desempenho e a eficácia da detecção. Os autores apresentam métodos baseados nos modelos de *Tukey* e entropia relativa para melhorar a precisão da detecção e diagnóstico de anomalias.

Uma abordagem sistemática para monitorar o tráfego de rede anormal é proposta por Qin *et al.* [44]. Comportamentos anormais, como o mau funcionamento do equipamento ou o uso indevido de recursos de rede, podem causar alterações significativas no fluxo normal. Os autores propõem um modelo de monitoramento de rede com base em uma análise de tráfego que identifica tais comportamentos, extraindo quatro características dos registros de fluxo para definição de um padrão. Esta abordagem extrai o comportamento anormal medindo as características do fluxo usando uma técnica chamada *Blind Source Separation* (BSS), sendo eficiente na identificação de comportamentos anormais que não induzem mudanças significativas no volume de tráfego.

Métodos para detecção de anomalias utilizando modelos de regressão em séries temporais têm sido consideravelmente investigados [45]. Um dos primeiros trabalhos envolvendo detecção de anomalias em séries temporais foi desenvolvido por meio de um processo auto-regressivo estacionário [46]. As anomalias eram reconhecidas comparando os dados reais, ou amostras, com a matriz de covariância do processo auto-regressivo.

Alguns trabalhos utilizam o modelo ARIMA para detecção de anomalias diretamente como [47] e [48].

Yaacob et al. [47] apresentam um sistema para detecção que gera alarmes para ataques do tipo *DoS*. Uma comparação direta da previsão gerada pelo modelo com o volume real (*bits/s*) é feita, evidenciando as mudanças de um comportamento normal para um comportamento anômalo. No entanto, os autores não fazem nenhuma investigação de alarmes falsos pertinentes aos resultados, gerados com base em um ambiente simulado, impossibilitando a implantação de um sistema real.

O trabalho desenvolvido por Zhu e Sastry [48] também utiliza um ambiente de simulação para geração de tráfego e anomalias, e explora as capacidades do modelo ARIMA para detecção de anomalias em uma dimensão (atributo *bits*). Os autores propõem um *framework* para estimação dos parâmetros do modelo. Além disso, uma técnica denominada *General Likelihood Ratio* é combinada com um filtro *Kalman* para melhorar a estimativa de parâmetros do modelo.

As Assinaturas Digitais de Segmento de Rede utilizando análise de Fluxos têm sido alvo de investigação no âmbito de caracterização de tráfego [49] e detecção de anomalias [50]. Abordagens pertencentes a diferentes grupos de algoritmos como estatísticos, heurísticos e de previsão têm provado a aplicação da ferramenta no gerenciamento de segurança.

Abordagens tradicionais para detecção de anomalias vêm sendo combinadas com as mais diversas técnicas, buscando melhorar a relação de falsos positivos identificados. Por exemplo, um sistema baseado em uma arquitetura de dois estágios combinando limite adaptável e soma cumulativa com a transformada *wavelet* é proposto em [51]. Uma combinação do algoritmo *k-Means* e árvores de decisão ID3 é proposta em [52]. O sistema proposto nesse trabalho combina as assinaturas digitais de segmento de rede com a Lógica Paraconsistente.

2.4 LÓGICA PARACONSISTENTE

A Lógica Paraconsistente (LP) têm sido estudada em diversos campos de aplicação, como os sistemas inteligentes [53], os circuitos digitais [54], e a medicina [55].

Guido *et al.* [56] desenvolve um modelo para reconhecimento de padrões capaz de representar o conceito de paraconsistência. A chamada *Discriminative Paraconsistent Machine* (DPM), é testada através de experimentos com análise de série temporal e detecção de patologias em vozes humanas. Torres *et al*. [54] emprega a LP na tomada de decisão para movimentação de um controlador para robôs. A Lógica Paraconsistente Anotada (LPA) é utilizada em [55] para auxiliar profissionais da área médica a detectar sintomas de hipertensão. A LPA também é utilizada em [53] na correlação de diversos critérios relacionados a transportes de carga.

Uma discussão sobre diversos aspectos da LP, que abrange desde sua terminologia até seu campo de aplicação é feito em Costa *et al.* [57]. A LP e suas extensões em geral trabalham o nível de certeza e contradição relacionados a solução de um problema. Dois

trabalhos que, apesar de não serem baseados na LP, investigam o níveis evidenciais na solução do problema de detecção de anomalias podem ser citados.

Kavitha *et al.* [58] lidam com o nível de incerteza na classificação de padrões anômalos com o uso da chamada lógica *neutrosófica*. A solução para segurança de rede é baseada em regras de classificação de intrusão, geradas com um algoritmo supervisionado e uma base de dados previamente classificada.

Chatzigiannakis *et al.* [59] utiliza uma abordagem para detecção de anomalias baseada na fusão de dados. Com base na perspectiva de alguns sensores, a fusão de dados classifica eventos na rede em diferentes grupos, gerando algumas hipóteses sobre o estado da rede. A teoria da evidência é então utilizada para modelar a incerteza gerada pelos sensores e detectar possíveis anomalias nas atividades da rede.

O presente trabalho apresenta uma proposta para integrar a LP na solução do problema de detecção de anomalias.

3 ASSINATURA DIGITAL DE SEGMENTO DE REDE UTILIZANDO ANÁLISE DE FLUXOS

A criação de um perfil para o tráfego normal de uma rede é de suma importância para diversas atividades da gerência de redes. Um grande desafio encontrado na caracterização do tráfego é a definição de um modelo capaz de representar adequadamente o conceito de normalidade, uma vez que o tráfego de rede apresenta um comportamento dinâmico, geralmente apresentando um nível considerável de ruídos. Este capítulo descreve os procedimentos adotados nesse trabalho para a caracterização de tráfego de rede, com o uso da Assinatura Digital de Segmento de Rede utilizando análise de Fluxos.

3.1 DEFINIÇÃO, APLICAÇÃO E FUNCIONAMENTO

A Assinatura Digital de Segmento de Rede utilizando análise de Fluxos, ou também DSNSF (do inglês *Digital Signature of Network Segment using Flow analysis*), pode ser definida como um conjunto básico de informações que constituem o perfil de operações normais do segmento ao qual ela se refere [60] e [61]. Tal perfil integra limiares mínimos e máximos que indicam qual o comportamento normal para o tráfego de rede ao longo de um dia determinado. Ainda mais, o conceito da assinatura digital aplica-se a qualquer equipamento ou dispositivo pertencente a uma rede plausível de monitoramento, como servidores, *switches* e roteadores.

A utilização da DSNSF visa auxiliar o administrador ou gerente a identificar limitações aparentes e a controlar o uso de recursos críticos, ajudando no planejamento estrutural da rede. Em respeito à detecção de anomalias, o comportamento normal é determinado pelos padrões de tráfegos especificados pela DSNSF. Qualquer evento que apresente um desvio significativo de tal especificação é considerado um potencial dano à rede, ou ainda, uma anomalia.

Uma importante característica do comportamento do tráfego de redes é a presença de ciclos diários que identificam determinadas atividades cotidianas de seu ambiente, diferindo para cada dia da semana. Por exemplo, o tráfego gerado em dias úteis, quando comparado com aquele gerado em fins de semana ou feriados tende a apresentar grande disparidade. Os mecanismos responsáveis pela caracterização do tráfego devem ser capazes de lidar com essa característica em particular. O sistema proposto nesse trabalho perfaz uma análise de dados históricos na rede, gerando DSNSFs específicos para cada dia da semana. Essa

abordagem é feita com intuito de diminuir a quantidade de erros consequentes de uma análise associando dias com comportamentos diferentes, como dias úteis e feriados.

Dentro do processo de criação das DSNSFs, a análise de fluxos é uma fase de preparação de dados responsável por extrair informações de fluxos IP, os chamados atributos de tráfego de rede. Essas informações servem como medidas das leituras reais do tráfego para um dia analisado, além de serem armazenadas periodicamente em uma base de dados. Tal base é utilizada por um determinado modelo estatístico, heurístico ou de mineração de dados para geração das DSNSFs.

Alguns modelos já foram investigados para geração de DSNSFs [49], [50], [62] e [63]. A possibilidade da utilização de diferentes modelos para geração de DSNSFs também foi um dos fatores que motivou a implementação de uma avaliação paraconsistente para detecção de anomalias, detalhada no Capítulo 4.

Esse trabalho aproveita os recursos deixados por trabalhos anteriores que sustentam a aplicação das assinaturas digitais na detecção de anomalias, mas também investiga e implementa um novo modelo, tradicionalmente utilizado na previsão de séries temporais.

É proposta a geração das DSNSFs com um modelo ARIMA, termo que deriva do inglês *AutoRegressive Integrated Moving Average* cuja tradução é modelo Auto-Regressivo Integrado de Médias Móveis. A principal característica do modelo ARIMA é sua capacidade de gerar previsões de séries temporais com base em análise histórica. A sua escolha foi motivada por trabalhos relacionados[62], por sua flexibilidade e por sua boa integração com a metodologia adotada nesse trabalho.

A Figura 3.1 ilustra a visão geral do sistema para detecção de anomalias proposto nesse trabalho. Os fluxos exportados por um *switch* ou roteador localizado em um segmento da rede são armazenados em um coletor de informações. Tal coletor é responsável por extrair um conjunto de atributos de tráfego e armazenar em uma base de informações históricas. É com base nessa base que os modelos adotados nesse trabalho são utilizados para geração das DSNSFs. As estimativas de cada assinatura digital e as leituras reais são relacionadas para que métricas paraconsitentes sejam obtidas e utilizadas na detecção de anomalias.

Da Seção 3.2 em diante são descritos conceitos e procedimentos para a preparação dos dados e adaptação do modelo ARIMA na geração de DSNSFs.

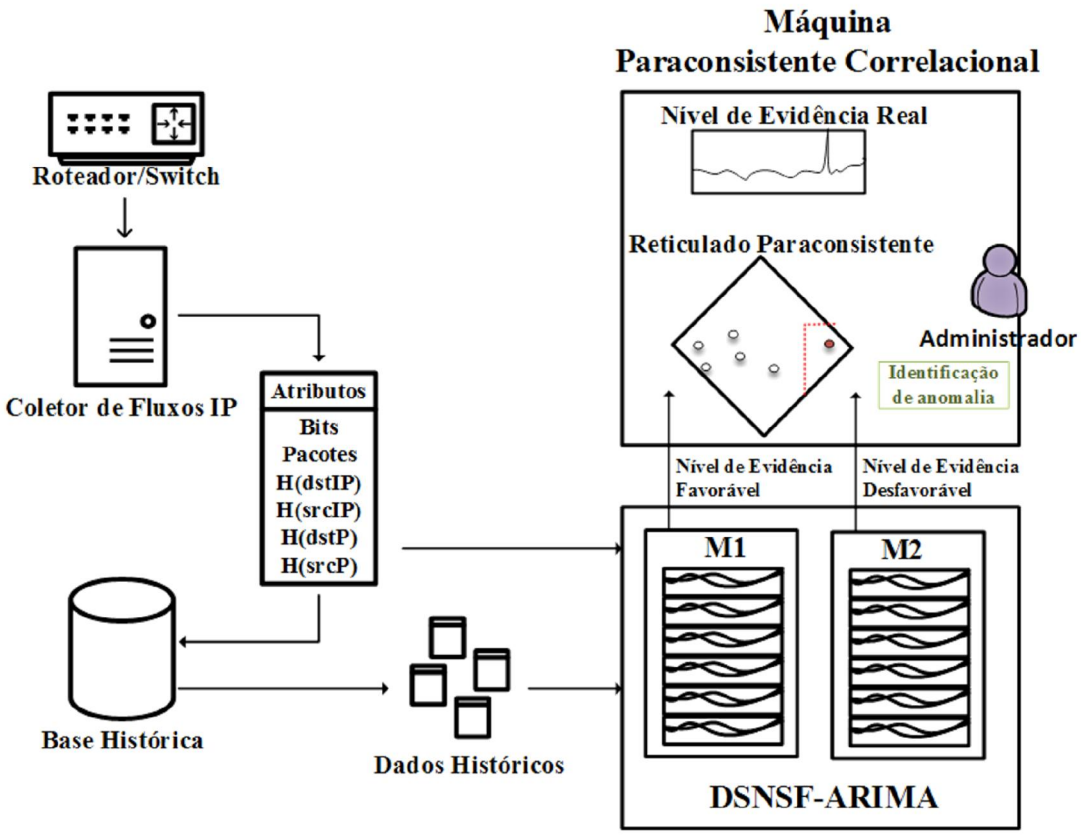


Figura 3.1 – Visão geral do sistema para detecção de anomalias proposto.

3.2 PREPARAÇÃO DOS DADOS

A fase de preparação dos dados é responsável por extrair algumas informações contidas em fluxos IP coletados em um determinado segmento. Além de ser a base para a geração das DSNSFs, essas informações devem auxiliar na abordagem de detecção de anomalias, evitando o aumento da complexidade e ineficiência na implantação em redes de larga escala. Dessa forma, esse trabalho busca conciliar o potencial de informação encontrado em fluxos com um baixo custo computacional.

As informações extraídas dos fluxos IP podem ser referidas como atributos de fluxos. Atributos relativos ao volume e dispersão de tráfego são extraídos minuto a minuto. Para atributos relativos ao volume de tráfego, *bits* e pacotes por segundo, são utilizadas médias aritméticas para o intervalo analisado. Os atributos relativos à dispersão de tráfego estão relacionados a endereços IP e aplicações, e necessitam de uma abordagem diferente daquela tomada com atributos de volume.

A análise de endereços IP de origem e destino bem como a de portas de origem e destino auxilia na detecção de anomalias, visto que a distribuição normal dos mesmos pode sofrer mudanças significativas em situações anômalas [36]. Informações relativas à dispersão do tráfego são diretamente retiradas do cabeçalho dos pacotes IP presentes nos fluxos analisados sendo: endereço IP de origem, endereço IP de destino, porta de origem e porta de destino. Uma forma de transformar tais informações em dados quantitativos que sumarizem o comportamento desses atributos é a utilização do ganho de informação, ou entropia. Este trabalho emprega a medida de entropia de Shannon [64], que determina o grau de concentração ou dispersão da distribuição dos atributos abordados. Dessa forma, um histograma a partir de amostras de cada atributo é gerado, dada uma sequência $A = \{n_1, \dots, n_i, \dots, n_N\}$, onde n_i é a quantidade de ocorrências da amostra i e N o número de elementos da sequência, a entropia H de A pode ser definida como na Equação (3.1). O termo s, definido com $s = \sum_{i=1}^{N} n_i$, é a soma das ocorrências presentes no histograma.

O valor da entropia é mínimo quando distribuições das amostras estão concentradas, nula quando todas as amostras são idênticas, e mais próximo de $\log N$ quando o grau de dispersão é elevado.

$$H(A) = -\sum_{i=1}^{N} \left(\frac{n_i}{s}\right) \log\left(\frac{n_i}{s}\right) \tag{3.1}$$

A Tabela 3.1 identifica os atributos utilizados nesse trabalho para geração das DSNSFs, em ordem de análise. Uma vez que as informações da rede estejam aptas para análise, um modelo capaz de gerenciar tais informações pode ser utilizado.

Tabela 3.1 – Conjunto de atributos da DSNSF.

Atributo	Referência	Descrição
Bits	Bits	Média de bits por segundo em um minuto
Pacotes	Pacotes	Média de pacotes por segundo em um minuto
Entropia IP de destino	H(dstIP)	Dispersão dos endereços IP de destino
Entropia IP de origem	H(orgIP)	Dispersão dos endereços IP de origem
Entropia porta de destino	H(dstP)	Dispersão dos números de porta de destino
Entropia porta de origem	H(orgP)	Dispersão dos números de porta de origem

3.3 ANÁLISE DE SÉRIES TEMPORAIS

Uma série temporal pode ser definida como um conjunto de observações sobre uma variável, ordenado no tempo, e registrado em períodos regulares [65]. Alguns exemplos simples de séries temporais incluem: temperaturas máximas e mínimas diárias em uma cidade, faturamento médio de uma empresa, valores de fechamentos diários da bolsa de valores, resultado de um eletrocardiograma, gráfico de controle de um processo produtivo, etc.

A análise de séries temporais é baseada em uma suposição básica: para um sistema causal relativamente constante, relacionado com o tempo, o qual exerceu influência sobre uma série de dados no passado, assume-se que o mesmo pode continuar a fazê-lo no futuro [65]. Esse sistema causal costuma gerar padrões não aleatórios de informações que podem ser detectadas através de um processo estatístico ou um gráfico de série temporal. Dessa forma, o principal objetivo da análise de séries temporais é a identificação padrões em uma variável de interesse. Uma vez observado, seu comportamento passado auxilia nas previsões sobre o futuro, orientando a tomada de decisões.

Uma série temporal pode ser composta por alguns elementos como tendência, ciclo, sazonalidade e variação irregular. A tendência diz respeito ao comportamento de longo prazo da série, podendo ser causada por diversos aspectos que afetem a variável de interesse em longo prazo. O ciclo, ou variação cíclica, é uma definição para flutuações nos valores da variável observada com duração superior a um ano, e que se repetem com certa periodicidade [66]. Sob o mesmo ponto de vista, as variações sazonais, ou sazonalidade, representam flutuações nos valores da variável de interesse com duração inferior a um determinado período. Finalmente, as variações irregulares representam flutuações inexplicáveis resultando de eventos inesperados. Contudo, nem sempre uma série temporal apresenta todos os componentes citados, dependendo muito do tipo de problema que está sendo modelado. A decomposição de uma série pode identificar quais componentes estão atuando em uma observação em particular [66].

Outra característica a ser considerada na análise de séries temporais é sua estacionariedade. Uma série temporal é considerada estacionária quando se desenvolve no tempo em torno de uma média constante, refletindo uma forma de estabilidade regular. No entanto, grande parte das séries encontradas na prática apresenta alguma forma de não estacionariedade.

3.3.1 Séries temporais na caracterização do tráfego

A caracterização do tráfego feita nesse trabalho é, em termos gerais, um tratamento de diversas séries temporais. A fase de preparação de dados é responsável por produzir séries para cada um dos seis atributos analisados, para cada dia, a partir do tráfego agregado em um segmento da rede em cada intervalo de tempo. Assim, ao final de um dia têm-se seis séries temporais que representam as leituras reais do tráfego. A representação dessas séries pode ser feita por uma matriz R de ordem $\boldsymbol{T} \times \boldsymbol{F}$, onde $\boldsymbol{T}$ representa o número de períodos ou intervalos da série e $\boldsymbol{F}$ o número de atributos analisados. A ordem em que os atributos são analisados é sempre a mesma (Bits, Pacotes, H(dstIP), H(orgIP), H(dstP), H(orgP)). A matriz R pode ser representada como a seguir:

$$R = \begin{bmatrix} \overbrace{r_{1,1}}^{\text{Bits}} & \overbrace{r_{1,2}}^{\text{Pacotes}} & \overbrace{r_{1,3}}^{\text{H(dstIP)}} & \overbrace{r_{1,4}}^{\text{H(orgIP)}} & \overbrace{r_{1,5}}^{\text{H(dstP)}} & \overbrace{r_{1,6}}^{\text{H(orgP)}} \\ r_{2,1} & r_{2,2} & r_{2,3} & r_{2,4} & r_{2,5} & r_{2,6} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{1440,1} & r_{1440,2} & r_{1440,3} & r_{1440,4} & r_{1440,5} & r_{1440,6} \end{bmatrix}$$

Esse trabalho considera apenas séries temporais discretas, ou seja, observações são feitas em períodos discretos $1,2,\dots,t,\dots,\boldsymbol{T}$. Além disso, as observações são realizadas em um intervalo fixo de um minuto, assim: $\boldsymbol{T} = 1440$, para um dia completo. Como são analisados seis atributos, tem-se $\boldsymbol{F} = 6$. Cada uma das colunas $r_{t,f}$, representa a leitura real do tráfego para um relativo atributo $(1,2,\dots,f,\dots,\boldsymbol{F})$ para todos um intervalo $(1,2,\dots,t,\dots,\boldsymbol{T})$ analisado.

As leituras reais do tráfego possuem dois propósitos. Seguindo a ordem da metodologia proposta, uma base de informações precisa ser utilizada para que os modelos ARIMA sejam construídos. Assim, algumas semanas de monitoramento são necessárias para que uma base de informação seja construída. Ao término de cada dia, as estruturas R são armazenadas. A base construída servirá para o treinamento de modelos ARIMA.

Assumindo modelos ARIMA adequados, cada nova leitura real de tráfego é confrontada com os mesmos por meio da Máquina Paraconsistente Correlacional (MPC) além de ser armazenada na base histórica para posterior uso em novos modelos.

Na discussão da modelagem ARIMA, séries temporais são representadas nesse trabalho com os termos Y e Z. Isso se faz apenas para convenção e simplicidade da

notação de um processo genérico. As colunas de R possuem o mesmo tratamento, apenas com símbolos diferentes.

3.4 Modelo Auto-Regressivo Integrado de Médias Móveis

Esta seção descreve os principais componentes do modelo Auto-Regressivo Integrado de Médias Móveis (*AutoRegressive Integrated Moving Average* - ARIMA). Tal modelo se firmou como uma importante ferramenta de previsão, sendo a base de muitos fundamentos na análise de séries temporais [66]. O seu uso se popularizou em diversas áreas, devido a sua amplitude de uso e versatilidade para uma variedade de situações.

O modelo ARIMA tem como objetivo principal prever o comportamento de séries estacionárias e não estacionárias. Uma série é considerada estacionária quando sua média e variância são constantes ao longo do tempo, e quando o valor de sua covariância depende de não mais do que dois períodos de tempo para composição. Entretanto, grande parte das séries representa algum tipo de não estacionariedade, com um deslocamento da média e a variância [66].

A metodologia de previsões com modelo ARIMA permite que valores futuros de uma determinada série temporal sejam previstos com base nos valores atuais e passados, através de uma correlação temporal. Essa metodologia se tornou muito popular com os trabalhos de Box e Jenkins [9], publicados a partir da década de 1970. Segunda ela, a construção de um modelo ARIMA pode ser dividida em três etapas: identificação, estimação e previsão.

Construir um modelo ARIMA pode se tornar uma tarefa árdua em alguns casos. Muitos parâmetros precisam ser ajustados e cada modelo precisa ser verificado. Além disso, esse trabalho lida com um grande número de séries temporais para a construção de DSNSFs. Dessa forma, faz-se necessária a descrição de um processo mais automatizado para seleção e estimação dos modelos ARIMA utilizados.

Primeiramente, são apresentadas as propriedades teóricas envolvidas nos modelos ARIMA com base nos trabalhos de Box e Jenkins [9], Makridakis e Hibon [67], entre outros [65; 66]. Em seguida, o processo computacional envolvido desde a seleção de modelos até a previsão de valores é descrito. Algumas diretrizes apresentadas em [68] e [45] foram adequadas para tornar todo o processo de modelagem o mais automatizado possível.

O modelo ARIMA é uma composição de dois modelos mais simples, o auto-regressivo e o de médias móveis, em conjunto com um processo de diferenciação.

3.4.1 O modelo auto-regressivo

O modelo auto-regressivo (do inglês, *AutoRegressive*), pode ser representado genericamente conforme a Equação (3.2), onde p determina o número de passos entre observações passadas. Os termos $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$ são independentes de α_t, chamado ruído branco. Dessa forma, é valido assumir que os valores da série Y_t são uma combinação linear dos p valores passados da série mais um termo α_t. Os termos $\phi_1, \phi_2, \ldots, \phi_p$ são parâmetros a serem estimados na construção do modelo. O processo auto-regressivo pode ser sucintamente definido por $AR(p)$.

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + a_t \tag{3.2}$$

3.4.2 O modelo de médias móveis

O modelo de médias móveis (do inglês, *Moving* Average) pode ser representado conforme a Equação (3.3), sendo uma combinação linear dos q termos de ruído branco passados. O modelo também pode ser referido como $MA(q)$.

$$Y_t = a_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \cdots - \theta_q \alpha_{t-q} \tag{3.3}$$

Esta terminologia refere-se ao fato de que o termo Y_t é obtido quando os pesos $1, -\theta_1, -\theta_2, \ldots, -\theta_q$, são aplicados às variáveis $\alpha_t, \alpha_{t-1}, \alpha_{t-2}, \ldots, \alpha_{t-q}$. O ruído branco, α_t, é uma sequência de variáveis aleatórias independentes. O modelo de médias móveis também pode representado por $MA(q)$.

3.4.3 O Modelo auto-regressivo de médias móveis

Considerando uma série formada por partes auto-regressivas e de médias móveis, um modelo mais geral de séries temporais pode ser representado como na Equação (3.4), o chamado modelo auto-regressivo de médias móveis, $ARMA(p, q)$, de ordens p e q .

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + a_t - \theta_1 \alpha_{t-1} - \theta_2 \alpha_{t-2} - \cdots - \theta_q \alpha_{t-q} \tag{3.4}$$

Os modelos apresentados até o momento são adequados para séries estacionárias, ou seja, aquelas onde a média é constante por todo tempo. Em geral, muitas séries apresentam algum tipo de não estacionariariedade [65].

3.4.4 O modelo auto-regressivo integrado de médias móveis

Antes da descrição do modelo Auto-Regressivo Integrado de Médias Móveis (ARIMA), define-se dois operadores comumente utilizados. O primeiro deles é o operador de defasagem (B), que representa o valor de uma variável aleatória em um número de períodos precedentes, $B^i Y_t = Y_{t-i}$, para $i = \{1,2,3, \dots\}$. O segundo deles é o operador diferença, onde $\Delta^d = (1 - B)^d$. Por exemplo, para $d = 1 Y_t - Y_{t-1} = Y_t - BY_t = (1 - B)Y_t = \Delta Y_t$.

Uma série Y pode ser diferenciada algumas vezes por um número finito d, tornando se estacionária em Z_t ($Z_t = \Delta^d Y_t$). Dessa forma, Z_t pode ser representada por um modelo ARMA(p, q), conforme Equação (3.5). Para simplificação coloca-se a parte autoregressiva do lado esquerdo da equação e toma-se $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ e $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$.

$$\phi(B) Z_t = \theta(B) \alpha_t, \tag{3.5}$$

Se Z_t é uma diferença de Y_t, então Y_t é integral de Z_t, sendo considerado um modelo auto-regressivo integrado de médias móveis descrito na Equação (3.6):

$$\phi(B) \Delta^d Y_t = \theta(B) \alpha_t \tag{3.6}$$

Esse modelo também pode ser referenciado por $ARIMA(p, d, q)$, de ordem (p, d, q), quando p e q são as ordens de $\phi(X)$ e $\theta(X)$, respectivamente, e d o número de diferenças.

3.4.5 Propriedades estatísticas para os modelos ARIMA

Algumas medidas da teoria da probabilidade são necessárias para identificação e estimação dos modelos AR, MA e ARIMA. Seguindo as definições propostas por Box e Jenkins [9], elas são descritas diretamente para aplicação na geração dos modelos.

Seja o conjunto de observações Y, a média desse processo é definida com a Equação (3.7), para um total de observações N.

$$\bar{Y} = \frac{1}{N}\sum_{t=1}^{N} Y_t \tag{3.7}$$

A Variância desse conjunto pode ser representada com σ_Y^2, sendo calculada conforme Equação (3.8).

$$\sigma_Y^2 = \frac{1}{N}\sum_{t=1}^{N} (Y_t - \bar{Y})^2 \tag{3.8}$$

A função de covariância entre duas observações Y_t e Y_{t+k}, para $k \in \mathbb{N}^*$, separada por k intervalos (B^k) pode ser estimada com a Equação (3.9)

$$c_k = \frac{1}{N}\sum_{t=1}^{N-k} (Y_t - \bar{Y})(Y_{t+k} - \bar{Y}) \tag{3.9}$$

O cálculo da função de covariância é utilizado na estimação da Função de AutoCorrelação (FAC) para um determinado passo ou defasagem k, estimada com a Equação (3.10).

$$\rho_k = \frac{c_k}{c_0} \tag{3.10}$$

Outra medida utilizada no processo de identificação dos modelos é a Função de AutoCorrelação Parcial (FACP). Essa medida corresponde a correlação deY_t eY_{t-k} eliminando a dependência das observaçõesY_{t-1}, Y_{t-2} e Y_{t-k+1}. O seu cálculo é representado pelos mesmos coeficientes do processo auto-regressivo ϕ, sendo que ϕ_{kj} é o j-éssimo coeficiente de ordem k, $k \in \mathbb{N}^*$. Para definir a função de autocorrelação faz-se necessário considerar uma estrutura autoregressiva $AR(p)$, como a da Equação (3.11).

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \cdots + \phi_p\rho_{k-p} \tag{3.11}$$

para $\phi(B)\rho_k = 0$, e assumindo $\rho_k = \rho_{-k}$. Fazendo $k = 1,2, \ldots p$, um sistema conhecido como equações de Yule-Walker [66] é obtido, e sua representação é vista na Equação (3.12).

$$\begin{array}{ccccccc} \rho_1 & = & \phi_1 & + & \phi_2\rho_1 & +\cdots+ & \phi_p\rho_{p-1} \\ \rho_2 & = & \phi_1\rho_1 & + & \phi_2 & +\cdots+ & \phi_p\rho_{p-2} \\ \vdots & & \vdots & & \vdots & \ldots & \\ \rho_p & = & \phi_1\rho_{p-1} & + & \phi_2\rho_{p-2} & +\cdots+ & \phi_p \end{array} \tag{3.12}$$

É possível se obter estimativas para os parâmetros ϕ_i's com a substituição dos valores da função de autocorrelação nas equações de Yule-Walker. Ao denotar os estimadores de cada ϕ_i por ϕ_{ii}, a autocorrelação parcial é dada pela sequência dos ϕ_{kk} obtidos pela resolução das equações de Yule-Walker reescritas como na Equação (3.13).

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix} \tag{3.13}$$

Resolve-se o sistema linear (Equação 3.13) em termos de determinantes, para $k = 1,2, \ldots$, sucessivamente, e obtêm-se os termos:

$$\phi_{11} = \rho_1$$

$$\phi_{22} = \frac{\begin{matrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{matrix}}{\begin{matrix} 1 & \rho_1 \\ \rho_1 & 1 \end{matrix}}$$

$$\phi_{22} = \frac{\begin{matrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{matrix}}{\begin{matrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{matrix}}$$

Que no geral podem ser representados pela Equação (3.14).

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|} \tag{3.14}$$

Onde P_k é a matriz de correlação e P_k^* é a matriz P_k com a última coluna substituída pelo vetor de autocorrelações. Resumidamente, a função de autocorrelação parcial de Y_t para uma defasagem k é representada por ϕ_{kk}, uma vez que esse termo é igual a correlação parcial entre as variáveis Y_t e Y_{t-k}, ajustados por intermédio das variáveis $Y_{t-1}, Y_{t-2}, \dots, Y_{t-k+1}$ [66].

Uma série de estudos discute as propriedades das funções de autocorrelação e autocorrelação parcial. Esse trabalho adota as estimativas apresentadas em [9] e [69].

3.4.6 Identificação dos modelos

Ao utilizar as Funções de AutoCorrelação (FAC) e de AutoCorrelação Parcial (FACP) é possível identificar quais as melhores ordens para os processos auto-regressivo s e de médias móveis.

3.4.6.1 Definição das diferenças d

Quando a série analisada apresenta algum tipo de não-estacionariedade, faz-se necessário algum tipo de diferenciação, uma vez que as autocorrelações não terão valores satisfatórios. Nesse caso o operador de diferenciação (Δ^d) é aplicado sucessivamente até que a série represente um processo estacionário, ou seja, seus valores oscilem ao redor de uma média constante. A série original é diferenciada para que o processo $\Delta^d Y_t$ possa ser tratado por meio de um modelo $ARMA(p,q)$. O número de diferenças, d , necessárias para que o processo se torne estacionário, é alcançado quando a função de autocorrelação de $Z_t = \Delta^d Y_t$ decresce rapidamente para zero. A principal razão para se fazer essa transformação a estabilização da variância.

3.4.6.2 Definição das ordens p e q

Com a série já diferenciada é necessária a escolha das ordens de p e q para $\phi(B)$ e $\theta(B)$. Para isso o comportamento das Funções de AutoCorrelação (FAC) e das AutoCorrelações Parciais (FACP) são analisadas. Esse trabalho acompanha a metodologia

proposta em [9], seguindo as propriedades e as características descritas no mesmo para a identificação teórica das ordens de p e q. A Tabela 3.2 descreve o comportamento teórico dos modelos $AR(p)$, $MA(q)$ e $ARMA(p,q)$ e a Tabela 3.3 descreve o comportamento dos modelos mais utilizados nesse trabalho.

O valor de p é determinado a partir da FACP da série diferenciada. Se a FACP se interrompe (tende a zero) depois de alguns passos de defasagem, o último valor da defasagem com valor alto é utilizado como p. Caso a função não se interrompa existem duas possibilidades: $p = 0$ ou um modelo $ARMA$ com valores positivos de p e q. Similarmente, o valor da FAC é utilizado para determinar o valor de q, se a FAC se interrompe depois de alguns passos de defasagem, o último valor da defasagem é utilizado como q. Caso contrário, as possibilidades $p = 0$ ou modelo $ARMA$ com p e q positivos devem ser investigadas.

Tabela 3.2 – Comportamento teórico dos modelos $AR(p)$, $MA(q)$ e $ARMA(p,q)$.

	Função de AutoCorrelação (FAC)	Função de AutoCorrelação Parcial (FACP)
$AR(p)$	Exponenciais ou senoidal amortecida. Não se anula bruscamente.	Finita. Anula-se bruscamente no passo p.
$MA(p)$	Finita. Anula-se bruscamente no passo q.	Dominada por exponenciais e/ou senoidal amortecidas. Não se anula bruscamente.
$ARMA(p,q)$	Exponenciais e/ou senoidal amortecidas para $k > q - p$. Não se anula bruscamente.	Exponenciais e/ou senoidal amortecidas para $k > q - p$. Não se anula bruscamente.

Quando tanto a FAC e a FACP não se interrompem, um modelo do tipo $ARMA(p,q)$ com p e q positivos é sugerido. As funções resultantes podem apresentar comportamento que variam entre decrescimento exponencial e senoidais amortecidas após os passos $q - p$.

Tabela 3.3 – Comportamento dos modelos mais recorrentes.

	Função de AutoCorrelação (FAC)	Função de AutoCorrelação Parcial (FACP)
$AR(1)$	Decresce exponencialmente.	Pico no passo 1 e $\phi_{11} \neq 0$.
$MA(1)$	Pico no passo 1 e. $\rho_1 \neq 0$.	Decresce exponencialmente.
$MA(2)$	Pico nos passos 1 e 2. $\rho_1 \neq 0 \rho_{\ 2} \neq 0$.	Exponenciais ou ondas senoidais amortecidas.
$ARMA(1,1)$	Decresce exponencialmente após o passo 1.	Decresce exponencialmente após o passo 1.
$ARMA(2,1)$	Decresce exponencialmente após o passo 2.	Decresce exponencialmente após o passo 2.

A fase de identificação de modelos é um passo um tanto quanto subjetivo [65]. Nesse trabalho algumas medidas são tomadas para tornar o processo mais automatizado, em nível de *software*. Os modelos vistos na Tabela 3.3 são os mais recorrentes na fase de identificação feita ao longo desse trabalho. Além disso, diferenças superiores a 1 não foram necessárias. Assim, depois de extensa análise das séries as ordens indicadas são prefixadas e a estimativa de parâmetros é feita para todos elas em paralelo. Os modelos cujas estimativas geram os menores erros entre observação são os escolhidos para previsão. O apêndice A ilustra alguns modelos e comportamentos das séries analisadas.

3.4.7 Estimação dos parâmetros ϕ e θ

Uma vez que as ordens do modelo $ARIMA(p, d, q)$ são definidas, os parâmetros ϕ e θ podem ser estimados. Esse trabalho segue um processo de estimação visto em Box e Jenkins [9] chamado de Máxima Verossimilhança. Para treinamento do modelo são utilizados dados (atributos na base histórica) das semanas que antecedem a previsão.

Para um dado conjunto de observações $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_N$, a função de verossimilhançaL é definida como a função densidade de probabilidade conjunta dos dados observados, em função dos parâmetros do modelo. Os estimadores de máxima verossimilhança são definidos como aqueles valores dos parâmetros que são os mais próximos ou prováveis, com relação aos dados observados (contidos na base histórica).

Considerando um modelo $ARIMA\ (p, d, q)$, com seus parâmetros alocados no vetor $\xi = (\phi, \theta, \sigma_a^2)$, onde $\phi = (\phi_1, \phi_2, \phi_3, \dots, \phi_p)$ e $\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_q)$, e as N observações $Y_1, Y_2, Y_3, \dots, Y_N$. A função de verossimilhança $L = (\xi | Y_1, \dots, Y_n)$ é considerada como uma função de ξ, então deve-se encontrar os valores de ξ que minimizem L. Inicialmente, as d diferenças das séries são tomadas, assim tem-se $N - d$ observações.

Os valores iniciais da série a_t precisam ser obtidos. Para isso valores presentes no dados históricos são utilizados. Supondo a_t, a função densidade conjunta de $a_1, \dots, a_N$ é dada pela Equação (3.15).

$$f(a_1, \dots, a_N) = (2\pi)^{-\frac{N}{2}} (\sigma_a)^N exp\left\{ - \sum_{t=1}^{N} \frac{a_t^2}{2\sigma_a^2} \right\} \tag{3.15}$$

Supondo os p valores Z_t (Y_t, depois de diferenciada) e q valores a_t (da base histórica), denotados por Z_t^* e a_t^*. A função de verossimilhança é dada por (3.16):

$$L = (\xi | Z, Z^*, a^*). \tag{3.16}$$

Segue que maximizar $L = (\xi | Z, Z^*, a^*)$ é o equivalente a minimizar $S(\delta | Z, Z^*, a^*)$, onde $\delta = (\phi, \theta)$. Os estimadores obtidos minimizando $S(\delta)$ serão aproximações para os estimadores de máxima verossimilhança. Para calcular a soma de quadrados para um dado δ, utiliza-se Z e a, gerando valores para o início da série e iterando os valores dos parâmetros com um valor pequeno (por exemplo: $0.01, 0.02$, e assim por diante). Essa forma de estimação fornece uma representação estacionária na qual Z_t é expressa somente em termos de valores futuros para Y_t e a_t, com a devida inversão para as diferenças d.

3.4.8 Padrões para o ruído branco

Alguns métodos tradicionais são utilizados para a distinção dos padrões de ruído branco nos modelos $ARIMA$ baseando-se na série temporal corrente. No entanto, esse trabalho aproveita os conceitos das assinaturas digitais e propõe uma direção que utiliza dados históricos para a definição de cada a_t. Esse aproveitamento é possível justamente devido a flexibilidade e adaptabilidade do modelo estudado.

Levando em consideração que os processos $ARIMA$ empregados nesse trabalho modelam séries temporais univariadas, então, um processo singular é utilizado para cada atributo estudado. Assim, seja um determinado intervalo t e um determinado atributo f. O escalar $r_{t,f}$ é relacionado na modelagem com Y_t. Têm-se duas possibilidades para seu relativo a_t: com base na mediana ou com base na moda de valores passados. Um exemplo é a melhor forma de ilustrar a situação. Seja $t = 1$ e $f = 1$, ou seja, o primeiro intervalo analisado para o primeiro atributo (Bits), representado por $r_{1,1}$. Supondo que o dia analisado seja uma segunda-feira, os valores passados considerados para o cálculo da moda e mediana são $r_{1,1}$'s referentes a segundas-feiras das semanas passadas ao dia analisado. A moda e a mediana são calculadas para esses valores, obtendo-se assim os ruídos brancos (a_t). Quando a média é utilizada para o cálculo tem-se: M1 → "Modelo ARIMA adaptado para média"; quando a mediana tem-se: M2 → "Modelo ARIMA adaptado para mediana". Todos os processos da modelagem $ARIMA$ são calculados para cada um dos modelos, essa abordagem permitirá que dois resultados diferentes sejam utilizados na abordagem paraconsistente.

A escolha do número de semanas históricas utilizadas é baseada nos resultados obtidos com previsões utilizando diferentes números de semanas. Esses resultados são interpretados por meio da diferença absoluta entre as séries temporais geradas (NMSE) e a correlação (mais detalhes destes resultados no Capítulo 5). As Figuras 3.2 e 3.3 apresentam os resultados obtidos para números de semanas que variam entre 1 e 7 para o ruído branco gerado com a média e mediana histórica, respectivamente. Para esse experimento foi utilizado a média da correlação e NMSE entre DSNSF e as leituras reais da semana que segue o limite dos dados de treinamento. Além disso, para que os testes fossem adequados para os dois ambientes testados nesse trabalho a média obtida nos mesmos foi apresentada.

A escolha do número de semanas também deve levar em consideração a viabilidade de implementação. A convergência do resultado a partir da quarta semana e a diferença dos resultados de quatro semanas em relação aos melhores valores apresentados variam entre 0.2% a 0.8% para NMSE e 0.3% a 1.2% para correlação.

O número de semanas escolhido não deve ser tão pequeno ao ponto de qualquer alteração em poucos dias analisados serem inferidas ao DSNSF. Por outro lado, valores elevados, além de terem um custo de processamento maior, fariam com que eventos legítimos demorassem mais tempo para serem incorporados ao perfil normal de comportamento. Dessa forma, padrões de ruído branco gerados com quatro semanas de base

histórica são suficientes para geração de DSNSFs aptos para assimilar de forma eficiente o comportamento operacional do tráfego de rede.

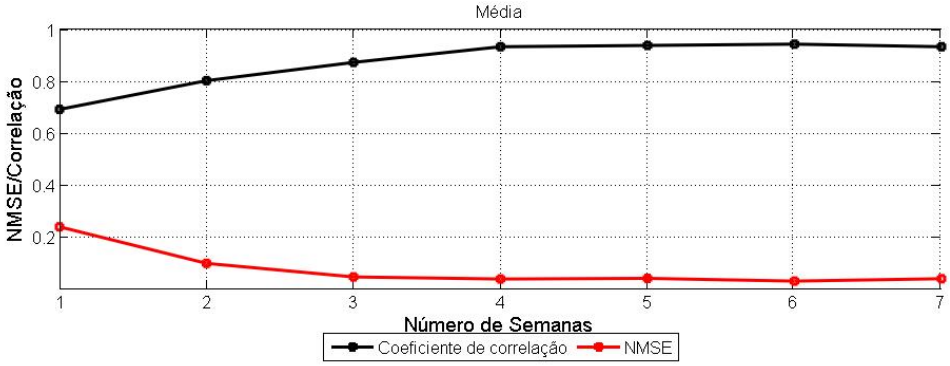


Figura 3.2 – Escolha do número de semanas para ruído branco baseado na média.

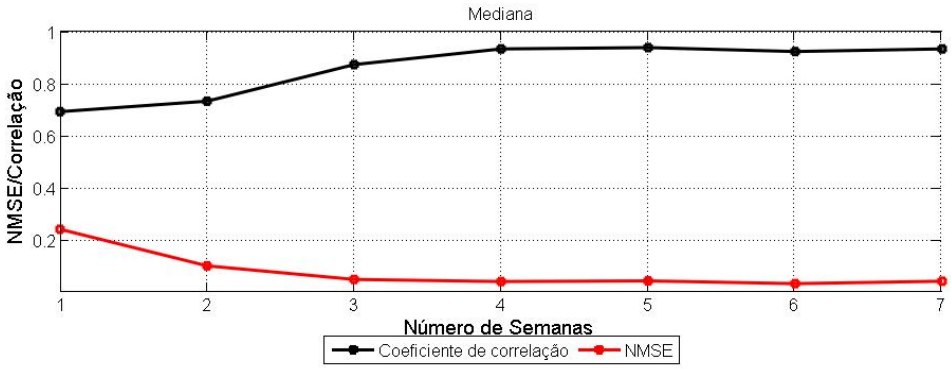


Figura 3.3 - Escolha do número de semanas para ruído branco baseado na mediana.

3.4.9 Visão geral da metodologia ARIMA

A metodologia adotada para construção de modelos ARIMA é dividida em três etapas: identificação, estimação e previsão. A Figura 3.4 ilustra o processo.

A etapa de identificação é baseada no comportamento das Funções de AutoCorrelação (FAC) e AutoCorrelação Parciais (FACP). Elas permitem identificar o quanto as observações dos períodos anteriores estão relacionadas com a observação do período seguinte. A observação dos resultados das funções para séries em seus estado original e diferenciado fornecem uma estimativa para a ordem dos parâmetros p, d e q. A série original juntamente com suas funções FAC e FACP são plotadas; a estacionariedade da mesma é obtida quando a FAC não termina bruscamente. A diferença entre a média da série e seus valores é outra opção para investigar a estacionariedade da série.

A etapa de estimação é responsável por escolher os valores de ϕ e θ. Utilizando a técnica de máxima verossimilhança é possível encontrar valores que produzem os menores erros de previsão possível. Quando os modelos estimados produzem erros muito elevados é possível que uma nova seleção de modelos aconteça e todo o restante do processo é iniciado. Finalmente, com o modelo identificado e estimado as previsões podem ser feitas de acordo com cada nova leitura real adquirida pelo sistema.

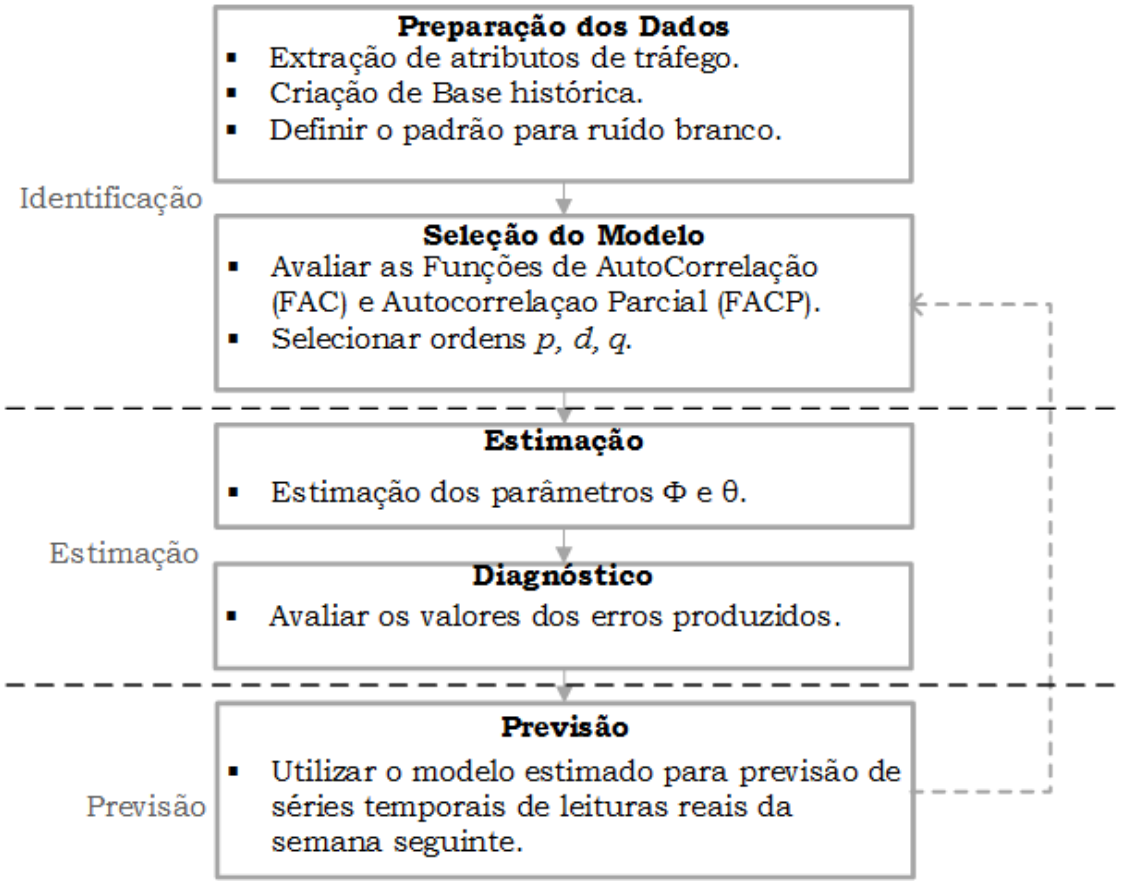


Figura 3.4 - Diagrama da utilização de modelos ARIMA.

4 DETECÇÃO DE ANOMALIAS COM A LÓGICA PARACONSISTENTE

Esse capítulo apresenta os aspectos fundamentais da Lógica Paraconsistente (LP) e uma de suas extensões, a Lógica Paraconsistente Anotada de anotação com dois valores (LPA2v), e como ela foi empregada na arquitetura da Máquina Paraconsistente Correlacional (MPC). Ao final do capítulo, um exemplo numérico é apresentado e o funcionamento do sistema proposto é resumido.

4.1 TRATAMENTO DE INCERTEZAS

Um recorrente problema entre sistemas de computação que modelam e analisam situações do mundo real é a grande quantidade de dados obtidos em situações não ideais, o que confere certo grau de incerteza em todo processamento de informação. Em virtude disso, é desejável que os projetos de tais sistemas estejam preparados para administrar situações adversas. Não raro, sistemas de processamento de dados devem ser capazes de manipular, representar e transportar informações imprecisas, incompletas, parcialmente ignoradas, ou até mesmo contraditórias [70].

Algumas abordagens buscam amenizar, de algum modo, o impacto que informações inconsistentes causam no processo computacional como um todo [53]. Dentre as mais tradicionais, estão presentes: a regra de Bayes e a regra modificada de Bayes [71]; Teoria da Confirmação [72]; Teorema de Dempster-shafer [73]; Lógica difusa [12]; entre outras. Segundo Costa et al. [57], em muitos casos, os procedimentos adotados por tais abordagens são baseados em procedimentos e representações de regras que não são suportadas por uma teoria bem definida ou não possuem propriedades semânticas estabelecidas. O presente trabalho propõe uma avaliação e modelagem de incertezas com base nos conceitos fundamentais da Lógica Paraconsistente.

A Lógica Paraconsistente (LP) tem sua origem em estudos sistematizados que investigam a possibilidade de restringir ou até mesmo rejeitar um dos fundamentos básicos da lógica clássica, o princípio da não contradição [57]. Em linhas gerais, esse princípio afirma que uma fórmula e sua negação não podem ser ambas as verdades, ou ainda, uma dada proposição com conotação verdadeira não pode ser falsa e uma proposição com notação falsa não pode ser verdadeira. A LP é uma lógica não clássica que admite contradição e que explora características evidenciais de uma determinada proposição.

Quando um sistema inteligente é utilizado para tomada de decisão, esse deve ser bem estabelecido em aspectos teóricos. A Lógica Paraconsistente fornece recursos para construção de um sistema que suporta o tratamento adequado de incertezas, seguindo algumas considerações [70]:

- *Resultados apresentam uma interpretação satisfatória.* As medidas de incerteza utilizadas possuem significado, clareza e precisão adequada para justificar decisões.
- *Capacidade de lidar com informação imprecisa*. As medidas de incerteza modelam possíveis conflitos de informação ou informação limitada. Além disso, combinações e operações com tais medidas podem ser feitas sob o domínio de regras bem estabelecidas.
- *Viabilidade computacional*. A modelagem e tratamento de incertezas opera com valores computacionalmente concretos.

A Lógica Paraconsistente possui duas extensões mais comuns: a Lógica Paraconsistente Anotada (LPA) e a Lógica Paraconsistente Anotada de anotação com dois Valores (LPA2v). O presente trabalho utiliza os fundamentos teóricos da LPA2v para implementação do sistema proposto.

4.2 LÓGICA PARACONSISTENTE ANOTADA DE ANOTAÇÃO COM DOIS VALORES (LPA2V)

A Lógica Paraconsistente Anotada de anotação com dois Valores (LPA2v) pode ser vista como uma extensão da Lógica Paraconsistente e possui fundamentos utilizados na extração de medidas paraconsistentes quantitativas, a serem combinadas computacionalmente em uma aplicação [53].

As evidências empregadas na aplicação LPA2v proposta nesse trabalho são provenientes de duas fontes de informação. A primeira delas produz um sinal que representa evidências favoráveis a uma proposição, assim:

$$\mu \longrightarrow Níveldeevidênciafavorável$$

Similarmente, a segunda fonte de informação produz um sinal que representa evidências desfavoráveis a uma proposição:

$$\lambda \longrightarrow Níveldeevidênciadesfavorável$$

O par de evidências (μ, λ) são valores reais, contidos no intervalo fechado $[0,1]$. Considerando uma proposição p e seu relativo par de evidências (μ, λ) , a anotação $p(\mu, \lambda)$ corresponde, intuitivamente, a seguinte leitura: "a proposição p apresenta um nível de evidência favorável μ , e um nível de evidência desfavorável λ ".

O seguinte exemplo é utilizado para contextualizar o presente trabalho com a análise da LPA2v. Dada a seguinte proposição p : "O intervalo t contém uma anomalia". As anotações a seguir sugerem as seguintes conotações:

$p(1.0, 0.0)$: "evidência totalmente positiva de que existe uma anomalia no intervalo t, ou seja, conotação lógica verdadeira";

$p(0.0, 1.0)$: "evidência totalmente negativa de que existe uma anomalia no intervalo t, ou seja, conotação lógica falsa";

$p(1.0, 1.0)$: "evidência total positiva e evidência total negativa de que existe uma anomalia no intervalo t, ou seja, conotação lógica inconsistente";

$p(0.0, 0.0)$: "evidência nula positiva e evidência nula negativa de que existe uma anomalia no intervalo t, ou seja, conotação lógica indeterminada".

Os níveis de evidência favoráveis e desfavoráveis de uma proposição são relacionados na análise paraconsistente. Muitos trabalhos utilizam o reticulado de Hasse para melhor representação e utilização das anotações no tratamento das incertezas, em função do nível de certeza e nível de contradição [10], [53] e [56]. Esse reticulado auxilia na interpretação dos resultados de uma proposição paraconsistente, sendo representado por um Quadro Unitário no Plano Cartesiano (QUPC), como ilustrado na Figura 4.1.

Com base no sistema de coordenadas do QUPC, o nível de evidência favorável é relacionado com o nível de evidência desfavorável. Primeiramente, com o nível de certeza, dado pela Equação (4.1).

$$\eta_c = \mu - \lambda \tag{4.1}$$

Outra relação é feita por meio do nível de contradição, dado pela Equação (4.2).

$$\eta_{ct} = \mu + \lambda - 1 \tag{4.2}$$

O nível de certeza é alocado horizontalmente no reticulado, enquanto o nível de contradição no eixo vertical. Ambos os valores pertencem ao conjunto dos números reais, dentro do intervalo fechado [-1,1].

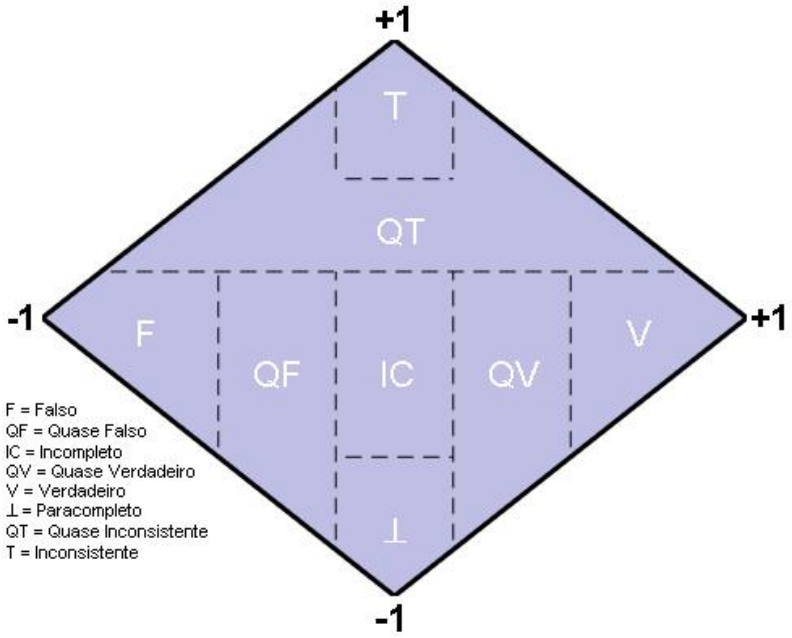


Figura 4.1 – O espaço paraconsistente representado pelo reticulado de Hasse, adaptado de [74]. As marcações (F, QF, IC, QV, V, ⊥, T) são relativas ao estado lógico paraconsistente (Tabela 4.1), em função do nível de certeza (eixo horizontal) e contradição (eixo vertical).

Quando $\eta_c = 1$, existe um nível de certeza máximo para a afirmação da proposição. Se $\eta_c = -1$, então existe um nível de certeza máximo para a negação da proposição. No que diz respeito ao nível de contradição, o valor $\eta_{ct} = 1$ representa uma contradição máxima positiva, enquanto o valor $\eta_{ct} = -1$uma contradição máxima negativa. Em uma interpretação paraconsistente, quanto mais a interpolação entre níveis de certeza e contradição estiver próxima ao $(1,0)$, maior a certeza de que a proposição é verdadeira, com o menor valor de contradição. É interessante notar que como os valores (η_c, η_{ct}) estão em função dos níveis de evidência favorável e desfavorável, pares ordenados do tipo $(\eta_c = 1, \eta_{ct} = 1)$, não ocorrem.

A Tabela 4.1 identifica a faixa de validade dos oito estados lógicos adotados nesse trabalho, adaptados de [74] e [57], resultantes de uma anotação paraconsistente e seus

níveis de certeza e contradição (η_c, η_{ct}). A interpretação associada aos estados lógicos é discutida posteriormente no texto.

Tabela 4.1 – Estados lógicos paraconsistentes. Adaptados de [74] e [57].

Estado lógico	η_c	η_{ct}
Falso (F)	$< -\frac{1}{2}$	$\geq -\frac{1}{2} e < \frac{1}{6}$
Quase Falso (QF)	$\geq -\frac{1}{2} e < -\frac{1}{6}$	$\geq -\frac{5}{6} e < \frac{1}{6}$
Incompleto (IC)	$\geq -\frac{1}{6} e < \frac{1}{6}$	$\geq -\frac{1}{2} e < \frac{1}{6}$
Quase Verdadeiro (QV)	$\geq \frac{1}{6} e < \frac{1}{2}$	$\geq -\frac{5}{6} e < \frac{1}{6}$
Verdadeiro (V)	$> \frac{1}{2}$	$\geq -\frac{1}{2} e < \frac{1}{6}$
Paracompleto (⊥)	$\geq -\frac{1}{6} e < \frac{1}{6}$	$< -\frac{1}{2}$
Inconsistente (T)	$\geq -\frac{1}{6} e < \frac{1}{6}$	$\geq \frac{1}{2}$
Quase Inconsistente (QT)	Outras delimitações	Outras delimitações

4.2.1 O nível de evidência real

Em uma análise paraconsistente para tomada de decisão, a avaliação de incertezas é feita considerando todas as informações evidenciais, sejam elas incompletas, contraditórias ou indefinidas. Para isso uma medida chamada nível de evidência real pode ser utilizada em conjunto com a análise do reticulado paraconsistente [70].

Para que o nível de evidência real seja estimado faz-se necessário o cálculo do nível de certeza real (η_{cr}), estimado por meio da distância entre níveis de certeza e contradição e os extremos (F, V) do reticulado, como na Equação (4.3).

$$\eta_{cr} = \begin{cases} 1 - \sqrt{(1 - |\eta_c|)^2 + \eta_{ct}^2} & se \eta_c > 0 \\ \sqrt{(1 - |\eta_c|)^2 + \eta_{ct}^2} - 1 & se \eta_c < 0 \\ 0 & se \eta_c = 0 \end{cases} \tag{4.3}$$

A partir de do nível de certeza real, o nível de evidência real (ω_{er}) é calculado por meio da Equação (4.4), sendo seu resultado a intensidade de certeza para uma dada proposição p.

$$\omega_{er} = \frac{\eta_{cr} + 1}{2} \tag{4.4}$$

4.3 MÁQUINA PARACONSISTENTE CORRELACIONAL (MPC)

A Máquina Paraconsistente Correlacional (MPC) integra DSNSFs e os conceitos da LPA2v para decidir sobre a proposição p: "A rede foi afetada por uma anomalia no intervalo t".

4.3.1 Representações

Seja R uma matriz representando as leituras reais do tráfego e X uma matriz representando uma DSNSF, ambas de ordem $\boldsymbol{T} \times \boldsymbol{F}$. Por convenção, R e X são relativas a um mesmo dia. Seja, $\boldsymbol{T} = 1440$ e $\boldsymbol{F} = 6$, sendo $\boldsymbol{T}$ o número total de intervalos e $\boldsymbol{F}$ o número total de atributos analisados. Dessa forma, cada linha r_t de R corresponde à leitura real do tráfego e cada linha x_t de X corresponde à previsão do DSNSF, para todos os seis atributos analisados, para um intervalo arbitrário t, $t \in \{1,2, \dots, \boldsymbol{T}\}$. A relação das colunas de R e X e seus atributos seguem a ordem: $(t, 1) \rightarrow$ "Bits", $(t, 2) \rightarrow$ "Pacotes", $(t, 3) \rightarrow$ "H(dstIP)", $(t, 4) \rightarrow$ "H(orgIP)", $(t, 5) \rightarrow$ "H(dstP)", $(t, 6) \rightarrow$ "H(orgP)". A seguir é apresentada a notação das matrizes R e X em sua forma expandida:

$$R = \begin{bmatrix} \overbrace{r_{1,1}}^{\text{Bits}} & \overbrace{r_{1,2}}^{\text{Pacotes}} & \overbrace{r_{1,3}}^{\text{H(dstIP)}} & \overbrace{r_{1,4}}^{\text{H(orgIP)}} & \overbrace{r_{1,5}}^{\text{H(dstP)}} & \overbrace{r_{1,6}}^{\text{H(orgP)}} \\ r_{2,1} & r_{2,2} & r_{2,3} & r_{2,4} & r_{2,5} & r_{2,6} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{1440,1} & r_{1440,2} & r_{1440,3} & r_{1440,4} & r_{1440,5} & r_{1440,6} \end{bmatrix};$$

$$X = \begin{bmatrix} \overbrace{x_{1,1}}^{\text{Bits}} & \overbrace{x_{1,2}}^{\text{Pacotes}} & \overbrace{x_{1,3}}^{\text{H(dstIP)}} & \overbrace{x_{1,4}}^{\text{H(orgIP)}} & \overbrace{x_{1,5}}^{\text{H(dstP)}} & \overbrace{x_{1,6}}^{\text{H(orgP)}} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} & x_{2,5} & x_{2,6} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1440,1} & x_{1440,2} & x_{1440,3} & x_{1440,4} & x_{1440,5} & x_{1440,6} \end{bmatrix}.$$

Cada tipo de anomalia pode afetar o comportamento normal de diferentes atributos de tráfego [36]. Esse trabalho avalia três tipos de anomalias, tipicamente encontradas em tráfego de rede agregado: *DoS*, *DDoS* e *Flash Crowd*. A relação entre essas anomalias e atributos afetados está baseada em trabalhos relacionados [36] e [75]. A Tabela 4.2 apresenta essa relação. Além do mais, é a partir dela que um vetor κ é construído.

Tabela 4.2 – Relação entre anomalias e atributos de tráfego afetados. Adaptado de [36] e [75].

Atributo	Anomalia		
	DoS	DDoS	Flash Crowd
Bits	-	-	✓
Pacotes	✓	✓	✓
H(dstIP)	✓	✓	✓
H(orgIP)	✓	-	-
H(dstP)	✓	✓	✓
H(orgP)	✓	✓	✓
	$\kappa_1 = \{0;1;1;1;1;1\}$	$\kappa_2 = \{0;1;1;0;1;1\}$	$\kappa_3 = \{1;1;1;0;1;1\}$

Seja uma determinada anomalia: para cada um dos seis elementos do vetor κ, cada qual representando um atributo de tráfego, o valor 1 é atribuído caso seu relativo atributo seja afetado ou 0 em caso negativo. Por exemplo, um DoS: não afeta o atributo "Bits", mas afeta "Pacotes", "H(dstIP)", "H(orgIP)", "H(dstP)" e "H(orgP)". Assim, levando em consideração a ordem em que os atributos são analisados na metodologia desse trabalho, o valor de κ para essa anomalia resulta em $\kappa = \{0;1;1;1;1;1\}$. Os três possíveis valores de κ são utilizados posteriormente na extração de informações evidenciais. Nesse trabalho a MPC foi limitada a investigação dos três tipos de anomalias descritas na, no entanto, diferentes atributos de tráfego podem ser explorados para que novos tipos de anomalias possam ser identificadas. Para isso, faz-se necessário a identificação desses atributos e a correlação entre eles.

4.3.2 Extração de Informações evidenciais

A proposta desse trabalho é avaliar minuto a minuto o tráfego agregado de rede em busca de possíveis anomalias. Dessa forma, informações evidenciais são extraídas com base nos valores do DSNSF e da leitura real de tráfego para um dado intervalo t.

Primeiramente, são calculadas medidas proporcionais entre o valor da leitura real de tráfego para um atributo e seu relativo correspondente na DSNSF. Esse processo é descrito na Equação (4.5). A função $\mathrm{men}(a;b)$ retorna o menor valor entre a e b e a operação $|a;b|$ reflete $|a;b| = \sqrt{(a-b)^2}$.

$$d_t = \left[\frac{|x_{t,1}; r_{t,1}|}{\mathrm{men}(x_{t,1}; r_{t,1})} \quad \frac{|x_{t,2}; r_{t,2}|}{\mathrm{men}(x_{t,2}; r_{t,2})} \quad \dots \quad \frac{|x_{t,2}; r_{t,2}|}{\mathrm{men}(x_{t,2}; r_{t,2})} \right] \tag{4.5}$$

Intuitivamente, a Equação 4.5 é utilizada para medir a discrepância entre o tráfego esperado e o tráfego real. A seguir, uma função de substituição, Equação (4.6), é utilizada para limitar os valores do vetor d_t em 1.

$$\varphi_{t,f} = \begin{cases} 1, & sed_{t,f} > 1 \\ d_{t,f}, & sed_{t,f} \leq 1 \end{cases} para f \in \{1,2,\dots,F\}. \tag{4.6}$$

Como mencionado anteriormente, a LPA2v utiliza níveis de evidência representados por valores reais, contidos no intervalo fechado $[0,1]$. Como d_t é utilizado na extração de tais níveis, sua substituição por φ_t pode ser interpretada como uma forma de assegurar valores contidos no intervalo $[0,1]$, que mantêm as menores discrepâncias entre leitura real e DSNSF com valores próximos de 0, enquanto as maiores com valores próximos ou iguais, mas não superiores a 1.

A MPC gera três saídas, uma para cada tipo de anomalia analisada, de acordo com a assinatura da anomalia de interesse representada por seu relativo κ. Esse fator é utilizado para obtenção do vetor e_t, como na Equação (4.7).

$$e_t = \kappa \varphi_t \tag{4.7}$$

Em seguida, o termo evidêncial (ψ_t) pode ser calculado por meio da média dos elementos de e_t, com relação a κ , como na Equação (4.8).

$$\psi_t = \frac{\sum_{j=1}^{F} e_{t,j}}{\sum_{j=1}^{F} \kappa_j} \tag{4.8}$$

Na metodologia adotada por esse trabalho, a análise de informação paraconsistente precisa de duas fontes de informação. Dessa forma, a DSNSF é considerada como fonte de informação. Como ela pode ser gerada com diversos modelos, assume-se:

$$X' \longrightarrow \text{DSNSF gerada com o modelo A;}$$

$$X'' \longrightarrow \text{DSNSF gerada com o modelo B.}$$

Assim, para cada linha t de X', x'_t, as Equações 4.6, 4.7 e 4.8 são utilizadas para gerar um fator evidêncial ψ'_t. Similarmente para cada linha t de X'', do qual é derivado um fator evidêncial ψ''_t. A partir de ψ'_t e ψ''_t os níveis de evidência favorável e desfavorável são gerados. Para o caso de ψ'_t, a Equação (4.9) é utilizada.

$$\mu = \psi'_t \tag{4.9}$$

Para ψ''_t, a Equação (4.10) é utilizada. A operação pode ser interpretada como uma forma de transformar níveis evidenciais positivos em negativos, em módulo.

$$\lambda = |\psi''_t - 1| \tag{4.10}$$

A transformação (4.10) é justificada uma vez que se uma fonte de informação afirma uma proposição com determinado nível de evidência favorável, ela nega a proposição com um nível inversamente proporcional a ele. A MPC se beneficia dessa característica e utiliza diferentes perspectivas para obter uma conclusão. Quando $X' = X''$ as interpolações (η_c, η_{ct}) seguem uma tendência linear. De outra forma, quando diferentes modelos são utilizados ($X' \neq X''$) os vieses obtidos por um único modelo são mitigados.

O Algoritmo 1 descreve os procedimentos para obtenção de níveis de evidência favoráveis e desfavoráveis, para cada anomalia avaliada, a partir dos valores estabelecidos por dois DSNSFs e leituras reais de tráfego. O mesmo é executado para cada intervalo t de um dia analisado.

Algoritmo 1 – Extração de informação evidêncial de assinaturas digitais

Entrada: Elementos de dois DSNSFs e leitura real de tráfego em um intervalo t: x'_t, x''_t, r_t.
Saída: Níveis de evidência favorável e desfavorável: μ, λ.

Use x'_t e r_t nas Equações (4.5, 4.6) e obtenha φ'_t.
Use x''_t e r_t nas Equações (4.5, 4.6) e obtenha φ''_t.

Para $i = 1$ **até** 3 **faça**
Use φ'_t e κ_i nas Equações (4.7,4.8) e obtenha ψ'_t.
Use ψ'_t na Equação (4.9) e obtenha μ. /* Nível de evidência favorável */
Use φ''_t e κ_i nas Equações (4.7,4.8) e obtenha ψ''_t.
Use ψ''_t na Equação (4.10) e obtenha λ. /* Nível de evidência desfavorável */
Use μ e λ na análise da LPA2v para anomalia i.
Fim

4.3.2.1 Exemplo numérico

Um exemplo numérico é apresentado para demonstração da extração de níveis evidênciais paraconsistentes a partir de valores hipotéticos gerados para um único intervalo.

Dado o primeiro intervalo de um dia, $t = 1$. Sejam dois DSNSFs gerados a partir de dois modelos hipotéticos A e B, respectivamente. Para $t = 1$, os valores numéricos desses dois DSNSFs são representados por dois vetores: $x'_{1,f}$ e $x''_{1,f}$, para $f \in \{1,2, \ldots, \boldsymbol{F}\}$. Para simplificação, os índices dos vetores ficam subentendidos, assim: $x' \rightarrow$ "Valores dos atributos de tráfego no intervalo$t = 1$do DSNSF gerado com modelo A"; e $x'' \rightarrow$ "Valores dos atributos de tráfego no intervalo$t = 1$ do DSNSF gerado com modelo B". A leitura real do tráfego, $r_{1,f}$, para $f \in \{1,2, \ldots, \boldsymbol{F}\}$, que subentende-se por: $r \rightarrow$ "Valores dos atributos da leitura real do tráfego no intervalo$t = 1$". Nesse exemplo, os valores para x', x'' e r, são os seguintes:

$$r = [2{,}9 \quad 2{,}7 \quad 3 \quad 3 \quad 4 \quad 4]$$
$$x' = [1{,}5 \quad 1{,}6 \quad 1{,}4 \quad 4{,}2 \quad 7{,}9 \quad 2]$$
$$x'' = [1{,}6 \quad 1{,}4 \quad 1{,}6 \quad 3{,}2 \quad 7{,}4 \quad 2{,}1]$$

A anomalia de interesse nesse exemplo é o *Flash crowd* com κ:

$$\kappa = [1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 1]$$

Primeiro, d' é calculado com r e x', como na Equação (4.5):

$$d' = \left[\frac{|x_{t,1}; r_{t,1}|}{\text{men}(x_{t,1}, r_{t,1})} \quad \frac{|x_{t,2}; r_{t,2}|}{\text{men}(x_{t,2}, r_{t,2})} \quad \ldots \quad \frac{|x_{t,2}; r_{t,2}|}{\text{men}(x_{t,2}, r_{t,2})}\right]$$

$$d' = \left[\frac{|1{,}5; 2{,}9|}{\text{men}(1{,}5; 2{,}9)} \quad \frac{|1{,}6; 2{,}7|}{\text{men}(1{,}6; 2{,}7)} \quad \frac{|1{,}4; 3|}{\text{men}(1{,}4; 3)} \quad \frac{|4{,}2; 3|}{\text{men}(4{,}2; 3)} \quad \frac{|7{,}9; 4|}{\text{men}(7{,}9; 4)} \quad \frac{|2; 4|}{\text{men}(2; 4)}\right]$$

$$d' = \left[\frac{1{,}4}{1{,}5} \quad \frac{1{,}1}{1{,}6} \quad \frac{1{,}6}{1{,}4} \quad \frac{1{,}2}{3} \quad \frac{3{,}9}{4} \quad \frac{2}{2}\right]$$

$$d' = [0{,}93 \quad 0{,}68 \quad 1{,}14 \quad 0{,}4 \quad 0{,}97 \quad 1]$$

Ao utilizar a substituição vista na Equação (4.6) tem-se:

$$\varphi' = [0{,}93 \quad 0{,}68 \quad 1 \quad 0{,}4 \quad 0{,}97 \quad 1]$$

Utilizando φ' e κ, como na Equação (4.7), obtém-se:

$$e' = [0{,}93 \times 1 \quad 0{,}68 \times 1 \quad 1 \times 1 \quad 0{,}4 \times 0 \quad 0{,}97 \times 1 \quad 1 \times]1$$

$$e' = [0{,}93 \quad 0{,}68 \quad 1 \quad 0 \quad 0{,}97 \quad 1]$$

O termo evidêncial ψ':

$$\psi' = \frac{0{,}93 + 0{,}68 + 1 + 0 + 0{,}97 + 1}{1 + 1 + 1 + 0 + 1 + 1} = \frac{4{,}59}{5}$$

$$\psi' = 0{,}91$$

Finalmente, utilizando a Equação (4.9) para obter o nível de evidência favorável $\mu = \psi' = 0{,}91$.

Os procedimentos utilizados na obtenção do nível de evidência desfavorável são similares. Primeiramente, d'' é calculado com r e x''.

$$d'' = \left[\frac{|x_{t,1}; r_{t,1}|}{\text{men}(x_{t,1}, r_{t,1})} \quad \frac{|x_{t,2}; r_{t,2}|}{\text{men}(x_{t,2}, r_{t,2})} \quad \ldots \quad \frac{|x_{t,2}; r_{t,2}|}{\text{men}(x_{t,2}, r_{t,2})}\right]$$

$$d'' = \left[\frac{|1{,}6;\ 2{,}9|}{\text{men}(1{,}6;\ 2{,}9)} \quad \frac{|1{,}4;\ 2{,}7|}{\text{men}(1{,}4;\ 2{,}7)} \quad \frac{|1{,}6;\ 3|}{\text{men}(1{,}6;\ 3)} \quad \frac{|3{,}2;\ 3|}{\text{men}(3{,}2;\ 3)} \quad \frac{|7{,}4;\ 4|}{\text{men}(7{,}4;\ 4)} \quad \frac{|2{,}1;\ 4|}{\text{men}(2{,}1;\ 4)}\right]$$

$$d'' = \left[\frac{1{,}3}{1{,}6} \quad \frac{1{,}3}{1{,}4} \quad \frac{1{,}4}{1{,}6} \quad \frac{0{,}2}{3} \quad \frac{3{,}4}{4} \quad \frac{1{,}9}{2{,}1}\right]$$

$$d'' = [0{,}81 \quad 0{,}92 \quad 0{,}87 \quad 0{,}06 \quad 0{,}85 \quad 0{,}90]$$

Com a substituição da Equação (4.6):

$$\varphi'' = [0{,}81 \quad 0{,}92 \quad 0{,}87 \quad 0{,}06 \quad 0{,}85 \quad 0{,}90]$$

Utilizando φ'' e κ na Equação (4.7), obtém-se:

$$e'' = [0{,}81 \times 1 \quad 0{,}92 \times 1 \quad 0{,}87 \times 1 \quad 0{,}06 \times 0 \quad 0{,}85 \times 1 \quad 0{,}90 \times]1$$

$$e'' = [0{,}81 \quad 0{,}92 \quad 0{,}87 \quad 0 \quad 0{,}85 \quad 0{,}90]$$

O termo evidêncial ψ'':

$$\psi'' = \frac{0{,}81 + 0{,}92 + 0{,}87 + 0 + 0{,}85 + 0{,}90}{1 + 1 + 1 + 0 + 1 + 1} = \frac{4{,}36}{5} = 0{,}87$$

$$\psi'' = 0{,}87$$

Utilizando a Equação (4.9) para obter o nível de evidência favorável: $\lambda = |\psi'' - 1| = 0{,}13$

Considerando os cálculos anteriores, as informações evidenciais das fontes de informações foram extraídas e resumidas em $\mu = 0{,}91$ e $\lambda = 0{,}13$, tornando a análise paraconsistente possível. Primeiramente o nível de certeza pode ser calculado utilizando a Equação (4.1):

$$\eta_c = \mu - \lambda = 0{,}91 - 0{,}13 = 0{,}78$$

Bem como o nível de contradição, Equação (4.2):

$$\eta_{ct} = \mu + \lambda - 1 = 0{,}91 + 0{,}13 - 1 = 0{,}04$$

O ponto (0.78,0.04) é marcado no reticulado. Seguindo as constantes de anotação da Tabela 4.1: $\eta_c > \frac{1}{2}$ e $-\frac{1}{2} \leq \eta_{ct} < \frac{1}{6}$, que representa o estado lógico Verdade (V). A Figura 4.2 representa a marcação do ponto ($n_c = 0.78, n_{ct} = 0.04$).

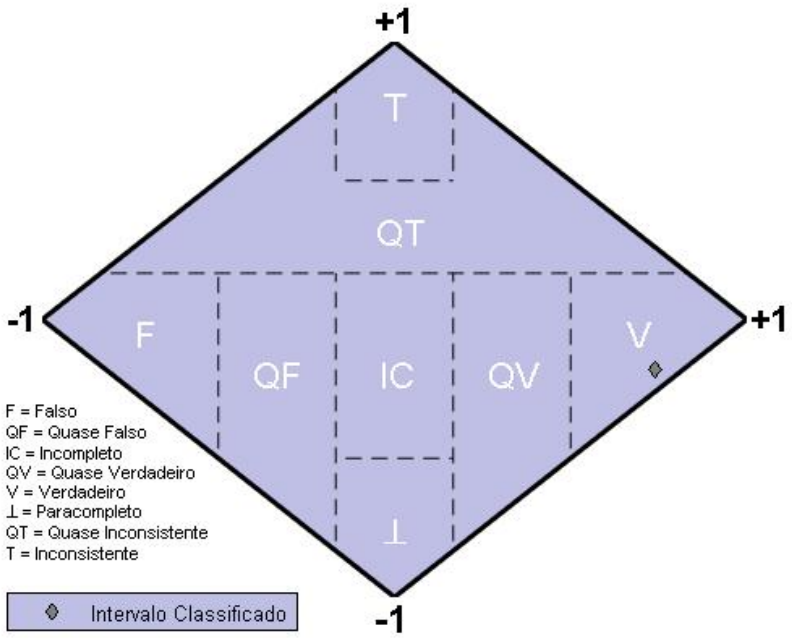


Figura 4.2 – Interpolação do ponto (0,78,0,04) no reticulado.

Adiante, o Nível de evidência real é calculado. Seguindo a Equação (4.3), e como $\eta_c > 0$, o nível de certeza real é calculado como a seguir:

$$\eta_{cr} = 1 - \sqrt{(1 - |\eta_c|)^2 + \eta_{ct}^2}$$

$$\eta_{cr} = 1 - \sqrt{(1 - |0.78|)^2 + 0.04^2}$$

$$\eta_{cr} = 1 - \sqrt{0.05}$$

$$\eta_{cr} = 1 - 0.2236$$

$$\eta_{cr} = 0{,}7764$$

Finalmente o Nível de evidência real (ω_{er}):

$$\omega_{er} = \frac{\eta_{cr} + 1}{2}$$
$$\omega_{er} = \frac{0.7764 + 1}{2}$$
$$\omega_{er} = 0{,}8882$$

O exemplo ilustrado apresenta um intervalo classificado como Verdadeiro e com elevado nível de evidência real. Na prática, cada intervalo pode ser classificado em qualquer um dos 8 estados lógicos.

4.3.2.2 Interpretação do reticulado

Os intervalos alocados no estado Verdadeiro (V) são considerados como anômalos. Isso acontece quando os níveis de evidência favoráveis estão próximos de um e os níveis de evidência desfavoráveis estão próximos de zero. Nessas situações as evidências vindas dos modelos empregados na geração dos DSNSFs ilustram distâncias consideráveis entre os perfis de atributo dos DSNSFs e as leituras reais de tráfego. A operação modular feita sob a segunda de fonte de informação permite que o nível de contradição seja o mais baixo possível quando os atributos não se comportam da forma esperada.

Na prática os intervalos tendem a ser alocados nos estados Falso, Quase Falso e Incompleto. Os estados F e QF representam uma situação em que os níveis de evidências favoráveis estão próximos de zero e os níveis de evidências desfavoráveis próximos de um, o que resulta nos menores valores para nível de certeza com os menores valores para o nível de contradição, ou seja, comportamentos normais. O estado Incompleto representa uma situação em que não existe informação ou evidência suficiente sobre uma determinada proposição, assim, uma anomalia não deve ser notificada. Isso evita que a MPC gere falsos alarmes quando há falta de informação.

Os estados Inconsistente e Quase Inconsistente referem-se a intervalos analisados com os maiores níveis de evidências favoráveis, ao mesmo tempo em que estão os maiores níveis de evidências desfavoráveis, resultando em níveis de contradição próximos a um e nível de certeza próximos de zero. Essa situação reflete um desacordo entre os modelos ou fontes de informações utilizadas. O primeiro modelo interpreta o intervalo como anômalo enquanto o segundo como normal. Esse é o motivo pelo qual pontos alocados nesse espaço não são considerados como anomalias.

O estado paracompleto representa uma situação em que não existe nem informação favorável nem informação desfavorável sobre a presença de anomalias, também representando um intervalo que não dever ser considerado como anomalia.

4.3.3 Funcionamento da MPC

Como os aspectos teóricos envolvidos na abordagem proposta já foram apresentados anteriormente, descreve-se o funcionamento do sistema baseado na Máquina Paraconsistente Correlacional de uma forma descritiva e prática.

Um coletor de informações é responsável por armazenar as informações no formato de exportação suportado pelo equipamento de rede ao longo do dia. Esse armazenamento é feito continuamente até que uma base de informações contenha a quantidade de informações suficiente para que os primeiros DSNSFs sejam construídos. O processo de monitoramento e armazenamento não para, e novas entradas de dados são utilizadas nos cálculos futuros. O armazenamento pode ser feito em um PC comum, servidor, ou qualquer outro dispositivo que forneça uma interface de acesso.

Com a base histórica pronta o sistema pode funcionar como o detector de anomalias proposto. O treinamento dos modelos ARIMA é feito utilizando-se quatro dias das quatro semanas precedentes ao dia a ser investigado. A assinatura de uma segunda-feira será baseada nas informações encontradas nas quatro segundas-feiras das semanas anteriores. A assinatura de uma terça-feira será baseada nas informações encontradas nas quatro terças-feiras das semanas anteriores, e assim por diante.

Para cada instante t o modelo ARIMA se baseia nas medidas estatísticas do ruído branco advindo da base de dados e das informações encontradas nos instantes $t-1, t-2, t-k$ do dia atual analisado, onde k é a defasagem estimada pelo modelo. Como foi proposto, dois modelos são utilizados, baseado na média e baseado na mediana histórica. Assim, duas previsões são feitas para o instante t analisado. Quando a leitura de tráfego real é obtida pelo coletor uma cópia dessa informação é armazenada na base de dados para posterior uso na construção das assinaturas e outra cópia dessa informação é utilizada em conjunto com as previsões para extração de informação evidencial.

Previsões e leituras reais são relacionadas e os níveis de evidência favorável e desfavorável são obtidos, para os três tipos de anomalias estudadas. Os mesmos são formalizados por meio da LPA2v e os níveis de certeza, contradição e de evidência real são obtidos. Esses valores são o suficiente para a tomada de decisão do administrador. Quando a

interpolação dos níveis de certeza e contradição estão localizados no estado lógico Verdadeiro uma anomalia é sinalizada. O nível de evidência real é uma métrica que reúne em um só valor os padrões evidenciais do comportamento normal da rede e demonstra as tendências presentes no tráfego relativas a comportamentos anômalos. Na prática, os intervalos classificados como Verdadeiro produzem níveis de evidência real elevados.

4.3.4 Complexidade Computacional

A complexidade do sistema está relacionada as diferentes fases de processamento empregadas. O cálculo da entropia de Shannon é de ordem linear, uma vez que refere-se a um somatório variando de 1 até N onde N é o número de amostras no intervalo. Como o cálculo é feito para cada intervalo, esse processo é repetido **T** vezes, para cada um dos 4 atributos que requerem essa transformação. Após a transformação a análise de series temporais inicia-se a modelagem ARIMA. O sistema apresentado processa seis atributos diferentes a serem tratados como séries temporais, assim, os cálculos executados pelo processo de caracterização de tráfego devem ser executados seis vezes. A complexidade do modelo ARIMA é de ordem quadrática para estimação e linear para previsão. Por fim, a complexidade da extração de informação evidencial e cálculo de medidas evidenciais é de ordem linear.

Na prática, o tempo de execução das aplicações envolvidas na classificação de um dia completo (incluindo treinamento de modelos ARIMA, extração evidencial, classificação e *plots* de gráficos) obteve média de 7.8 segundos. Os experimentos foram feitos com um computador pessoal comum e assumindo a base de informações e leituras reais prontas para análise.

5 TESTES E RESULTADOS

Este capítulo descreve os ambientes de rede e monitoramento e métricas de avaliação utilizadas, além de apresentar e discutir os testes e resultados da caracterização de tráfego e detecção de anomalias.

5.1 AMBIENTES DE REDE E MONITORAMENTO

Os testes foram conduzidos em duas redes diferentes, avaliando o comportamento da solução proposta em diferentes ambientes.

A primeira rede foi a da Universidade Estadual de Londrina (UEL), que contém aproximadamente sete mil dispositivos. Os dados foram coletados em um *switch* BD8801 Extreme, localizado no *core* da rede, o qual agrega todo o tráfego da Universidade. Devido à grande quantidade de informação, as informações coletadas foram amostradas a uma taxa de 1 para cada 256 pacotes. O formato *sFlow* foi utilizado na exportação, que se deu com o uso do aplicativo analisador de trafego *Softflowd*.

A segunda rede testada é a da Universidade Tecnológica Federal do Paraná (UTFPR) – Campus Toledo. O tráfego agregado no *gateway* principal da rede é monitorado e exportado com o padrão Netflow v9. Por se tratar de uma rede de menor porte, nenhuma técnica de amostragem foi utilizada.

5.2 MÉTRICAS DE AVALIAÇÃO

As métricas descritas a seguir são utilizadas para avaliar a eficiência do sistema proposto na caracterização de tráfego bem como na detecção de anomalias. Como a caracterização de tráfego é uma etapa essencial na metodologia adotada para detecção de anomalias, é importante que seus resultados sejam consistentes. Com relação à detecção de anomalias, é crucial que o sistema alerte os administradores e gerentes sobre anomalias apenas em situações que mereçam atenção.

Dessa forma, os coeficientes de correlação e o *Erro Quadrático Médio Normalizado* (NMSE) são utilizados para avaliar a caracterização de tráfego enquanto a detecção de anomalias é avaliada por meio da acurácia e das curvas ROC (*Receiver Operating Characteristics*).

5.2.1 Coeficiente de correlação

A primeira métrica utilizada é denominada coeficiente de correlação (cc). Ela mede o grau de relação linear entre duas variáveis quantitativas. Nos ambientes avaliados, essas variáveis dizem respeito aos valores gerados pelas assinaturas digitais e às leituras reais monitoradas. Os valores obtidos estão dentro do intervalo de -1 a +1, em que 1 indica correlação total, 0 indica correlação nula e -1 especifica uma correlação inversa. Na prática, valores próximos de 1 indicam uma correlação perfeita positiva, ou um ótimo nível de concordância entre o DSNSF e as leituras reais da rede. A Equação (5.1) é utilizada na estimativa desses coeficientes para duas séries temporais, onde X é o observado e Y é o previsto. Como definido anteriormente, $\boldsymbol{T} = 1440$, ou seja, o tamanho das séries temporais utilizadas ao longo desse trabalho.

$$cc = \frac{\sum_{i=1}^{\boldsymbol{T}}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{\boldsymbol{T}}(X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^{\boldsymbol{T}}(Y_i - \bar{Y})^2}} \tag{5.1}$$

5.2.2 Erro quadrático médio normalizado (NMSE)

A segunda métrica utilizada foi o *Erro Quadrático Médio Normalizado* (*Normalized Mean Square Error* - NMSE). Ela avalia a diferença absoluta entre o que foi predito pelo modelo e o que é apresentado nas leituras do tráfego real. Valores elevados indicam divergências entre o DSNSF e o movimento do tráfego, enquanto valores próximos de zero indicam excelente caracterização do tráfego. A Equação (5.2) é utilizada para calcular o NMSE.

$$NMSE = \frac{\sum_{i=1}^{\boldsymbol{T}}(X_i - Y_i)^2}{\sum_{i=1}^{\boldsymbol{T}}(Y_i)^2} \tag{5.2}$$

5.2.3 Acurácia

A medida de acurácia é um nível percentual dos intervalos classificados corretamente pelo sistema. Ela é calculada com base na quantidade de Verdadeiros Positivos (VP) e Verdadeiros Negativos (VN).

Os Verdadeiros Positivos (VP) são referentes aos intervalos realmente anômalos reconhecidos pelo sistema. Por sua vez, os Verdadeiros Negativos (VN) são

referentes aos intervalos normais, também classificados pelo sistema como normais. Nessa mesma linha de raciocínio, a noção de Falso Positivo e Falso Negativo pode ser descrita.

Um Falso Positivo (FP) é um intervalo classificado pelo sistema como anômalo, mas que na verdade é normal. Já um Falso Negativo (FN) é um intervalo realmente anômalo, que não foi identificado pelo sistema.

Um sistema tem boa acurácia quando reconhece a maior quantidade de intervalos anômalos reais e a maior quantidade de intervalos normais. A Equação (5.3) demonstra como a acurácia é calculada.

$$Acurácia = \frac{VP + VN}{IntervalosNormais \ + IntervalosAnômalos} \tag{5.3}$$

5.2.4 Curvas ROC (*Receiver Operating Characteristic*)

As curvas ROC são utilizadas nesse trabalho como auxílio visual na interpretação das Taxas de Verdadeiros Positivos (TVP) e das Taxas de Falsos Positivos (TFP). Elas são representadas por um gráfico cujo eixo horizontal (ou das abscissas) ilustra o valor percentual das Taxas de Falsos Positivos enquanto o eixo vertical (ou das ordenadas) representa o das Taxas de Verdadeiros Positivos. A Equação (5.4) é utilizada para calcular a Taxa de Verdadeiros Positivos.

$$TVP = \frac{VP}{IntervalosAnômalos} \tag{5.4}$$

Por sua vez, a Equação (5.5) é utilizada no cálculo da Taxa de Falsos Positivos (TFP).

$$TFP = \frac{FP}{Intervalosnormais} \tag{5.5}$$

O resultado do contraste entre TVP e TFP é uma curva. Quanto mais rápido ela convergir para 100% de TVP, melhores são os resultados obtidos pelo sistema de classificação.

5.3 CENÁRIOS DE TESTE

Algumas modificações nas configurações do sistema proposto foram feitas para que parâmetros de comparação fossem obtidos. Além das anomalias inerentes ao tráfego real, alguns dias foram selecionados para que anomalias fossem neles injetadas com auxílio de uma ferramenta para simulação de anomalias.

A ferramenta *Scorpius* [76], desenvolvida pelo grupo *Rigel* de pesquisa em redes do Departamento de Computação da Universidade Estadual de Londrina, foi desenvolvida com o objetivo de auxiliar o processo de análise e validação de desempenho de sistemas de detecção de anomalias baseados na análise de fluxos IP. Arquivo de fluxos reais são utilizados como entrada da ferramenta. Ela injeta fluxos anômalos com as características requeridas, em um intervalo de tempo, e produz um arquivo de saída de fluxos IP com as anomalias incorporadas. Este arquivo pode ser utilizado para os testes da mesma forma que o original.

Por meio de inspeção manual exaustiva dos fluxos IP dos ambientes de rede estudados, um grupo de administradores de rede sinalizou um gabarito, o qual é utilizado na avaliação da eficiência do sistema na detecção de anomalias. A sinalização é feita logo após a injeção de anomalias ser realizada pela ferramenta *Scorpius*.

A metodologia proposta nesse trabalho gera dois tipos de DSNSFs para cada dia da semana, cada qual relativo aos seis atributos de tráfego analisados. O primeiro tipo de DSNSF é referenciado como M1 enquanto o segundo tipo como M2. Essa diferenciação entre os tipos de DSNSF está ligada a qual modelo foi utilizado para o cálculo dos perfis de comportamento.

Este trabalho apresentou uma descrição do modelo ARIMA e duas maneiras de implementação: com base na média ou na mediana histórica de dados para geração do ruído branco. Os DSNSFs baseados nessas duas medidas estatísticas servem como as fontes de informações para Máquina Paraconsistente Correlacional (MPC) tradicional, que extrai informações evidenciais sobre o estado da rede em um determinado intervalo e o classifica de acordo com seu estado lógico paraconsistente.

A detecção de anomalias baseada na MPC é comparada com uma abordagem tradicional encontrada na literatura baseada em limites de confiança ou *thresholds* [77]. Os limites de confiança indicam variações ou desvios entre previsão, e observação que são considerados normais. Os procedimentos utilizados no cálculo desses limites são descritos no Apêndice C.

O processo de comparação é feito com base nos três tipos de anomalias analisadas pela MPC. As mesmas características do comportamento das anomalias avaliadas são utilizadas (Tabela 4.2) e um evento só é classificado como anômalo quando existe alteração (movimento excede os limites dos *thresholds*) nos atributos referentes a essas características. Ao final de uma intersecção entre as três análises (*DoS, DDoS* e *Flash Crowd*) o intervalo é classificado como "Normal" ou "Anômalo".

A metaheurística *Ant Colony Optimization for Digital Signature* (ACODS) [63] é utilizada para comparação com a MPC. O ACODS também é um modelo para geração de DSNSFs. O objetivo é comparar duas classes de algoritmos diferentes, tanto em modo competitivo como em modo cooperativo. Os detalhes do ACODS são descritos no Apêndice B.

O sistema foi avaliado em quatro cenários diferentes. Os três primeiros cenários foram configurados no ambiente de rede da UEL e o último no da UTFPR. A Tabela 5.1 descreve os cenários testados.

Tabela 5.1 – Cenários de teste.

	Experimento	**Rede**	**Período**
Cenário 1	Caracterização de tráfego com os modelos ARIMA e detecção de anomalias com a MPC.	UEL	Treinamento: 24/09/2012 a 19/10/2012 e 01/04/2013 a 26/04/2013
			Testes: 22/10/2012 a 09/11/2012 e 29/04/2013 a 17/05/2013
Cenário 2	Comparação entre MPC e ACODS.	UEL	Treinamento: 24/09/2012 a 19/10/2012
			Testes: 22/10/2012 a 09/11/2012
Cenário 3	Avaliação da MPC baseada na cooperação entre ARIMA e ACODS.	UEL	Treinamento: 24/09/2012 a 19/10/2012
			Testes: 22/10/2012 a 09/11/2012
Cenário 4	Caracterização de tráfego com o modelo ARIMA e detecção de anomalias com a MPC.	UTFPR	Treinamento: 01/04/2013 a 26/04/2013
			Testes: 29/04/2013 a 17/05/2013

5.3.1 Cenário 1 – Caracterização de tráfego e detecção de anomalias (UEL)

O primeiro cenário utiliza o ambiente de rede da Universidade Estadual de Londrina, modelos ARIMA (com ruído branco baseado na média e mediana), e Máquina Paraconsistente Correlacional para avaliar os seguintes aspectos:

- Efetividade dos modelos ARIMA na caracterização de tráfego;
- Efetividade da MPC baseada nos modelos ARIMA na detecção de anomalias;
- Comparação entre abordagem baseada na MPC e abordagem baseada em limites de confiança.

Esse cenário foi construído com registros de fluxos IP de quatorze semanas da rede UEL. A caracterização de tráfego utilizou os seguintes períodos: entre 1 e 26 de abril de 2013 para treinamento; e entre 29 de abril e 16 de maio de 2013 para testes. Por sua vez, a detecção de anomalias utilizou os períodos entre 24 de setembro e 19 de outubro de 2012 para treinamento e entre 22 de outubro e 9 de novembro de 2012 para testes.

A ferramenta *Scorpius* foi utilizada nos dias 5 e 6 de novembro de 2012 para inclusão de anomalias sintéticas. Um ataque *DoS* foi injetado utilizando endereços fictícios, um endereço IP de origem e um endereço IP de destino, com a porta de origem de número 80 e de destino com número 8081. Um ataque *DDoS* foi injetado utilizando um conjunto de endereços IP de origem fictícios, todos com porta número 80 e um de endereço IP de destino fictício, com porta número 8081. A Tabela 5.2 apresenta as informações dos ataques simulados.

Tabela 5.2 – Informações sobre a simulação de anomalias (UEL).

Dia	Anomalia	Intervalo de Tempo	IP de origem	Porta de origem	IP de destino	Porta de destino
05/11/2012	*DoS*	12h00 até 14h00	11.11.11.11	80	12.12.12.12	8081
06/11/2012	*DDoS*	17h00 até 19h00	500 endereços gerados aleatoriamente	80	12.12.12.12	8081

Os modelos utilizados foram: M1 → “ARIMA baseado em ruído branco gerado com a média de 4 semanas de dados históricos” e M2 → “ARIMA baseado em ruído branco gerado com a mediana de 4 semanas de dados históricos”.

5.3.1.1 Caracterização de tráfego

O conjunto de Figuras de 5.1 a 5.7 ilustram o comportamento das leituras reais de tráfego e as especificações geradas pelas DSNSFs entre os dias 22 e 30 de outubro de 2012 no ambiente de rede da Universidade Estadual de Londrina. O volume das leituras reais coletado na rede da Universidade aparece em áreas verde-claro limitado por linhas pontilhadas na cor preta. As linhas sólidas representam as DSNSFs geradas com M1 (cor vermelha) e M2 (cor preta).

As mudanças na intensidade dos atributos são visualmente mais evidentes no período de funcionamento da instituição, entre 8 e 23 horas. No geral, é possível identificar o ajustamento entre leituras reais e perfis encontrados nas DSNSFs, um indício da eficiência dos modelos proposto na caracterização de comportamento normal.

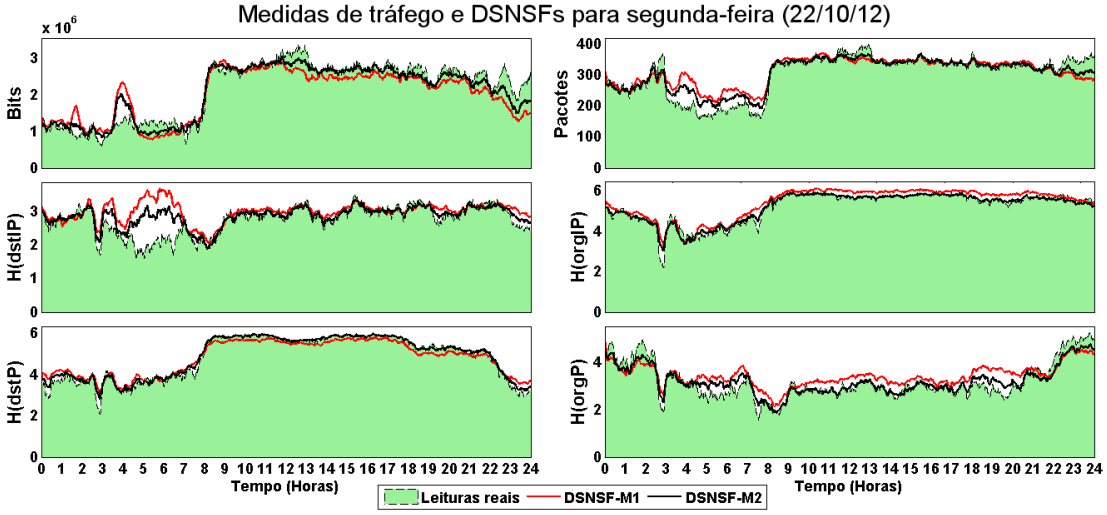


Figura 5.1 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 22/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

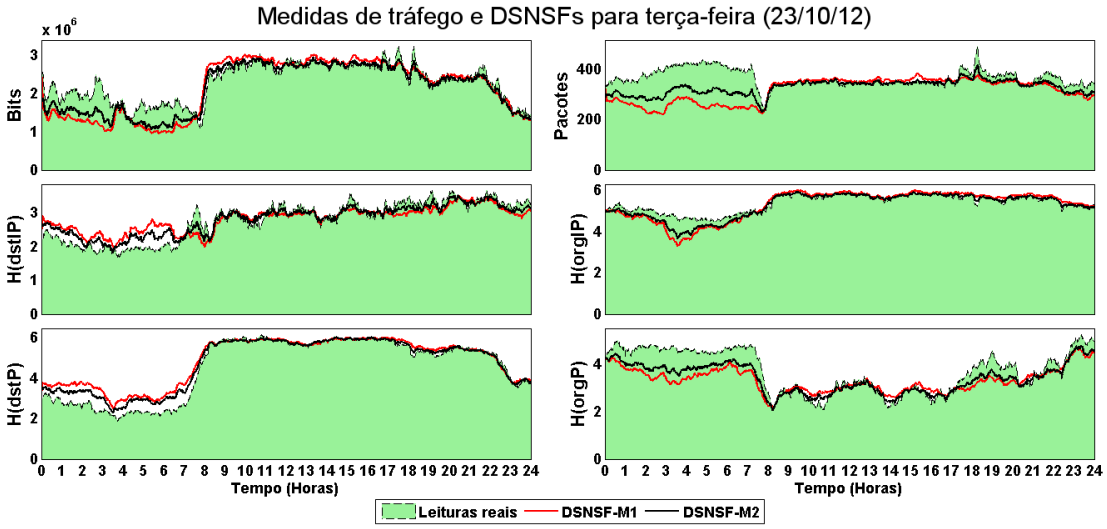


Figura 5.2 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 23/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

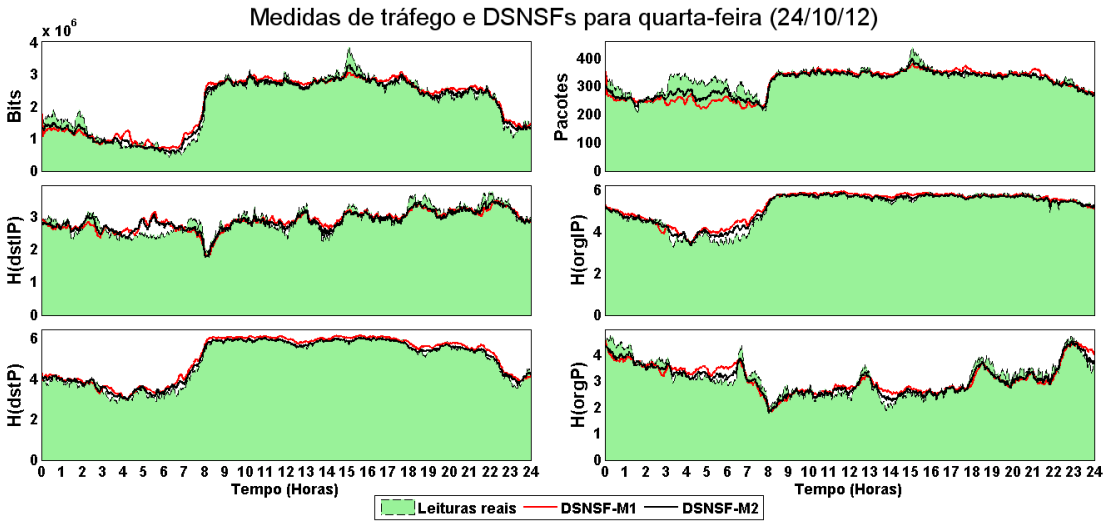


Figura 5.3 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 24/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

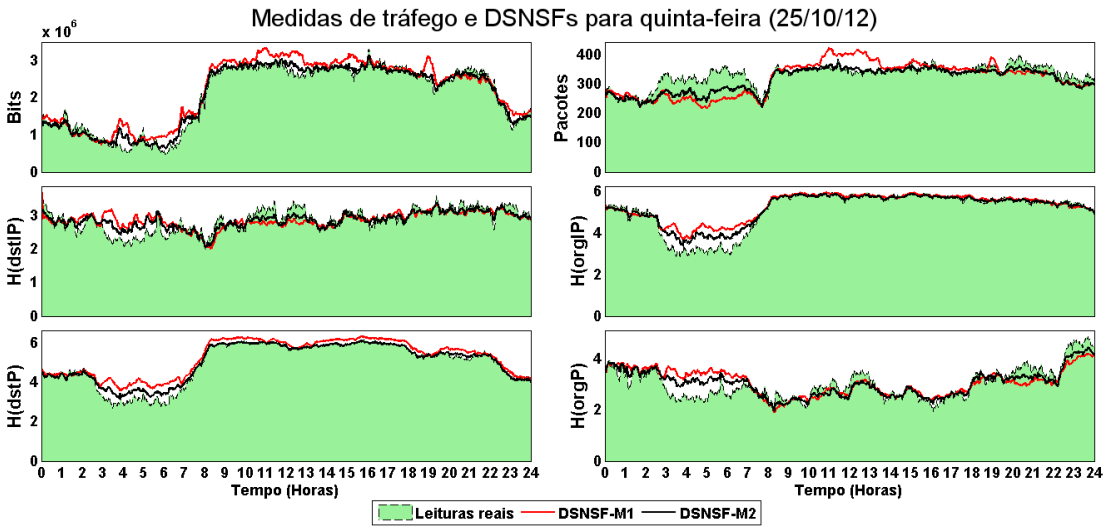


Figura 5.4 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 25/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

É possível detectar que no período da madrugada, da 0 as 7 horas, as variações são mais acentuadas. Isso ocorre por causa da falta de movimentação na rede analisada nesses horários. Isso, aliado a atividades como *backup,* aumentam a possibilidade dos modelos sub ou supra estimarem os limites para o movimento real.

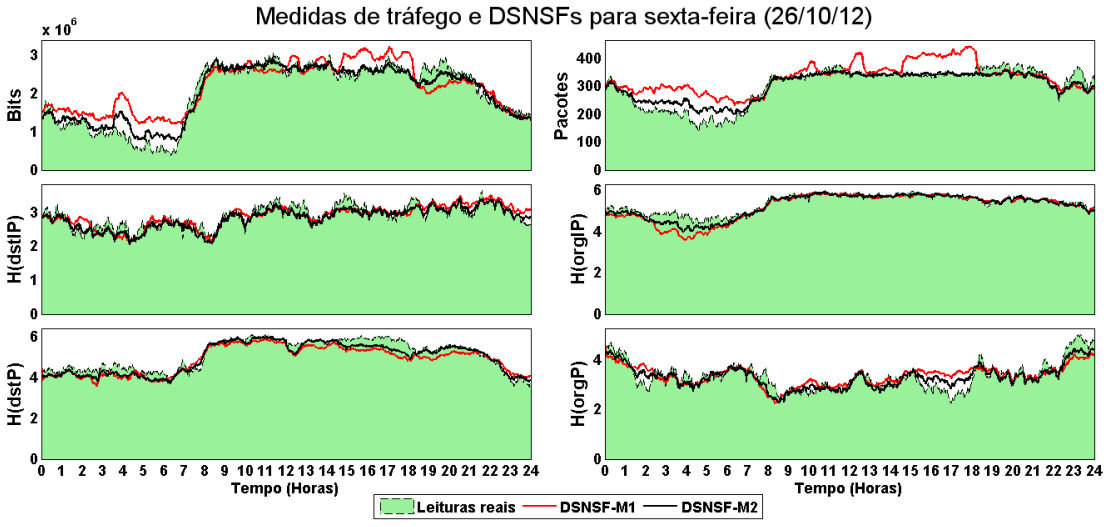


Figura 5.5 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 26/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

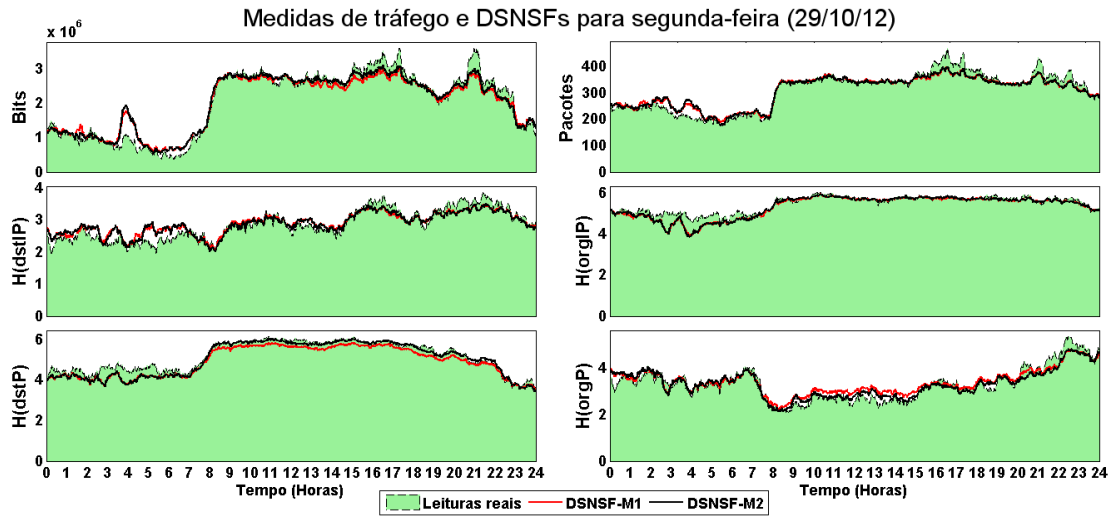


Figura 5.6 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 29/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

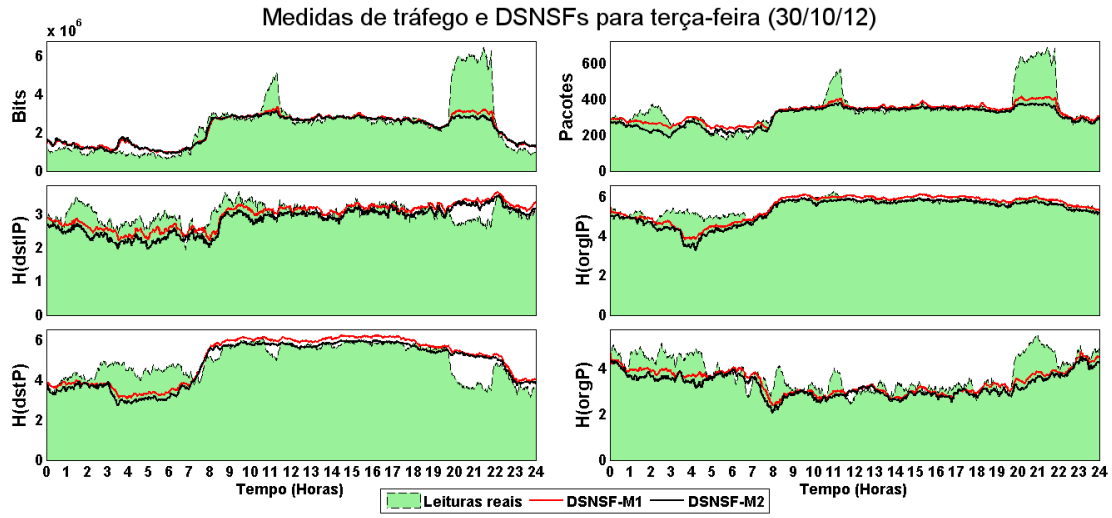


Figura 5.7 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 30/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

O conjunto de Figuras 5.8, 5.9, 5.10 e 5.11 investigam numericamente a qualidade dos modelos ARIMA na caracterização de tráfego. A Figura 5.8 apresenta os resultados do cálculo do coeficiente correlação entre as DSNSFs geradas com os dois modelos e as leituras reais do tráfego do período de 29 de abril a 17 de maio de 2013. Cada dia apresenta a média aritmética do coeficiente de correlação entre os seis atributos analisados. As barras na

cor preta ilustram os resultados do modelo M1, já as barras na cor branca ilustram os resultados do modelo M2.

Os dois modelos obtiveram bons resultados, com média geral acima de 0.9. Os piores coeficientes foram resultado de um feriado nacional (1 de maio), aproximadamente 0.5 e 0.65 para M1 e M2, respectivamente. Apesar disso, os demais resultados indicam forte correlação entre o previsto e o observado. Nota-se também que o modelo M2 apresentou resultados melhores que o modelo M1.

A Figura 5.9 ilustra os resultados do coeficiente de correlação obtidos com os modelos em função de cada atributo da DSNSF. Esse teste foi realizado com o objetivo de investigar a correlação entre cada atributo individualmente, identificando quais são mais afetados, e quais apresentam maior estabilidade. As barras na cor preta ilustram os resultados do modelo M1, sendo a média de correlação de cada atributo dos dias referentes às 3 semanas investigadas na caracterização. Semelhantemente, as barras na cor branca indicam os resultados do modelo M2. Nesse teste, o atributo entropia de porta de origem, H(orgP), foi o que obteve os menores valores, enquanto a entropia de endereço IP de origem, H(orgIP), atingiu o maiores valores. Apesar disso, todos os atributos obtiveram resultados iguais ou superiores a 0.8, o que corresponde a forte correlação.

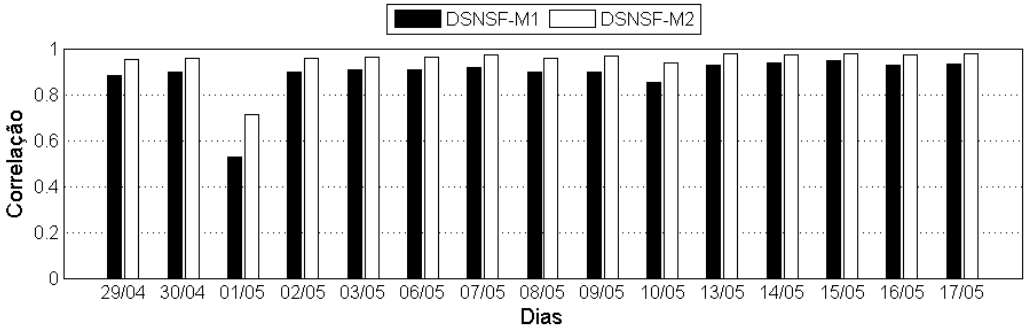


Figura 5.8 - Média dos coeficientes de correlação, calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 29 de abril a 17 de maio de 2013.

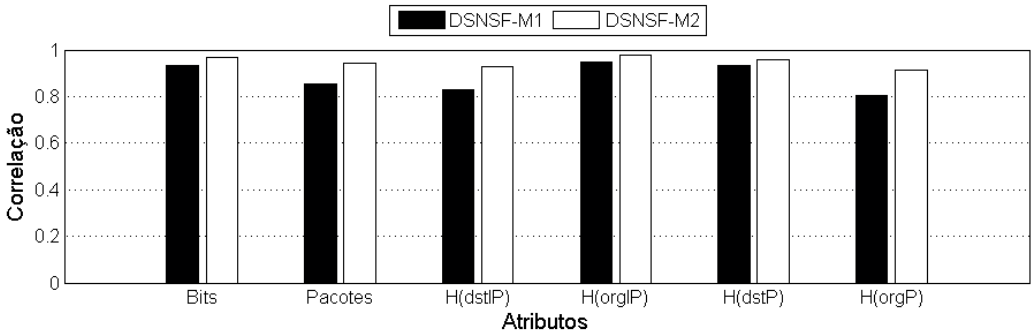


Figura 5.9 - Coeficientes de correlação em relação a cada atributo analisado para o período de 29 de abril a 17 de maio de 2013.

Os testes conduzidos com a métrica NMSE são similares aos conduzidos com o coeficiente de correlação. A Figura 5.10 apresenta os resultados do cálculo dos erros NMSE, entre as DSNSFs de cada modelo e as leituras do tráfego reais. A média aritmética do erro dos seis atributos para cada dia analisado são apresentadas. Novamente, o modelos propostos obtiveram excelentes resultados, com valores inferiores a 0.02 para M1 e inferiores a 0.01 para M2, na maior parte do tempo. Quanto menor o erro, mais adequado o modelo de previsão. Novamente o dia mais afetado foi 1 de maio, feriado nacional, onde o movimento da Universidade caiu consideravelmente. Como os modelos utilizam uma base de dados históricos no processo de previsão de valores, a tendência de alto movimento de tráfego das semanas precedentes levam as expectativas de tráfego serem condizentes para dias de atividade. Como não está no escopo desse trabalho, a investigação de modelos específicos para fins de semana e feriados será feita em trabalhos futuros.

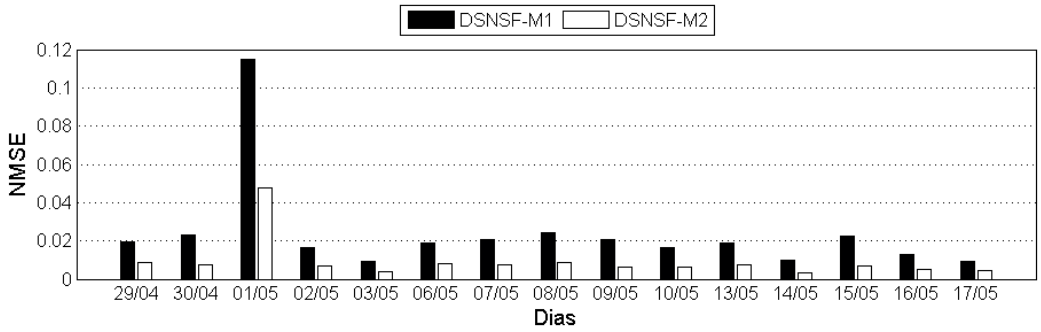


Figura 5.10 - Média dos erros NMSE calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 29 de abril a 17 de maio de 2013.

Os resultados da métrica NMSE em relação a cada atributo são ilustrados na Figura 5.11. Os atributos bits, pacotes e entropia de porta de origem, H(orgP), obtiveram os maiores erros. Ainda assim, as médias se mantiveram um nível muito bom para previsão de valores.

Nota-se que o modelo M2, baseado na mediana, apresentou resultados superiores aos de M1, com base na média, em todos os testes. Alguns fatores podem contribuir para tanto. A métrica estatística média leva em consideração todos os valores de um conjunto de dados, até mesmo aqueles com maior discrepância. O mesmo não ocorre com a mediana, baseada nos valores centrais de um conjunto de dados.

Mesmo com aproximação dos parâmetros ϕ e θ sendo feita para que o modelo ARIMA construído minimize a distância entre previsões e observações, valores muito pequenos ou muito altos no conjunto de dados de treinamento podem causar uma sobre ou superestimação dos perfis de cada atributo.

Ainda que os resultados de M1 sejam inferiores, tal modelo não deve ser descartados. Ele apresenta um perfil baseado em outro comportamento do conjunto histórico de dados. Essa é uma característica a ser aproveitada na detecção de anomalias baseada na lógica paraconsistente.

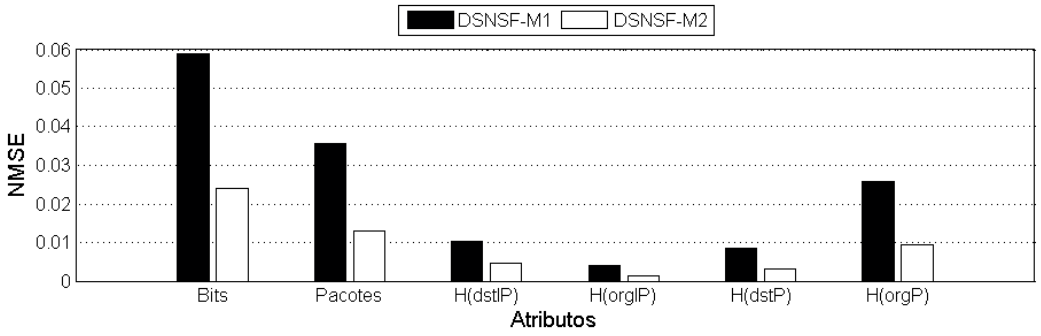


Figura 5.11 - NMSE em relação a cada atributo analisado para o período de 29 de abril a 17 de maio de 2013.

Uma vez que a avaliação da caracterização de tráfego é concluída, inicia-se a avaliação detecção de anomalias por meio da MPC.

5.3.1.2 Detecção de anomalias

Antes dos resultados da acurácia do sistema serem apresentados, toma-se como exemplo as saídas da Máquina Paraconsistente Correlacional (MPC) para dois dias diferentes. Primeiro para uma segunda-feira, 29 de outubro de 2012.

As Figuras 5.12, 5.13 e 5.14 ilustram os níveis de evidência real (ω_{er}) ao longo das 24 horas do dia 29, com uma MPC configurada para detecção das anomalias do tipo *DoS, DDoS e Flash Crowd*, respectivamente. Os modelos utilizados na extração de informação são M1 e M2, os quais produziram os DSNSF contendo os perfis de atributos ilustrados na Figura 5.6 (linhas continuas preta e vermelha). Para cada intervalo t está associado um ω_{er} que é traduzido como a possibilidade de existir uma anomalia nesse dado intervalo. Esse valor varia entre 0 e 1, quanto mais próximo de 1 maior a chance da rede estar sendo alvo de uma das anomalias estudadas.

Os reticulados de Hasse associados as saídas da MPC para os três tipos de anomalia são ilustradas nas Figuras 5.15 (*Dos*), 5.16 (*DDos*) e 5.17 (*Flash Crowd*). Cada ponto (losango) no interior do reticulado representa a interpolação entre os Níveis de Certeza e Contradição (η_c, η_{ct}) gerados na avaliação paraconsistente de um intervalo, feita por meio de uma MPC configurada para cada tipo de anomalia. Assim, um total de 1440 pontos são alocados em cada reticulado.

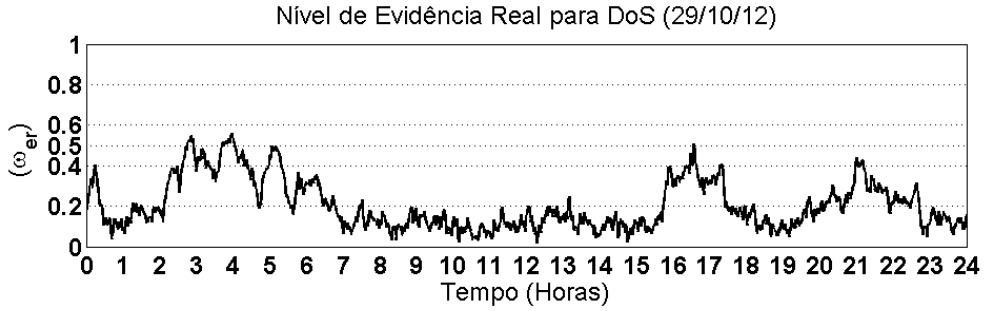


Figura 5.12 – Nível de evidência real da MPC configurada para anomalia do tipo *DoS*, segunda-feira - 29/10/2012.

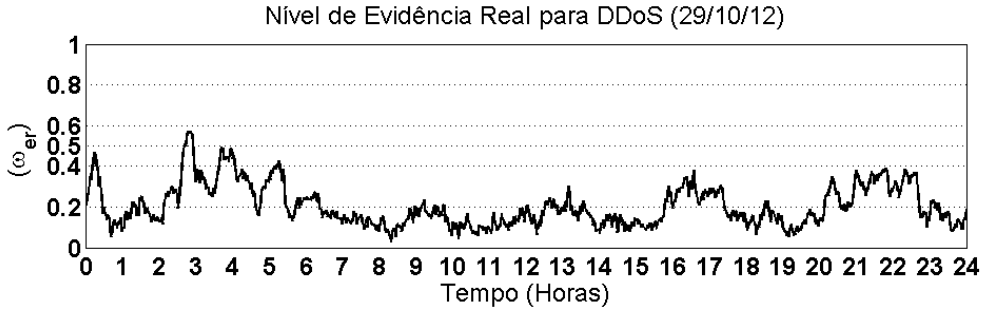


Figura 5.13 - Nível de evidência real da MPC configurada para anomalia do tipo *DDoS*, segunda-feira - 29/10/2012.

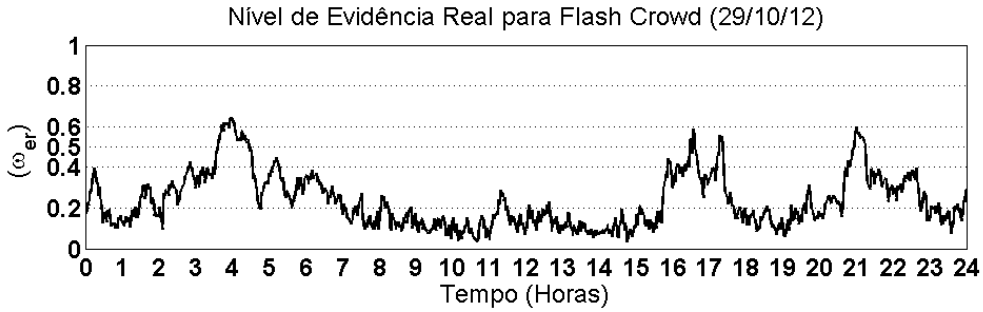


Figura 5.14 - Nível de evidência real da MPC configurada para anomalia do tipo *Flash Crowd*, segunda-feira - 29/10/2012.

Para a MPC configurada para anomalia *DoS* a intensidade de ω_{er} atinge seu máximo por volta das 4 horas. Outros horários também possuem uma leve elevação em seus níveis de evidência real: aproximadamente entre 2h30 e 4h30, 5 horas e 5h30, às 16h20 e entre 21 horas e 21h20. O comportamento dos níveis de evidência real da MPC configurada para *DDoS* e *Flash Crowd* são semelhantes ao da configurada para *DoS*, sendo que a configuração *Flash Crowd* produziu picos de ω_{er} um pouco maiores.

Ao analisar os reticulados associados a cada anomalia para o dia 29, nota-se que a distribuição dos pontos está localizada em sua maioria nos limites dos estados lógicos conotados como Falso e Quase Falso, com uma pequena parcela de pontos nos limites do estado lógico Incompleto. A exceção fica para o reticulado associado ao *Flash Crowd*, com alguns intervalos alocados entre os limites do estado lógico Quase Verdadeiro.

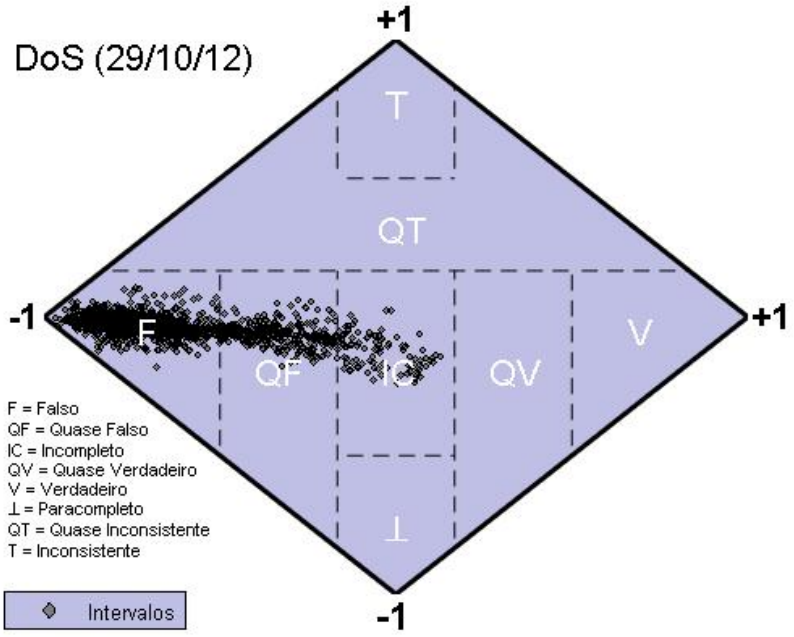


Figura 5.15 – Reticulado de Hasse de uma MPC configurada para anomalia do tipo *DoS*, segunda-feira - 29/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (η_c, n_{ct}).

Intuitivamente, percebe-se que os pontos (η_c, n_{ct}) localizados mais à direita do reticulado são aqueles com o maior nível de evidência real. No entanto, em nenhum dos 3 casos houve a necessidade de um alerta, pois os pontos classificados se mantiveram distantes do estado lógico Verdadeiro. A configuração que o obteve pontos mais próximos desse estado foi para a anomalia *Flash Crowd*. Esse tipo de anomalia costuma se iniciar por atitudes não maliciosas, como acesso a um recurso legítimo da rede por grande número de pessoas.

Assim, com base nas saídas da MPC (reticulado e níveis de evidência real), mesmos os períodos com os maiores níveis evidenciais para presença de anomalias são classificados como "Normais", com pequenos indícios de um possível problema, mas sem a necessidade de alerta. As saídas da MPC sintetizam o comportamento das leituras reais do tráfego, perfis descritos pelas DSNSFs, e características comuns das anomalias estudadas em medidas simples, utilizadas na tomada de decisão.

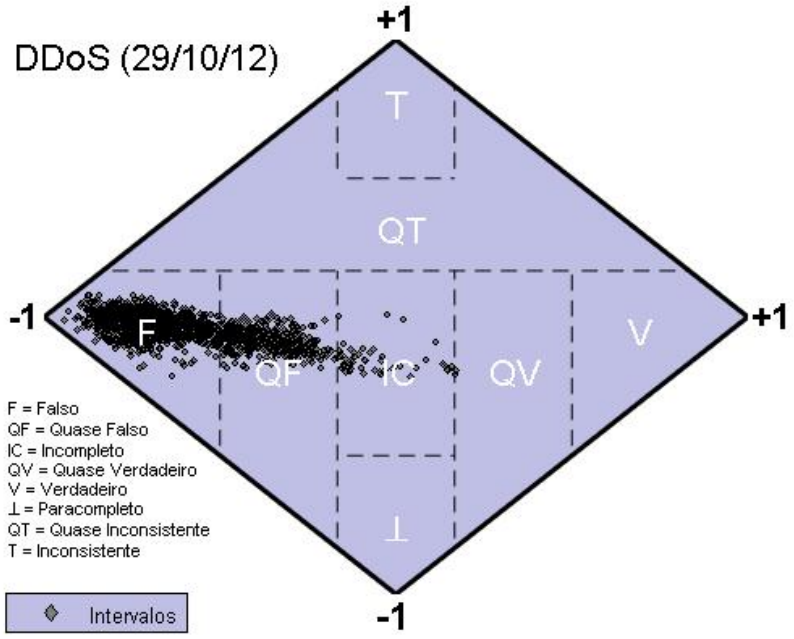


Figura 5.16 - Reticulado de Hasse de uma MPC configurada para anomalia do tipo *DDoS*, segunda-feira - 29/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (η_c, n_{ct}).

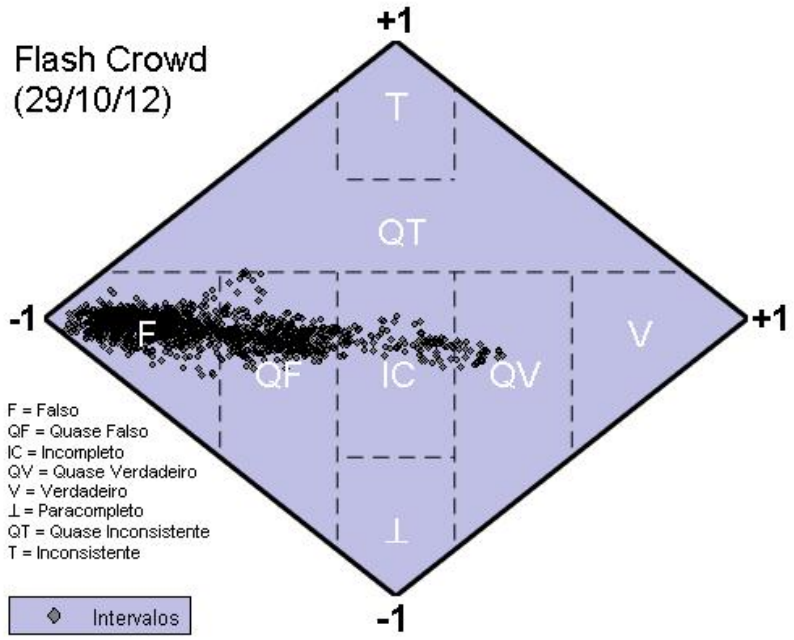


Figura 5.17 - Reticulado de Hasse de uma MPC configurada para anomalia do tipo *Flash Crowd*, segunda-feira - 29/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (η_c, n_{ct}).

Outro exemplo para ilustração das saídas da MPC é o do comportamento de suas saídas para uma terça-feira, 30 de outubro de 2012. As Figuras 5.18, 5.19 e 5.20 ilustram os níveis de evidência real ao longo das 24 horas do dia 30, cada figura representando uma das três configurações da MPC. Os reticulados associados às saídas da MPC são ilustradas nas Figuras 5.21 (*DoS*), 5.22 (*DDoS*) e 5.23 (*Flash Crowd*).

Figura 5.18 - Nível de evidência real da MPC configurada para anomalia do tipo *DoS*, terça-feira - 30/10/2012.

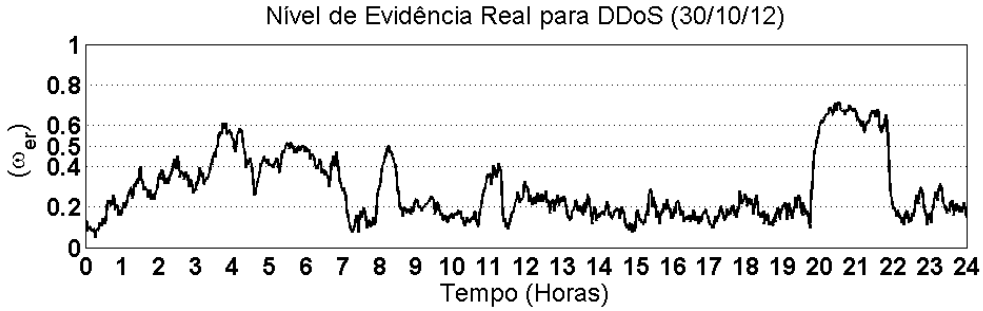


Figura 5.19 - Nível de evidência real da MPC configurada para anomalia do tipo *DDoS*, terça-feira - 30/10/2012.

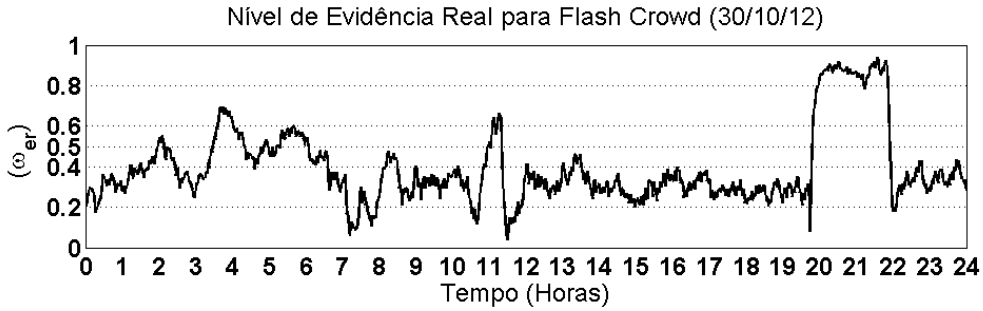


Figura 5.20 - Nível de evidência real da MPC configurada para anomalia do tipo *Flash Crowd*, terça-feira - 30/10/2012.

A intensidade de ω_{er} no dia 30 atinge seu máximo entre 20 e 22 horas, para as três configurações. Nesse horário, A MPC configurada para identificar *Flash Crowd* produziu os níveis de evidência real com picos muito próximos de 1.

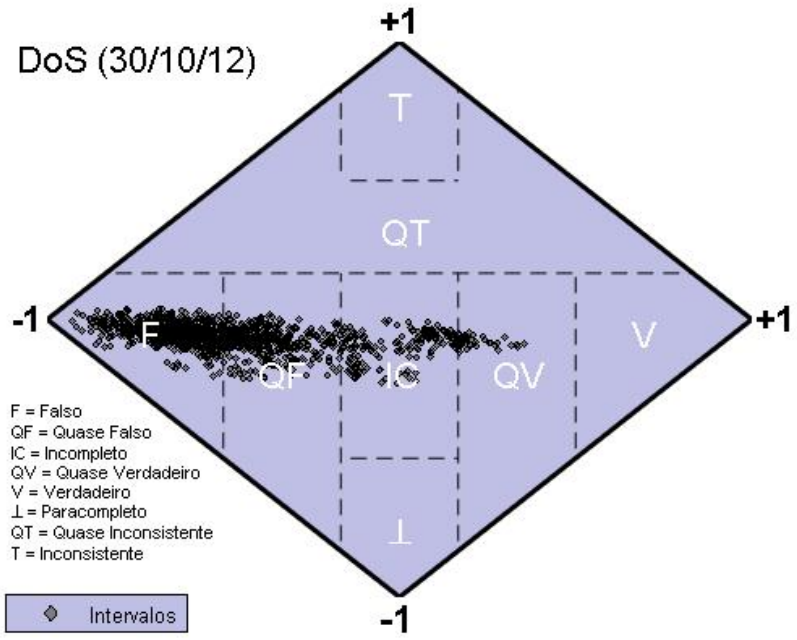


Figura 5.21 - Reticulado de Hasse de uma MPC configurada para anomalia do tipo *DoS*, terça-feira - 30/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (η_c, n_{ct}).

Ao se observar os indícios apresentados pelos níveis de evidência real e os reticulados relativos ao dia 30, nota-se que, para a MPC configurada para *DoS* e *DDoS*, alguns pontos se limitam ao estado Quase Verdadeiro. No entanto, para a configuração *Flash Crowd* o estado Verdadeiro é alcançado. Toda vez que a interpolação (η_c, η_{ct}) atinge tal estado um "Alarme" é disparado.

O comportamento de tráfego do dia 30 pode ser observado na Figura 5.7. É visível a discrepância entre assinaturas digitais e leituras reais de tráfego. Com exceção da entropia de endereço IP de origem, todos os atributos foram afetados drasticamente. Essa é justamente a característica ou assinatura da anomalia *Flash Crowd*.

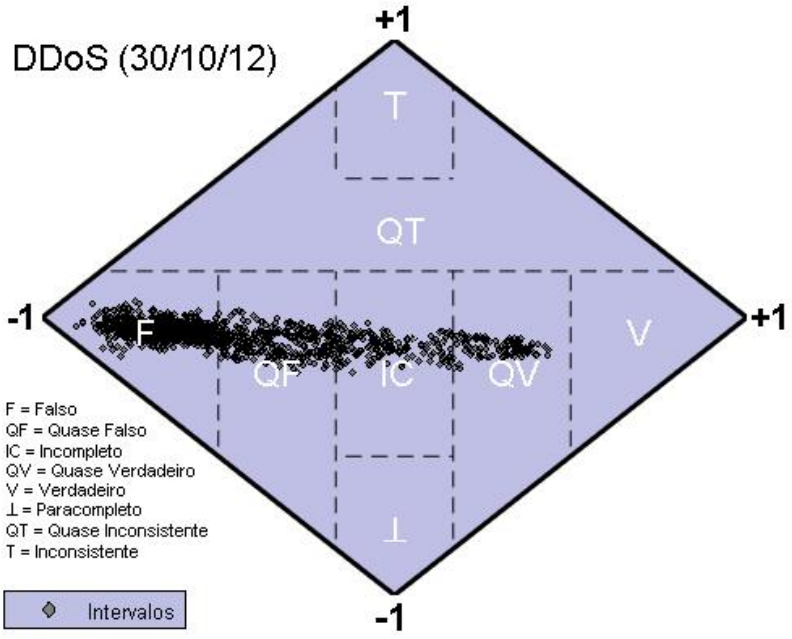


Figura 5.22 - Reticulado de Hasse de uma MPC configurada para anomalia do tipo *DDoS*, terça-feira - 30/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (η_c, n_{ct}).

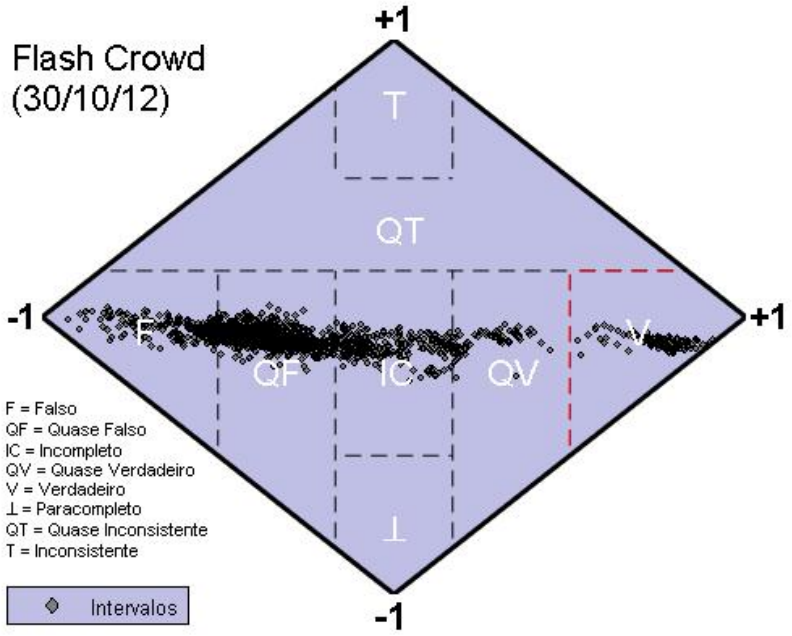


Figura 5.23 - Reticulado de Hasse de uma MPC configurada para anomalia do tipo *Flash Crowd*, terça-feira - 30/10/2012. Cada ponto (◊) representa um intervalo do dia classificado com sua relativa localização (η_c, n_{ct}). Quando a interpolação (η_c, n_{ct}) atinge o estado Verdadeiro um alarme é disparado.

Ao utilizar o potencial que a análise de fluxo fornece, foi possível identificar os endereços e portas de destino mais acessadas. Os mesmos referiam-se a servidores *web* responsáveis por divulgar resultados de concursos e vestibulares. No dia analisado era prevista a divulgação do resultado de um concurso público, motivo que acarretou o grande volume de acessos. A identificação desse tipo de anomalia, mesmo que não maliciosa, é de extrema importância para prevenção de possíveis quedas ou comprometimento do funcionamento da rede e dos diversos serviços prestados.

Na sequência, com as classificações produzidas pela MPC e o gabarito sinalizado pelos administradores de rede para o período entre 22 de outubro e 9 de novembro, a métrica acurácia pode ser avaliada. A Figura 5.24 ilustra os resultados obtidos pelo sistema por meio da utilização dessa métrica. Ela também apresenta os resultados obtidos para a abordagem tradicional utilizando limites de confiança e o modelo ARIMA baseado na mediana, escolhido por causa dos melhores resultados na caracterização de tráfego.

Com exceção dos dias 22 e 23 de outubro e 9 de novembro, o sistema baseado na MPC apresenta índices de acurácia superiores a 95%. Mesmo com os valores inferiores desses três dias a média de acurácia do sistema se manteve em 96,35%. A abordagem baseada em limites de confiança obteve uma média de 94,63%. Se comparadas as médias das duas abordagens, o sistema proposto obtém uma vantagem de 1,72 pontos percentuais. Dessa forma, os resultados indicam uma boa efetividade do sistema em classificar intervalos como anômalos ou normais, o que auxilia administradores de rede na detecção de comportamentos anômalos com precisão.

Complementando a taxa de acurácia e evidenciando o *trade-off* (equilíbrio) entre as Taxas de Verdadeiros Positivos (TVP) e Falsos Positivos (TFP) de cada abordagem as Curvas ROC são ilustradas na Figura 5.25. As curvas são relativas a todo o período avaliado na detecção de anomalias, entre 22 de outubro e 9 de novembro. Cada ponto da curva representa a sensibilidade da abordagem proposta com relação a TFP, passando por extremos onde a classificação feita não encontra falsos positivos mas também não encontra verdadeiros positivos ou todos os verdadeiros positivos são classificados, mas com alta TFP. Esses extremos são vistos na Figura 5.25 nas proximidades dos pontos (0, 60) e (50, 60), respectivamente. Um bom classificador consegue atingir um equilíbrio entre TVP e TFP, e um classificador perfeito obtém resultado igual a TVP=100% e TFP=0%.

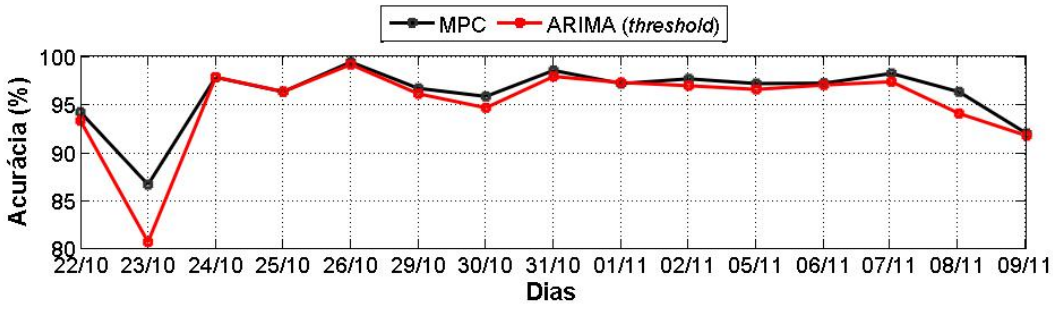


Figura 5.24 – Acurácia do sistema proposto utilizando a MPC e comparação com abordagem baseada em limites de confiança. Resultados para o período entre 22 de outubro e 9 de novembro.

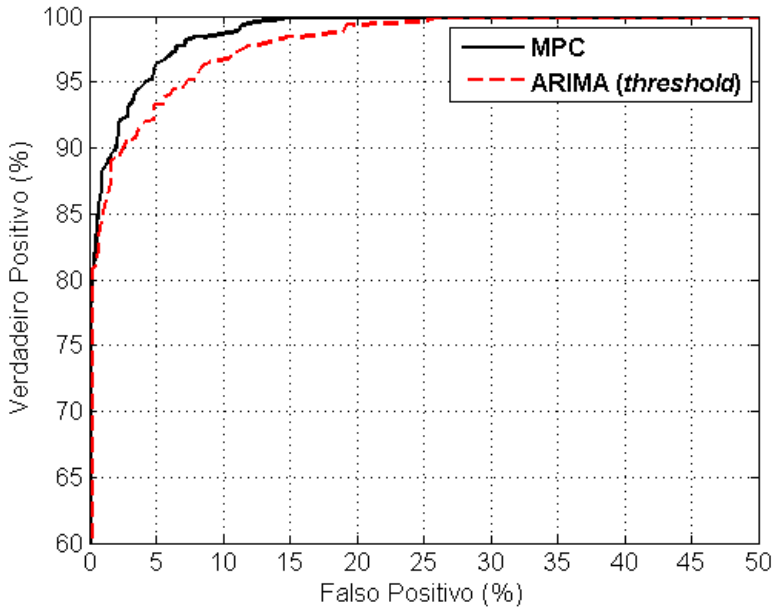


Figura 5.25 - Curvas ROCs do sistema proposto e abordagem baseada em limites de confiança. Resultado para o período entre 22 de outubro e 9 de novembro.

Cada linha ou curva representa o desempenho da abordagem utilizada para classificação de intervalos em função do ganho entre as Taxas de Falso Positivo (TFP) e Verdadeiro Positivo (TVP). Quanto mais próxima uma linha estiver das taxas de 0% de TFP e 100% de TVP, melhores são os resultados de classificação, ou também, quanto maior a área abaixo das curvas, melhores são os resultados do sistema. Um exemplo de *trade-off* atingido pelo sistema proposto é o de aproximadamente 96,5% de TVP para 5% de TFP, mostrando boa precisão na classificação de intervalos anômalos. Nota-se também que a área abaixo da curva

relativa à MPC tem uma área maior do que a curva relativa à abordagem baseada nos limites de confiança, indicando melhores resultados para proposta feita nesse trabalho.

5.3.2 Cenário 2 – Detecção de anomalias com a MPC e com o ACODS (UEL)

O segundo cenário utiliza o ambiente de rede da Universidade Estadual de Londrina e compara a abordagem proposta nesse trabalho baseada na MPC gerada com Modelos ARIMA (com ruído branco baseado na média e mediana) e modelo ACODS (*Ant Colony Optimization for Digital Signature*) na detecção de anomalias.

As assinaturas digitais construídas com o modelo ACODS foram utilizadas na detecção de anomalias baseada em limites de confiança. Essa abordagem leva em consideração apenas os atributos afetados por uma determinada anomalia.

Os registros de fluxos IP utilizados para testar a detecção de anomalias compreendem os períodos entre 24 de setembro e 19 de outubro de 2012 para treinamento e entre 22 de outubro e 9 de novembro de 2012 para testes. A inclusão de anomalias sintéticas seguiu as mesmas configurações do cenário 1. Nesse mesmo período também foi testada a caracterização de tráfego utilizando o modelo ACODS.

5.3.2.1 Caracterização de tráfego

O conjunto de Figuras de 5.26 a 5.30 ilustram o comportamento das leituras reais de tráfego e as especificações geradas pelas DSNSFs entre os dias 22 e 26 de outubro de 2012 no ambiente de rede da Universidade Estadual de Londrina. Nesse conjunto de ilustrações foram incluídas as caracterizações feitas pelo modelo ACODS (linhas pontilhadas na cor azul).

Nota-se que os três modelos apresentam tendências similares. Algumas variações ocorrem em intervalos específicos, principalmente onde o tráfego agregado tem menor intensidade. Assim, cada um dos modelos apresenta características particulares referentes às expectativas de tráfego normal. Essas características motivam o uso de um recurso apropriado para que anomalias sejam identificadas apenas quando os modelos concordam entre si. Como será visto no próximo cenário, os modelos ARIMA e ACODS podem ser empregados em modo cooperativo por meio da MPC. Nesse cenário, a MPC com os modelos ARIMA e a abordagem baseada no ACODS com limites de confiança são comparados.

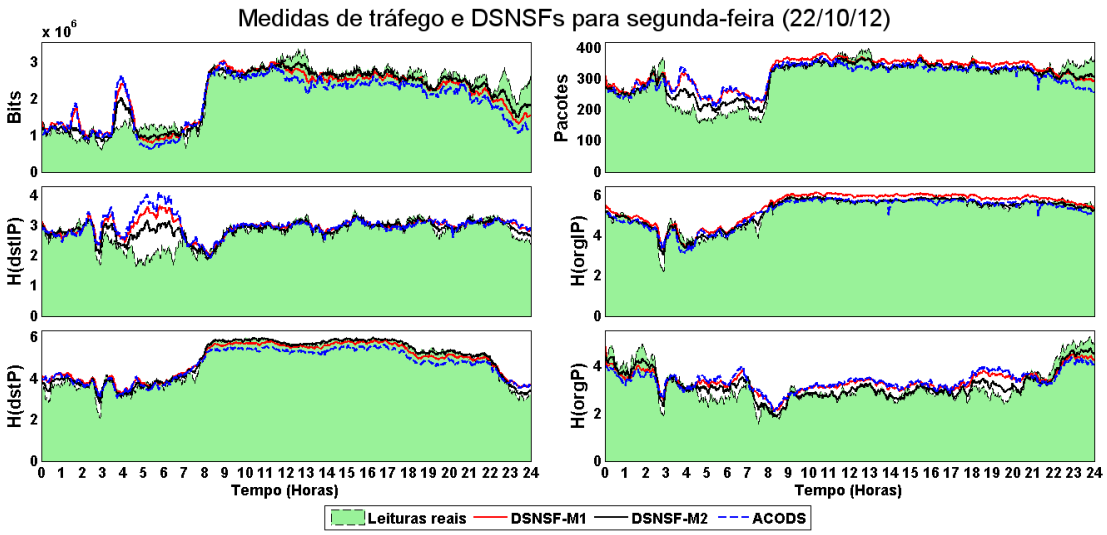


Figura 5.26 – Leituras reais de tráfego e DSNSFs para rede da UEL no dia 22/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS.

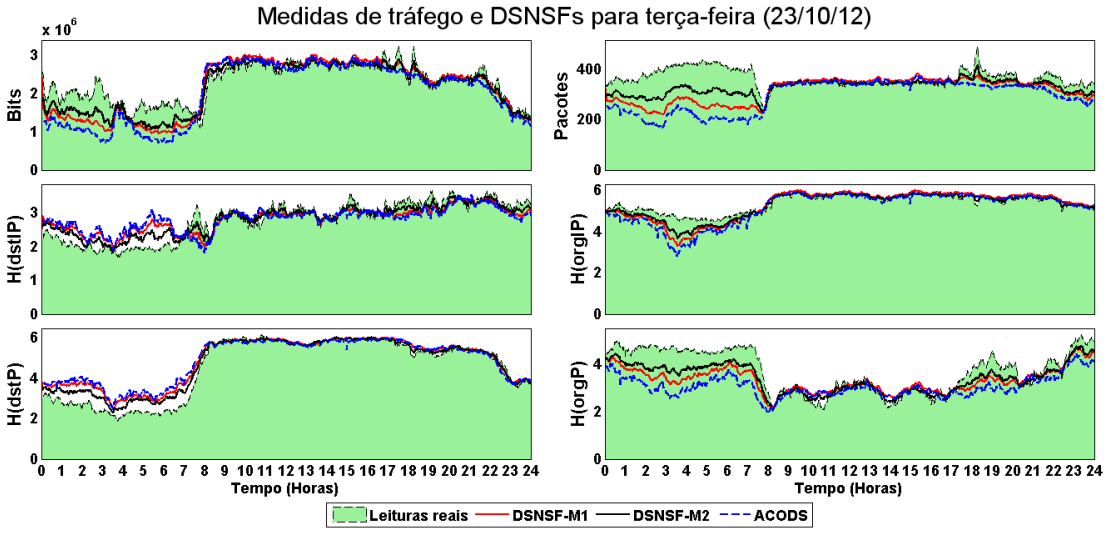


Figura 5.27 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 23/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS.

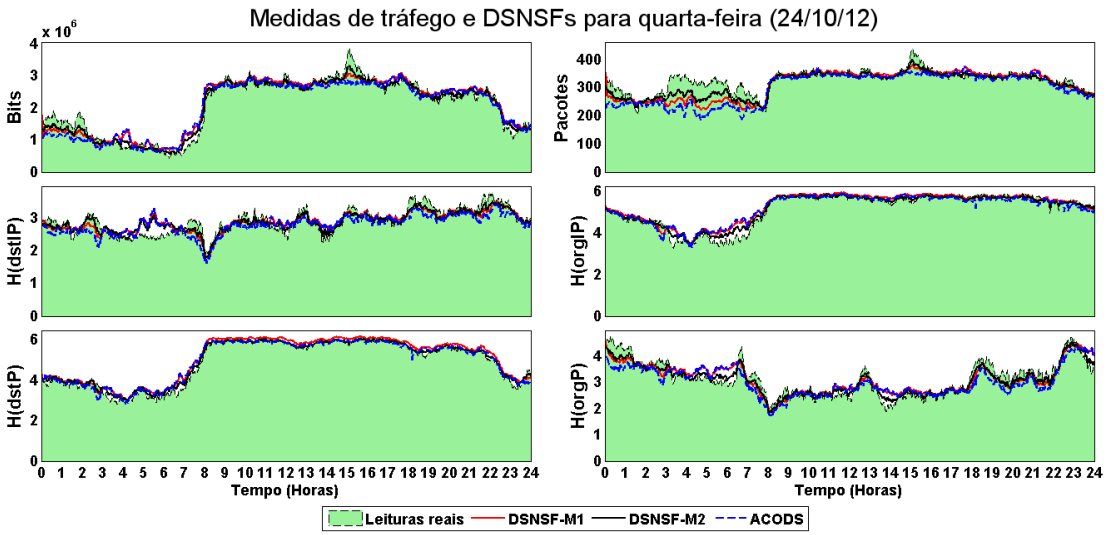


Figura 5.28 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 24/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS.

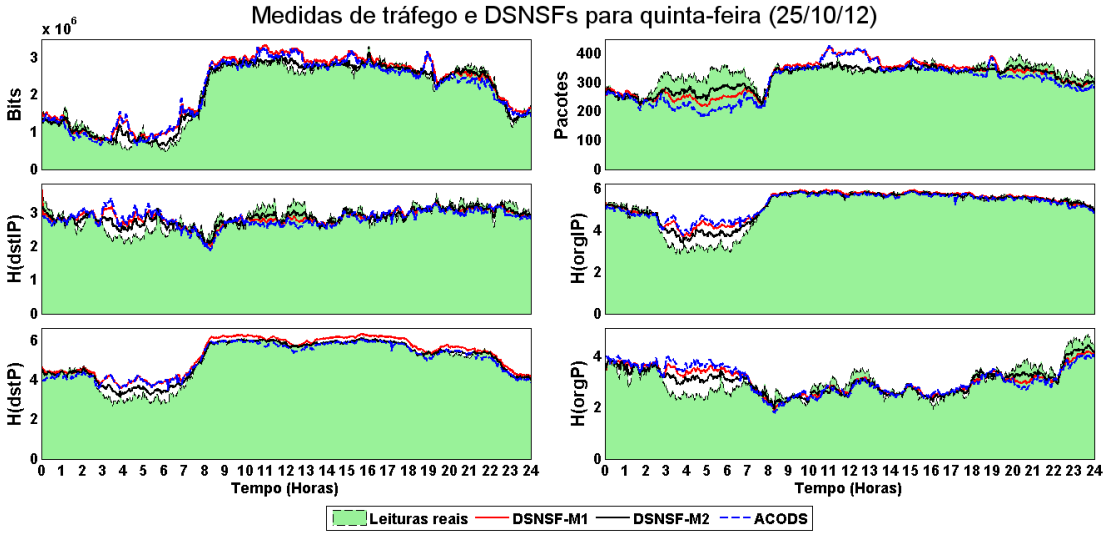


Figura 5.29 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 25/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS.

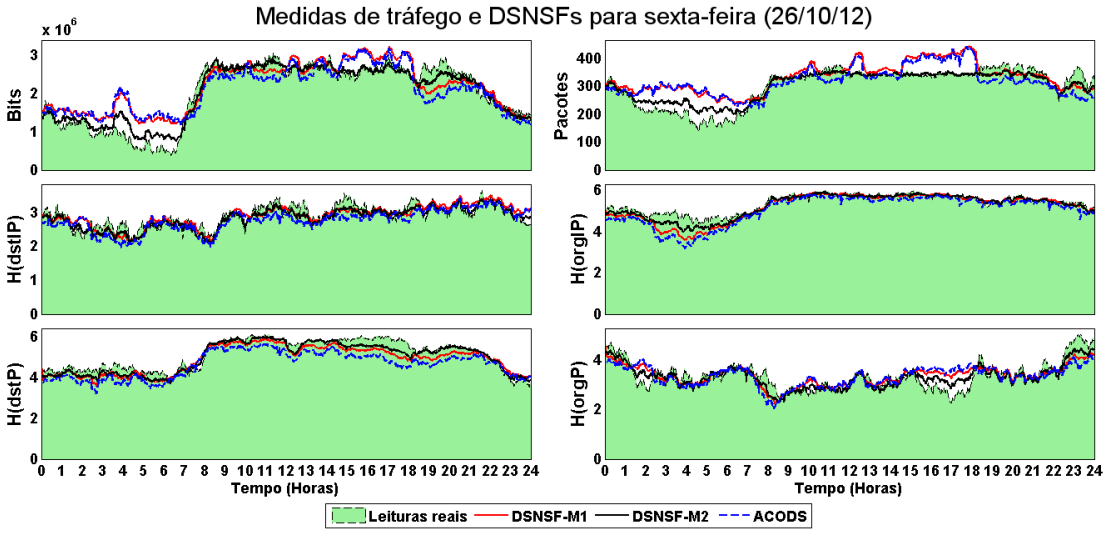


Figura 5.30 - Leituras reais de tráfego e DSNSFs para rede da UEL no dia 26/10/2012. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana; ACODS.

As Figuras 5.31 e 5.32 ilustram os resultados dos modelos ARIMA (M1 e M2) e ACODS na caracterização de tráfego em função do coeficiente de correlação. Os resultados obtidos pelo ACODS são em geral próximos aos obtidos por M1, e inferiores a M2. O atributo H(dstIP) foi o que apresentou os coeficientes de correlação mais baixos.

Em relação ao NMSE, as Figuras 5.33 e 5.34 ilustram os resultados obtidos pelos três modelos avaliados. O modelo M2 obteve os menores erros enquanto o modelo M1 obteve os maiores, na maior parte dos dias. Nesse aspecto o atributo mais vulnerável foi Bits.

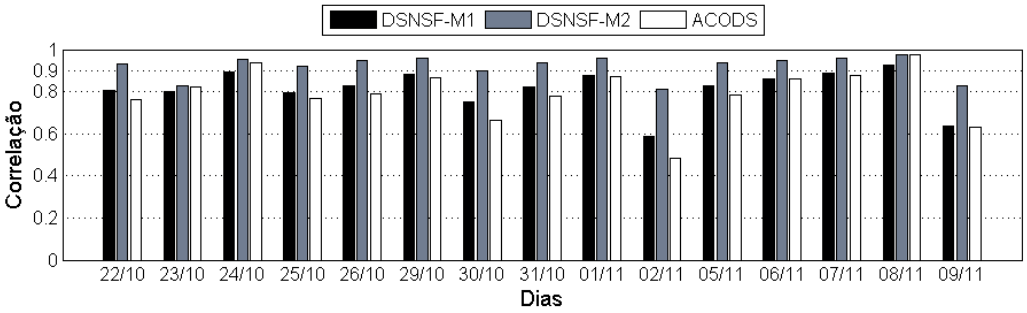


Figura 5.31 - Média dos coeficientes de correlação, calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 22 de outubro a 9 de novembro de 2012. As medidas ilustradas referem-se a média dos seis atributos analisados, para M1, M2 e ACODS.

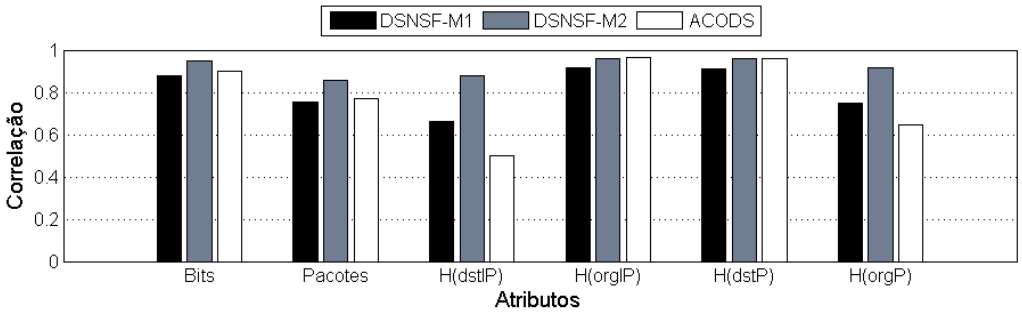


Figura 5.32 - Coeficientes de correlação em relação a cada atributo analisado para o período de 22 de outubro a 9 de novembro de 2012, para M1, M2 e ACODS.

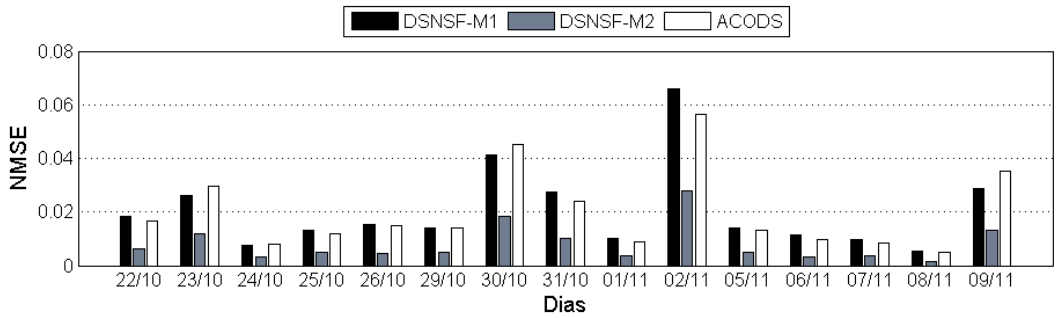


Figura 5.33 - Média dos erros NMSE calculados entre os DSNSFs e as leituras reais do tráfego da rede da UEL no período de 22 de outubro a 9 de novembro de 2012. As médias ilustradas referem-se aos seis atributos analisados, para M1, M2 e ACODS.

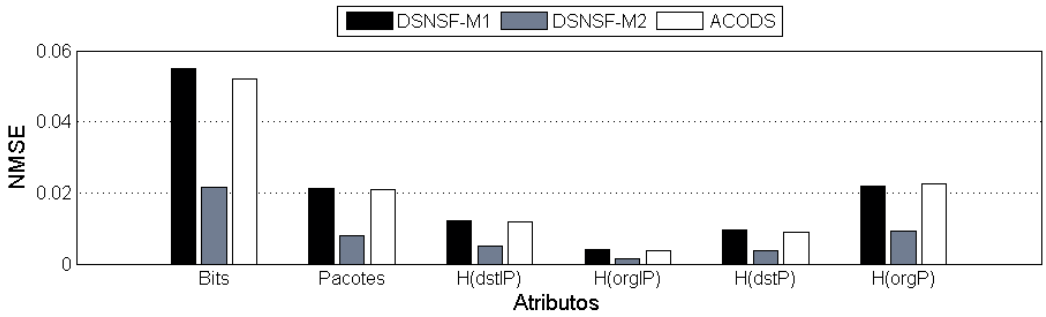


Figura 5.34 - Erros NMSE em relação a cada atributo analisado para o período de 22 de outubro a 9 de novembro de 2012, para M1, M2 e ACODS.

Em geral, os resultados obtidos na caracterização de tráfego no período relativo a esse cenário foram inferiores aos obtidos no período de testes da caracterização do cenário 1. Atividades ao longo das madrugadas como *backups* contribuíram para tanto, apesar disso, os resultados obtidos continuaram satisfatórios.

5.3.2.2 Detecção de anomalias

A Figura 5.35 ilustra os resultados obtidos pelo sistema proposto e pela abordagem ACODS por meio da utilização da acurácia. A Figura 5.36 ilustra os resultados da detecção em termos das curvas ROC.

A abordagem baseada no ACODS obteve uma média de acurácia de 94,51%. Se comparada com a média do sistema proposto (96,35%), uma diferença de -1,84 pontos percentuais é observada. Assim, os resultados do modelo proposto continuam sendo os melhores obtidos. Esse melhor desempenho também pode ser observado nas curvas ROC, em que a MPC atinge um *trade-off* aproximado de 96,5% de TVP para 5% de TFP enquanto o ACODS de aproximadamente 95% de TVP para 8% de TFP.

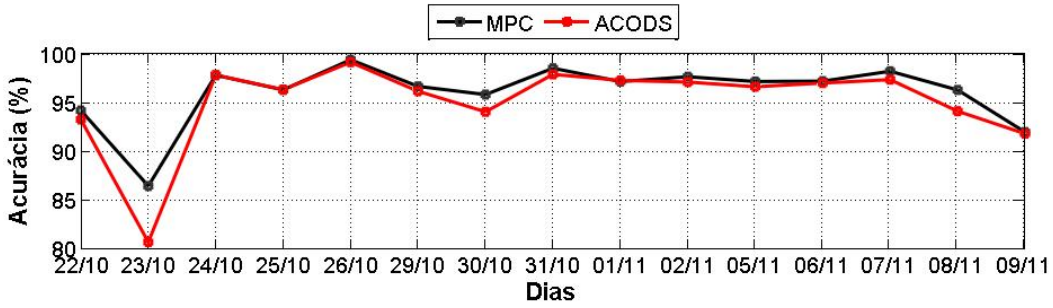


Figura 5.35 – Acurácia do sistema proposto utilizando a MPC e comparação com ACODS. Resultados para o período entre 22 de outubro e 9 de novembro.

Os resultados obtidos com o modelo ACODS, tanto nesse trabalho como em trabalhos encontrados na literatura, foram motivadores e levaram a criação de um terceiro cenário cujos modelos ARIMA e ACODS são utilizados em cooperação por meio da abordagem paraconsistente.

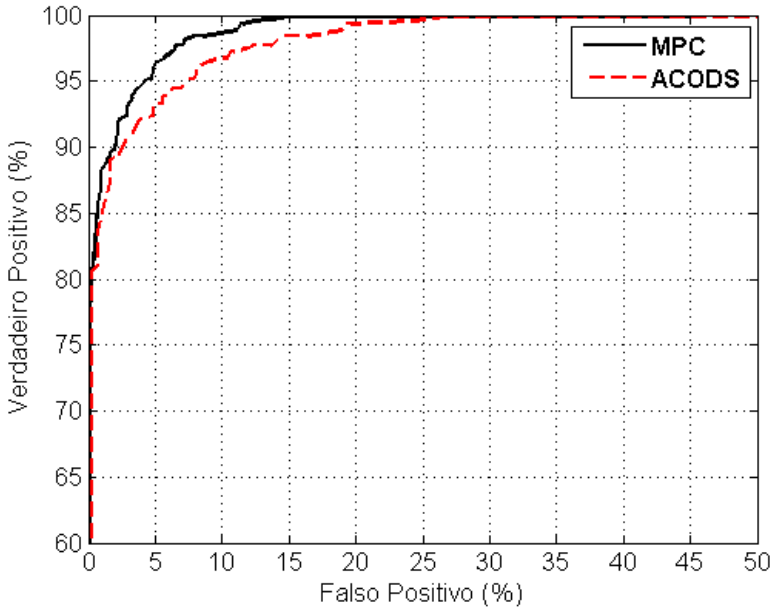


Figura 5.36 - Curva ROC do sistema proposto e comparação com abordagem ACODS. Resultado para o período entre 22 de outubro e 9 de novembro.

5.3.3 Cenário 3 –MPC baseada no ARIMA e ACODS (UEL)

O terceiro cenário utiliza o ambiente de rede da Universidade Estadual de Londrina e uma MPC construída com base no ARIMA e ACODS. Esse cenário foi construído em cima dos mesmos registros de fluxos IP utilizados para testar a detecção de anomalias nos cenário 1 e 2, incluindo injeção de anomalias. Os testes de caracterização dos dois modelos utilizados já foram feitos nos cenários 1 e 2.

Os modelos utilizados como fontes de informações para a MPC tradicional foram os mesmos dos cenários anteriores e para a segunda MPC foram: M1 → "ARIMA baseado em ruído branco gerado com a mediana de 4 semanas de dados históricos" e M2 → "ACODS baseada em treinamento com 4 semanas de dados históricos". Os testes com essa adaptação da MPC tem como objetivo verificar o desempenho da mesma quando alimentada com fontes de informações diversificadas, que nesse caso são modelos para cálculo das DSNSF de diferentes classes de algoritmos.

5.3.3.1 Detecção de anomalias

A Figura 5.37 ilustra os níveis de evidência real (ω_{er}) ao longo das 24 horas do dia 30, para a MPC adaptada com o ACODS configurada para a anomalia *Flash Crowd*. O reticulado associado à saída dessa MPC é ilustrado na Figura 5.38.

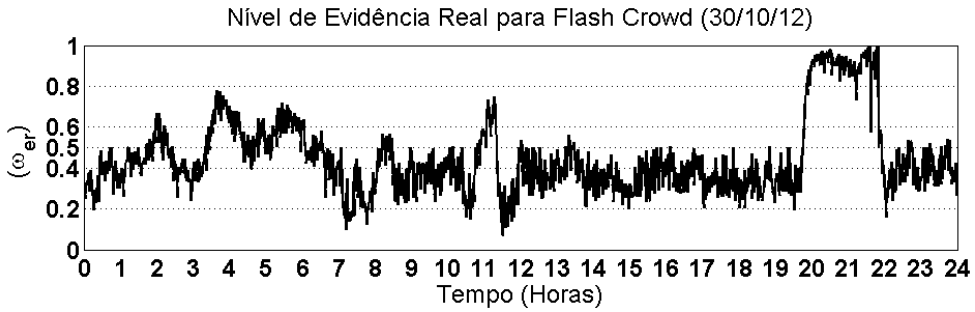


Figura 5.37 - Nível de evidência real da MPC adaptada com o ACODS configurada para anomalia do tipo *Flash Crowd*, terça-feira - 30/10/2012.

Um diferença visível entre a saída da MPC baseada nos dois modelos ARIMA (Cenário 1 e 2) e da MPC baseada nos modelos ARIMA e ACODS (Cenário atual) é a granularidade das intensidades $\omega_{er} em$ função do tempo. Enquanto as diferenças entre intensidades ω_{er} da primeira MPC são mais regulares ou suaves as diferenças entre as intensidades ω_{er} apresentadas pela segunda MPC são mais acentuadas de um intervalo para outro.

Outra diferença é a distribuição de pontos referentes à interpolação (η_c, η_{ct}) no reticulado. Enquanto no cenário 1 os pontos seguem uma disposição que cria uma impressão de proximidade, os pontos desse cenário (Figura 5.38) são mais espaçados. Além disso, estados como Quase Inconsistente (QT) Inconsistente (T) e Paracompleto ($\perp$), ainda que em poucos casos, são alcançados. Isso ocorre quando os níveis de contradição atingem valores próximos dos extremos $(+1, -1)$.

Quando os dois modelos ARIMA foram utilizados os níveis evidenciais extraídos das duas fontes de informações produziram diferenças menores. Isso ocorre por causa das características similares dos dois modelos, com diferenciação apenas no cálculo estatístico do ruído branco.

Quando os modelos ACODS (clusterização) e ARIMA (previsão de séries temporais) são utilizados na MPC as diferenças evidenciais entre os mesmos são mais elevadas.

Seja um intervalo alocado no estado Inconsistente. A fonte de informação que produziu o nível de evidencia favorável para o comprometimento da rede sinaliza uma grande divergência entre expectativa (DSNSF) e leitura real de tráfego. Por sua vez, a fonte de informação que produziu o nível de evidência desfavorável sinaliza uma divergência próxima de zero entre expectativa e leituras reais. Isso é traduzido em níveis de evidência favorável e desfavorável com os valores próximos de um. O contrário ocorre para um intervalo alocado no estado Paracompleto, onde o nível de evidencia favorável é resultado de divergências entre expectativa e leituras reais próximas de zero e o nível de evidencia desfavorável é resultado de divergências entre expectativa e leituras reais mais elevadas.

De modo geral, os estados relacionados a inconsistência ou paracompletude dizem respeito a avaliações divergentes e evitam que classificações do tipo Verdadeiro sejam feitas sem a total concordância das fontes de informações.

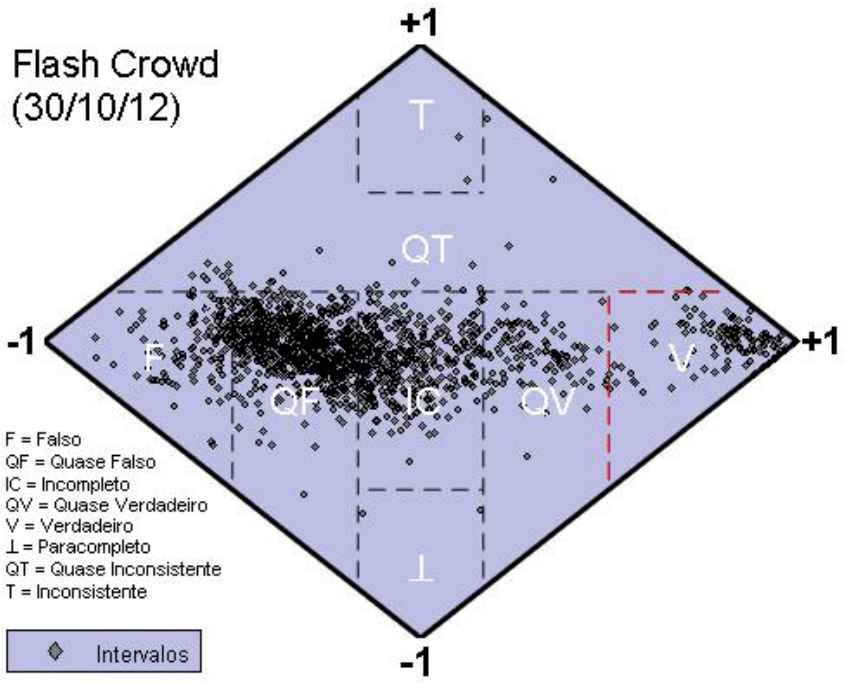


Figura 5.38 - Reticulado de Hasse de uma MPC adaptada com o ACODS configurada para anomalia do tipo *Flash Crowd*, terça-feira - 30/10/2012.

Com relação aos resultados de precisão desse cenário, a Figura 5.39 ilustra os resultados obtidos pelo sistema proposto e pela abordagem com MPC adaptada com o modelo

ACODS por meio da utilização da acurácia. A Figura 5.40 ilustra os resultados da detecção em termos das curvas ROC.

A abordagem baseada na MPC adaptada com o ACODS obteve uma média de acurácia de 95,7%. uma diferença de -0,65 pontos percentuais com relação a MPC original (96,35%). Com relação à curva ROC, a MPC adaptada consegue atingir um *trade-off* aproximado de 95% de TVP para 5% de TFP. Nota-se que a adaptação feita surtiu um efeito positivo, a abordagem baseada apenas no ACODS obteve resultados inferiores a ela.

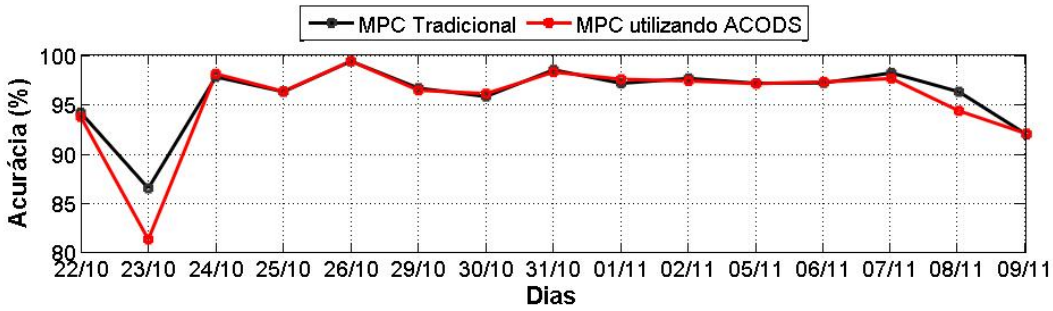


Figura 5.39 – Acurácia do sistema proposto utilizando duas MPCs.

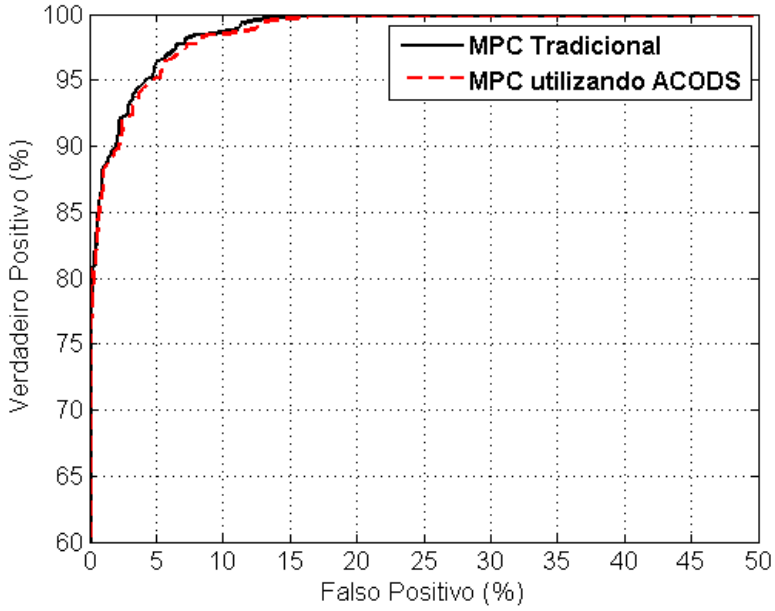


Figura 5.40 - Curva ROC do sistema proposto comparado com adaptação da MPC com ACODS. Resultado para o período entre 22 de outubro e 9 de novembro.

5.3.4 Cenário 4 – Caracterização e detecção de anomalias (UTFPR)

O quarto cenário utiliza o ambiente de rede da Universidade Tecnológica Federal do Paraná (UTFPR), Modelos ARIMA (com ruído branco baseado na média e mediana), e Máquina Paraconsistente Correlacional. Essa rede foi incluída nos testes devido a suas características diferentes da rede UEL, como sua dimensão, protocolo de exportação de fluxos e padrão de amostragem de pacotes de um para um.

Este cenário foi construído em cima de registros de fluxos IP de sete semanas da rede UTFPR. A caracterização de tráfego e detecção de anomalias foram testadas no período entre 1 de abril e 17 de maio de 2013. Dados entre 1 e 26 de abril para treinamento, enquanto o restante do período para teste.

Três anomalias foram sintetizadas com a ferramenta *Scorpius* e combinadas ao tráfego analisado nos dias 6, 7 e 8 de maio de 2013. As informações das mesmas são apresentadas na Tabela 5.3.

Tabela 5.3 – Informações sobre a simulação de anomalias (UTFPR).

Dia	Anomalia	Intervalo de Tempo	IP de origem	Porta de origem	IP de destino	Porta de destino
06/05/2013	*Flash Crowd*	19h00 até 22h00	100 endereços gerados aleatoriamente	80	12.12.12.12	8081
07/05/2013	*DoS*	19h00 até 22h00	11.11.11.11	80	12.12.12.12	8081
08/05/2013	*DDoS*	19h00 até 22h00	50 endereços gerados aleatoriamente	80	12.12.12.12	8081

Os modelos utilizados foram: M1 → “ARIMA baseado em ruído branco gerado com a média de 4 semanas de dados históricos” e M2 → “ARIMA baseado em ruído branco gerado com a mediana de 4 semanas de dados históricos”.

5.3.4.1 Caracterização de tráfego

Exemplos dos resultados visuais da caracterização de tráfego realizada nos dados da rede da Universidade Tecnológica Federal do Paraná são ilustrados no conjunto compreendendo as Figuras 5.41, 5.42 e 5.43, 5.44, 5.45.

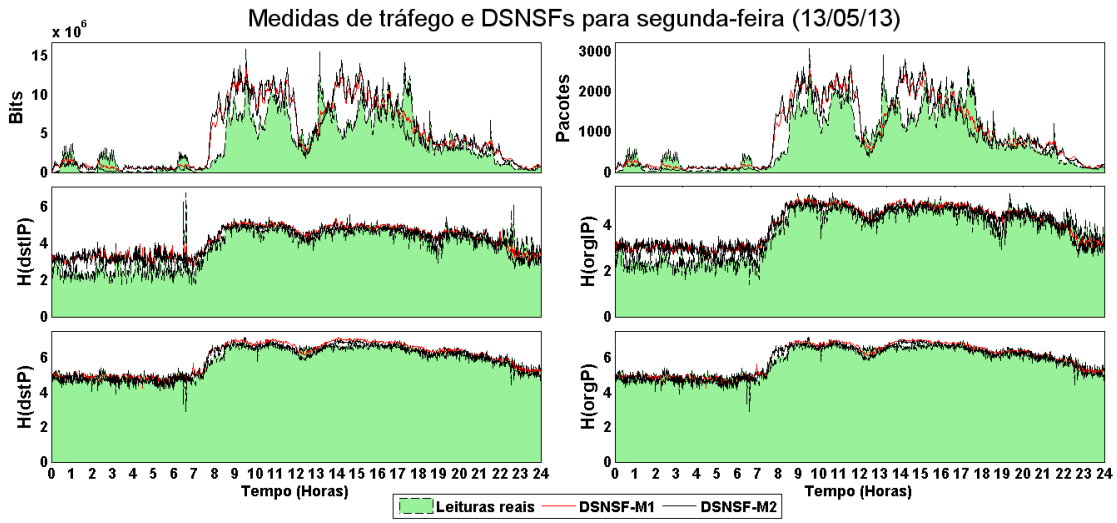


Figura 5.41 - Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 13/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

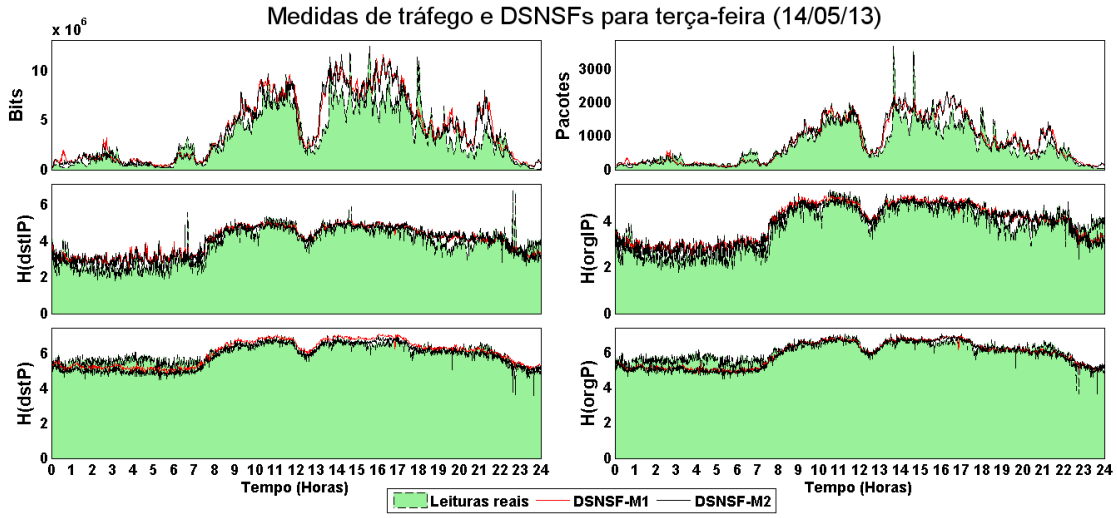


Figura 5.42 - Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 14/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

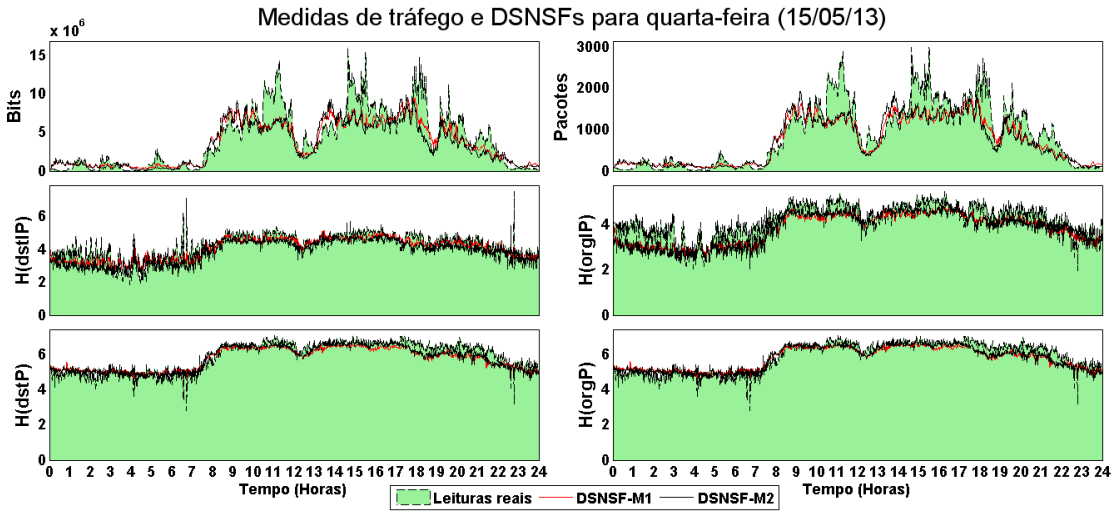


Figura 5.43 - Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 15/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

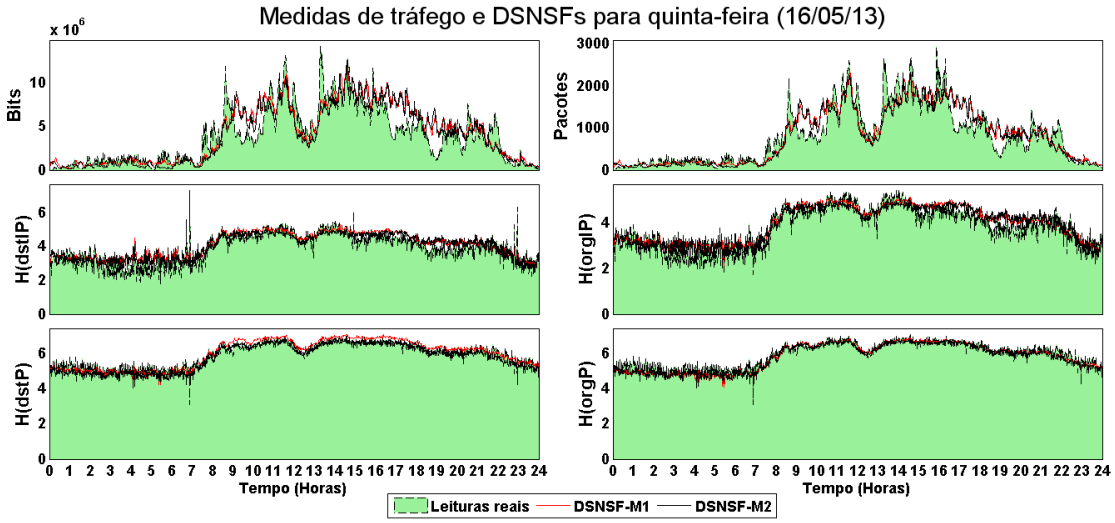


Figura 5.44 - Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 16/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

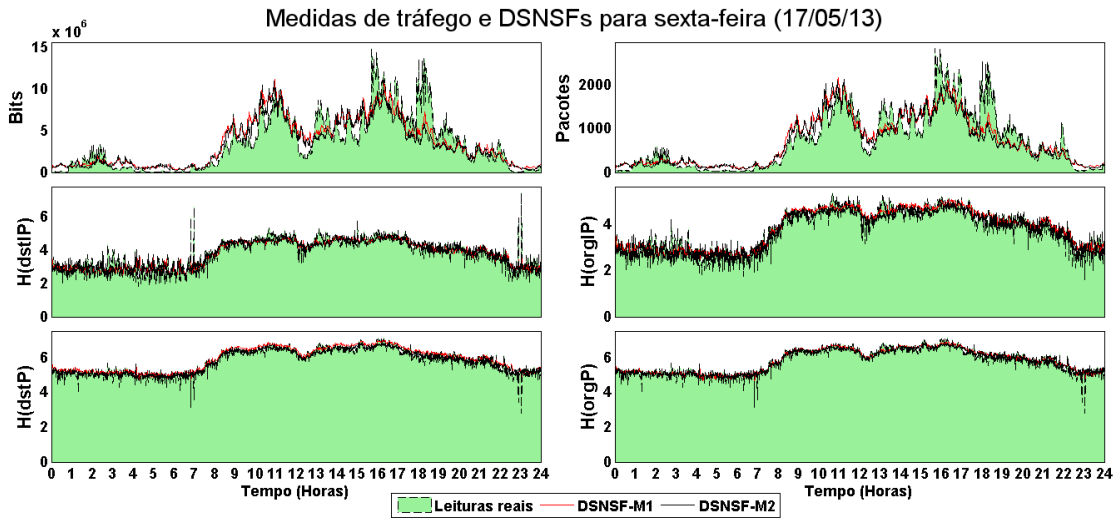


Figura 5.45 - Leituras reais de tráfego e DSNSFs para rede da UTFPR no dia 17/05/2013. DSNSF-M1: ARIMA / Média; DSNSF-M2: ARIMA / Mediana.

As atividades no campus da UTFPR são mais intensas no período das 8 horas às 18 horas. Como pode ser observado nas Figuras, a maior parte do tráfego está concentrada nesse período, resultado das atividades acadêmicas e administrativas da Universidade. Essa rede tem a capacidade mais reduzida, tornando a caracterização do seu tráfego mais difícil. Ainda assim, a caracterização conseguiu ser eficiente nesse ambiente. Prova disso são os coeficientes de correlação, Figura 5.46, com valores entre 0,8 e 0,9 tanto para M1 quanto para M2, com exceção do dia 1 de maio. Esse dia é referente a um feriado nacional. Devido a seu valor muito inferior, as médias relativas a cada atributo (Figura 5.47) caíram consideravelmente. Apesar disso, se mantiveram com média de aproximadamente 0,8 para M1 e M2.

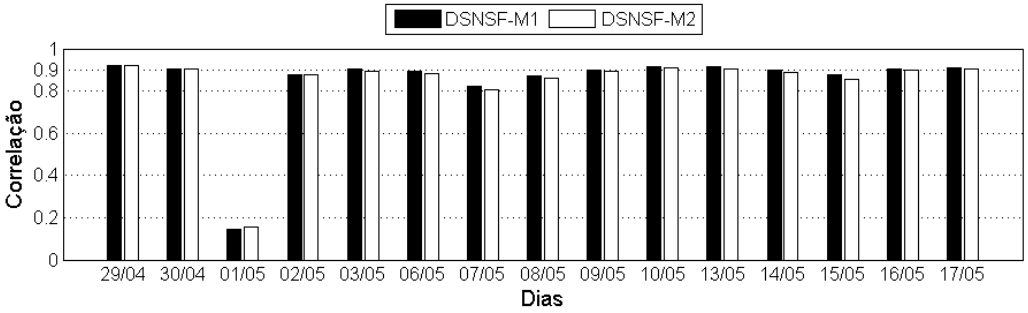


Figura 5.46 - Média dos coeficientes de correlação, calculados entre os DSNSFs e as leituras reais do tráfego da rede da UTFPR no período de 29 de abril a 17 de maio de 2013.

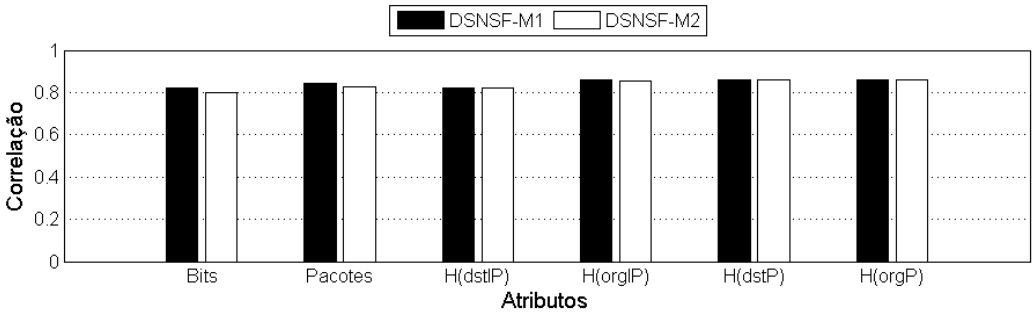


Figura 5.47 - Coeficientes de correlação em relação a cada atributo analisado para o período de 29 de abril a 17 de maio de 2013, UTFPR.

Os modelos também apresentaram bons resultados relativos aos erros NMSE, Figuras 5.48 e 5.49. No entanto se comparado ao cenário da UEL os resultados não são tão constantes. Percebe-se que as dimensões das redes afetam diretamente no processo de caracterização de tráfego. Esse é um fator que deve ser contornado na abordagem de detecção de anomalias. Nota-se que os atributos Bits e Pacotes são aqueles que produzem os maiores erros. Isso acontece devido a escala dos dados trabalhadas, sendo que as entropias costumam estar na casa das primeiras unidades. Mesmo visualmente os atributos de distribuição se comportam de maneira mais regular.

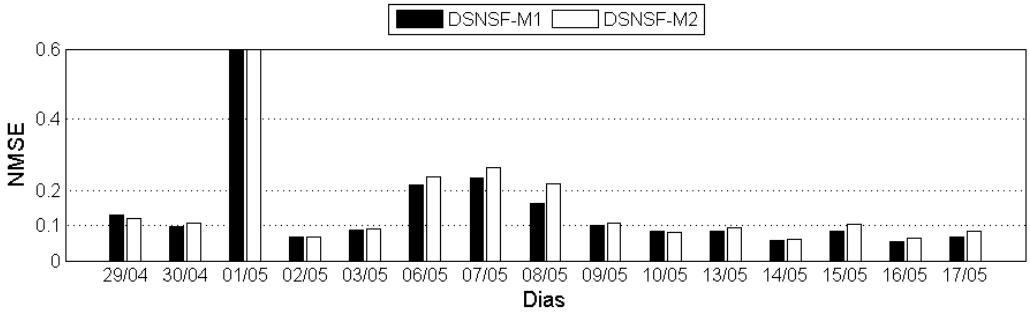


Figura 5.48 - Média do erro NMSE calculados entre os DSNSFs e as leituras reais do tráfego da rede da UTFPR no período de 29 de abril a 17 de maio de 2013.

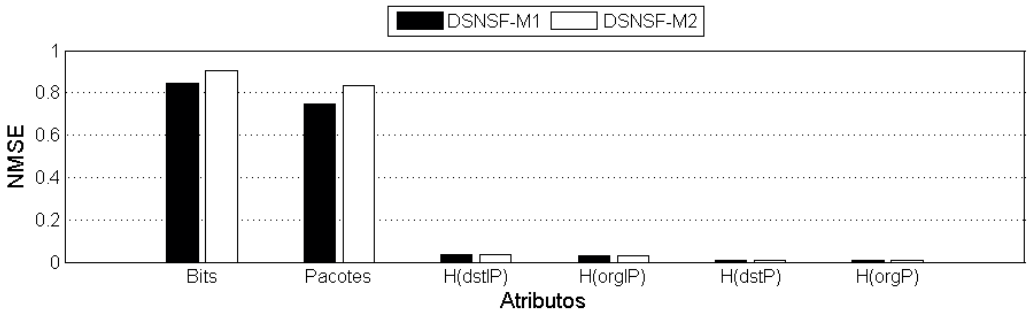


Figura 5.49 - NMSE em relação a cada atributo analisado.

A Figura 5.50 ilustra os resultados obtidos pelo sistema proposto e pela abordagem baseada em limites de confiança, assim como no cenário 1, por meio da utilização da acurácia. A Figura 5.51 ilustra os resultados da detecção em termos das curvas ROC.

Com exceção do dia 1 de abril, feriado nacional, o sistema baseado na MPC apresenta índices de acurácia superiores a 90%, com média geral de acurácia de 94,68%. A abordagem baseada em limites de confiança obteve uma média de 89,62%, uma diferença de 5,06%.

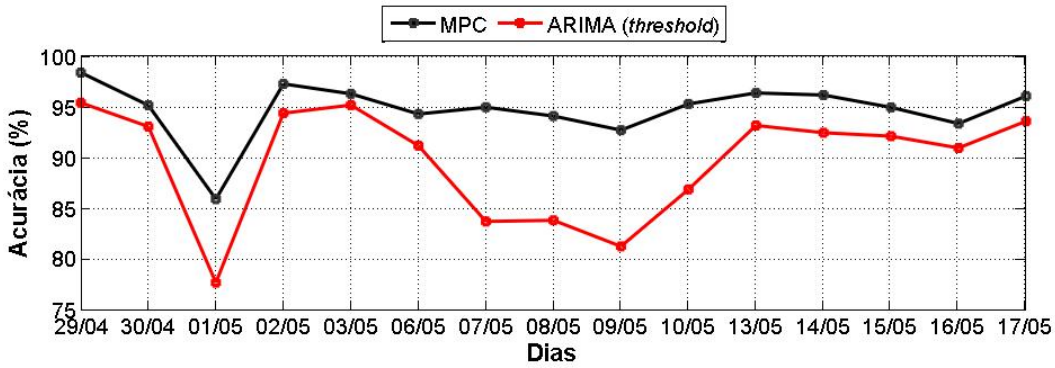


Figura 5.50 – Acurácia do sistema proposto utilizando a MPC e comparação com abordagem baseada em limites de confiança.

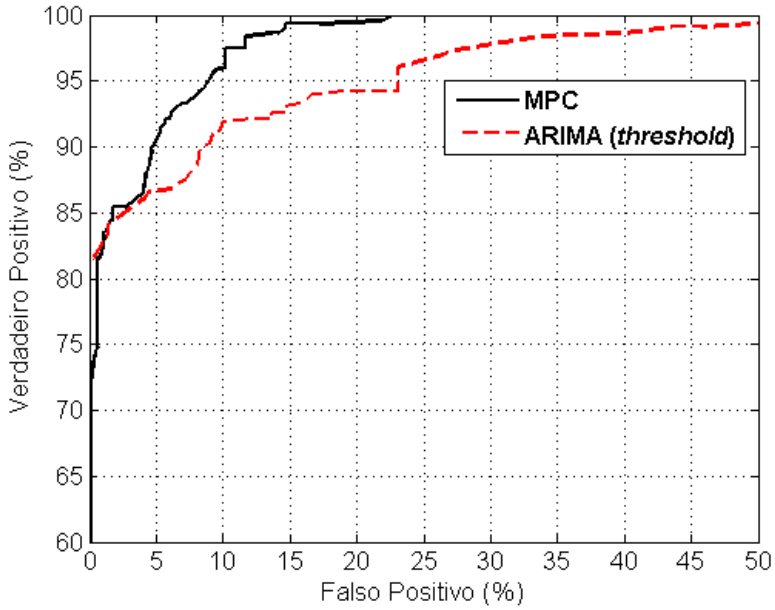


Figura 5.51 - Curva ROC do sistema proposto e comparação com abordagem baseada em limites de confiança. Resultado para o período entre 29 de abril e 17 de maio.

6 CONCLUSÃO

O desenvolvimento de instrumentos que tenham a capacidade de descrever e monitorar o comportamento normal das redes de comunicação representa um grande desafio na área de gerência de redes. Este trabalho apresentou uma solução capaz de avaliar o tráfego de segmentos de redes para identificação de possíveis falhas ou eventos anômalos. A solução foi baseada em uma ferramenta para caracterização de tráfego, chamada Assinatura Digital de Segmento de Rede utilizando análise de Fluxo (também referida como DSNSF), em combinação com a Lógica Paraconsitente (LP).

As assinaturas digitais cumprem seu papel em esboçar perfis referentes ao fluxo de informações de um determinado segmento de rede ao longo do dia. Esse esboço é baseado na análise de fluxos IP, e representa a caracterização do comportamento normal para os atributos: bits por segundo, pacotes por segundo, distribuição de endereços IP de origem e destino e distribuição de números de portas de origem e destino.

Este trabalho desenvolveu uma abordagem baseada na análise de séries temporais utilizando o modelo ARIMA. Assim, utilizando a flexibilidade do mesmo, assinaturas digitais foram construídas com base na moda e mediana de dados históricos e nos valores das leituras reais do tráfego em períodos anteriores à previsão. Essa construção fornece fontes de informações utilizadas pela LP.

A abordagem paraconsistente utiliza DSNSFs e leituras reais do tráfego na extração de informações evidenciais em uma formalização matemática de incertezas referentes ao reconhecimento e classificação de irregularidades em níveis de tráfego. A combinação DSNSF e Lógica Paraconsistente permitiu que uma ferramenta denominada Máquina Paraconsistente Correlacional (MPC) fosse desenvolvida e utilizada como um recurso que auxilia administradores e gerentes de redes na identificação de anomalias. A MPC também fornece um panorama das incertezas e contradições obtidas quando diferentes modelos são utilizados na detecção de anomalias. Esse recurso permite identificar possíveis falhas ou inconsistências nos modelos utilizados.

A MPC foi configurada para identificar três tipos de anomalias: *DoS*, *DDoS* e *Flash Crowd.* Cada anomalia apresenta um comportamento característico e foi identificada por meio da avaliação dos distúrbios que ela causa em determinados atributos.

Outro aspecto apresentado pela MPC na detecção de anomalias foi a classificação de intervalos com base em algumas propriedades de uma extensão da LP chamada Lógica Paraconsistente Anotada de anotação com dois valores (LPA2v). A interpretação do

resultado da interpolação entre níveis de certeza e contradição extraídos das assinaturas digitais e relativos a classificação de um intervalo permitiu que novos estados lógicos fossem analisados. A introdução de estados lógicos como “Inconsistente”, com elevado nível de contradição, e “Paracompleto”, quando as informações são insuficientes, permitiram que classificações do tipo “Verdadeiro” só fossem assumidas quando todas as fontes de informações concordassem entre si. Além disso, os reticulados combinados com os níveis de evidência reais fornecem uma perspectiva contínua do grau de vulnerabilidade da rede.

O sistema proposto foi avaliado em dois ambientes de rede, e apresentou bons resultados nas tarefas de caracterização de tráfego e detecção de anomalias. Os resultados obtidos com os testes de coeficientes de correlação e NMSEs nos dois ambientes mostraram que a metodologia adotada para previsão com o modelo ARIMA é eficaz na caracterização do comportamento normal do tráfego das redes. A correlação obteve uma média de 0,78 no pior cenário testado e de 0,90 no melhor cenário testado. Os erros NMSEs obtidos demonstraram que atributos de volume estão mais vulneráveis em ambiente com menos tráfego agregado. Ainda assim uma média de erros de 0,13 foi obtida no pior cenário testado e de 0,02 no melhor deles.

Com relação a detecção de anomalias, o sistema consegue atingir taxas de Verdadeiros Positivos de aproximadamente 96,5% com taxas de Falsos Positivos de 5%. A Máquina Paraconsistente Correlacional (MPC) teve um papel fundamental no bom desempenho do sistema. Os níveis de acurácia da MPC obtiveram ganho de 1,72 pontos percentuais, em média, em relação ao modelo ARIMA utilizando limites de confiança.

Outra característica do sistema proposto é a possibilidade de utilização de diferentes modelos como fontes de informação para a MPC. Em um dos cenários avaliados a MPC foi alimentada com assinaturas digitais geradas com dois modelos: ARIMA e ACODS. A MPC permitiu uma redução na taxa de falsos positivos e aumento da acurácia em relação aos resultados obtidos com o emprego dos modelos de maneira tradicional utilizando limites de confiança. Assim, a combinação de diferentes fontes de informações e análise paraconsitente permitiu que maior precisão na identificação de anomalias fosse alcançada.

De maneira geral, destaca-se como principais contribuições deste trabalho a especificação e implementação de um sistema para detecção de anomalias utilizando análise de Fluxos IP, aplicação de diferentes atributos de tráfego para caracterizar e detectar anomalias, adaptação do modelo ARIMA para geração das Assinaturas Digitais de Segmento de Rede utilizando análise de Fluxos e a aplicação da Lógica Paraconsistente para detecção de anomalias.

Dessa forma, este trabalho alcança seus objetivos e contribui com um sistema que auxilia administradores de rede em suas atividades cotidianas como o monitoramento do tráfego de rede, a identificação e prevenção de anomalias. Assim, a implantação do sistema proposto pode contribuir diretamente para a disponibilidade e qualidade de serviços prestados pelas redes de computadores.

Os trabalhos futuros incluem: a adaptação do sistema descrito em cenários com características diferentes dos testados neste trabalho; aprimorar o uso da correlação dos atributos de tráfego na composição das assinaturas das anomalias; investigar o impacto da amostragem no processo de detecção; utilizar outros modelos como fontes de informações da MPC.

REFERÊNCIAS

[1] HAMDI, M.; BOUDRIGA, N. Computer and network security risk management: theory, challenges, and countermeasures: Research Articles. **Int. J. Commun. Syst.,** v. 18, n. 8, p. 763-793, 2005. ISSN 1074-5351.

[2] PLONKA, D.; BARFORD, P. **Network anomaly confirmation, diagnosis and remediation**. Proceedings of the 47th annual Allerton conference on Communication, control, and computing. Monticello, Illinois, USA: IEEE Press**:** 128-135 p. 2009.

[3] TAVALLAEE, M.; STAKHANOVA, N.; GHORBANI, A. A. Toward credible evaluation of anomaly-based intrusion-detection methods. **Trans. Sys. Man Cyber Part C,** v. 40, n. 5, p. 516-524, 2010. ISSN 1094-6977.

[4] CATANIA, C. A.; GARINO, C. G. Automatic network intrusion detection: Current techniques and open issues. **Computers & Electrical Engineering,** v. 38, n. 5, p. 1062-1072, 9// 2012. ISSN 0045-7906. Disponível em: < http://www.sciencedirect.com/science/article/pii/S0045790612001073 >.

[5] PATCHA, A.; PARK, J.-M. An overview of anomaly detection techniques: Existing solutions and latest technological trends. **Comput. Netw.,** v. 51, n. 12, p. 3448-3470, 2007. ISSN 1389-1286.

[6] STAKHANOVA, N.; BASU, S.; WONG, J. On the symbiosis of specification-based and anomaly-based detection. **Computers & Security,** v. 29, n. 2, p. 253-268, 3// 2010. ISSN 0167-4048. Disponível em: < http://www.sciencedirect.com/science/article/pii/S0167404809000893 >.

[7] ESTEVEZ-TAPIADOR, J. M.; GARCIA-TEODORO, P.; DIAZ-VERDEJO, J. E. Anomaly detection methods in wired networks: a survey and taxonomy. **Comput. Commun.,** v. 27, n. 16, p. 1569-1584, 2004. ISSN 0140-3664.

[8] YAMANISHI, K.; TAKEUCHI, J.-I.; WILLIAMS, G.; MILNE, P. On-Line Unsupervised Outlier Detection Using Finite Mixtures with Discounting Learning Algorithms. **Data Min. Knowl. Discov.,** v. 8, n. 3, p. 275-300, 2004. ISSN 1384-5810.

[9] BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time Series Analysis: Forecasting and Control**. Wiley, 2008. ISBN 9780470272848. Disponível em: < http://books.google.com.br/books?id=lJnnPQAACAAJ >.

[10] ABE, J. M. **Paraconsistent logics and applications**. Soft Computing Applications (SOFA), 2010 4th International Workshop on. 15-17 July 2010, 2010. 11-18 p.

[11] WASSERKRUG, S.; GAL, A.; ETZION, O. Inference of Security Hazards from Event Composition Based on Incomplete or Uncertain Information. **IEEE Trans. on Knowl. and Data Eng.,** v. 20, n. 8, p. 1111-1114, 2008. ISSN 1041-4347.

[12] MENDEL, J. M. Uncertainty, fuzzy logic, and signal processing. **Signal Processing,** v. 80, n. 6, p. 913-933, 6// 2000. ISSN 0165-1684. Disponível em: < http://www.sciencedirect.com/science/article/pii/S0165168400000116 >.

[13] UDUPA, D. K. **Network Management Systems Essentials**. McGraw-Hill, Inc., 1995. 557 ISBN 0070657661.

[14] KERSHENBAUM, A.; MALEK, M.; WALL, M. **Network Management and Control**. Perseus Publishing, 1990. 448 ISBN 030643587X.

[15] SUBRAMANIAN, M. **Network Management: Principles and Practice**. Dorling Kindersley, 2010. ISBN 9788131727591. Disponível em: < http://books.google.com.br/books?id=VGDMIIfL6XcC >.

[16] TODINOV, M. T. 1 - Flow Networks – Existing Analysis Approaches and Limitations. In: TODINOV, M. T. (Ed.). **Flow Networks**. Oxford: Elsevier, 2013. p.1-12. ISBN 978-0-12-398396-1.

[17] PAUL, S.; PAN, J.; JAIN, R. Architectures for the future networks and the next generation Internet: A survey. **Comput. Commun.,** v. 34, n. 1, p. 2-42, 2011. ISSN 0140-3664.

[18] IETF. CLAISE, B. RFC5101 - Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information., 2008. Disponível em: < https://tools.ietf.org/html/rfc5101 >.

[19] KERR, D. R.; BRUINS, B. L. **Network flow switching and flow data export**: Google Patents 2009.

[20] CISCO. Introduction to Cisco IOS NetFlow - A Technical Overview. 2012. Disponível em: < http://www.cisco.com/c/en/us/products/collateral/ios-nx-os-software/ios-netflow/prod_white_paper0900aecd80406232.html >.

[21] PHAAL, P.; PANCHEN, S. InMon Corporation's sFlow: A Method for Monitoring Traffic in Switched and Routed Networks. 2001. Disponível em: < http://tools.ietf.org/search/rfc3176 >.

[22] PRAS, A. et al. Using NetFlow/IPFIX for Network Management. **J. Netw. Syst. Manage.,** v. 17, n. 4, p. 482-487, 2009. ISSN 1064-7570.

[23] LI, B.; SPRINGER, J.; BEBIS, G.; HADI GUNES, M. A survey of network flow applications. **Journal of Network and Computer Applications,** v. 36, n. 2, p. 567-581, 3// 2013. ISSN 1084-8045. Disponível em: < http://www.sciencedirect.com/science/article/pii/S1084804512002676 >.

[24] REESE, B. Cisco's NetFlow vs. Inmon's sFlow: Which will prevail? , 2007. Disponível em: < http://www.networkworld.com/community/node/22667 >.

[25] RASPALL, F. Efficient packet sampling for accurate traffic measurements. **Comput. Netw.,** v. 56, n. 6, p. 1667-1684, 2012. ISSN 1389-1286.

[26] MAI, J.; CHUAH, C.-N.; SRIDHARAN, A.; YE, T.; ZANG, H. **Is sampled data sufficient for anomaly detection?** Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. Rio de Janeriro, Brazil: ACM**:** 165-176 p. 2006.

[27] BRAUCKHOFF, D.; TELLENBACH, B.; WAGNER, A.; MAY, M.; LAKHINA, A. **Impact of packet sampling on anomaly detection metrics**. Proceedings of the 6th ACM SIGCOMM conference on Internet measurement. Rio de Janeriro, Brazil: ACM**:** 159-164 p. 2006.

[28] BARTOS, K.; REHAK, M.; KRMICEK, V. **Optimizing flow sampling for network anomaly detection**. Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International. 4-8 July 2011, 2011. 1304-1309 p.

[29] RICCIATO, F.; STROHMEIER, F.; DORFINGER, P.; COLUCCIA, A. **One-way loss measurements from IPFIX records**. Measurements and Networking Proceedings (M&N), 2011 IEEE International Workshop on. 10-11 Oct. 2011, 2011. 158-163 p.

[30] KUNDU, S. R.; PAL, S.; BASU, K.; DAS, S. K. An Architectural Framework for Accurate Characterization of Network Traffic. **IEEE Trans. Parallel Distrib. Syst.,** v. 20, n. 1, p. 111-123, 2009. ISSN 1045-9219.

[31] YUANYUAN, L. **Study of the monitoring model for securities trading network Quality of Service**. Information Science and Engineering (ICISE), 2010 2nd International Conference on. 4-6 Dec. 2010, 2010. 1-4 p.

[32] LIU, D.; HUEBNER, F. **Application profiling of IP traffic**. Local Computer Networks, 2002. Proceedings. LCN 2002. 27th Annual IEEE Conference on. 6-8 Nov. 2002, 2002. 220-229 p.

[33] KALAFUT, A. J.; VAN DER MERWE, J.; GUPTA, M. **Communities of Interest for Internet Traffic Prioritization**. INFOCOM Workshops 2009, IEEE. 19-25 April 2009, 2009. 1-6 p.

[34] CARACAS, A.; KIND, A.; GANTENBEIN, D.; FUSSENEGGER, S.; DECHOUNIOTIS, D. **Mining semantic relations using NetFlow**. Business-driven IT Management, 2008. BDIM 2008. 3rd IEEE/IFIP International Workshop on. 7-7 April 2008, 2008. 110-111 p.

[35] YINGYING, C.; JAIN, S.; ADHIKARI, V. K.; ZHI-LI, Z.; KUAI, X. **A first look at inter-data center traffic characteristics via Yahoo! datasets**. INFOCOM, 2011 Proceedings IEEE. 10-15 April 2011, 2011. 1620-1628 p.

[36] LAKHINA, A.; CROVELLA, M.; DIOT, C. Diagnosing network-wide traffic anomalies. **SIGCOMM Comput. Commun. Rev.,** v. 34, n. 4, p. 219-230, 2004. ISSN 0146-4833.

[37] CAMPBELL, P. L. The denial-of-service dance. **Security & Privacy, IEEE,** v. 3, n. 6, p. 34-40, 2005. ISSN 1540-7993.

[38] JANG-JACCARD, J.; NEPAL, S. A survey of emerging threats in cybersecurity. **Journal of Computer and System Sciences,** v. 80, n. 5, p. 973-993, 8// 2014. ISSN 0022-0000. Disponível em: < http://www.sciencedirect.com/science/article/pii/S0022000014000178 >.

[39] RINGBERG, H.; SOULE, A.; REXFORD, J.; DIOT, C. Sensitivity of PCA for traffic anomaly detection. **SIGMETRICS Perform. Eval. Rev.,** v. 35, n. 1, p. 109-120, 2007. ISSN 0163-5999.

[40] HAJJI, H. **Baselining network traffic and online faults detection**. Communications, 2003. ICC '03. IEEE International Conference on. 11-15 May 2003, 2003. 301-308 vol.1 p.

[41] THOTTAN, M.; CHUANYI, J. Anomaly detection in IP networks. **Signal Processing, IEEE Transactions on,** v. 51, n. 8, p. 2191-2204, 2003. ISSN 1053-587X.

[42] JIANG, J.; PAPAVASSILIOU, S. Detecting Network Attacks in the Internet via Statistical Network Traffic Normality Prediction. **J. Netw. Syst. Manage.,** v. 12, n. 1, p. 51-72, 2004. ISSN 1064-7570.

[43] CHENGWEI, W. et al. **Statistical techniques for online anomaly detection in data centers**. Integrated Network Management (IM), 2011 IFIP/IEEE International Symposium on. 23-27 May 2011, 2011. 385-392 p.

[44] QIN, T.; GUAN, X.; LI, W.; WANG, P.; HUANG, Q. Monitoring abnormal network traffic based on blind source separation approach. **J. Netw. Comput. Appl.,** v. 34, n. 5, p. 1732-1742, 2011. ISSN 1084-8045.

[45] HAU, M. C.; TONG, H. A practical method for outlier detection in autoregressive time series modelling. **Stochastic Environmental Research and Risk Assessment (SERRA),** v. 3, n. 4, p. 241-260, // 1989.

[46] FOX, A. J. Outliers in time series. **Journal of the Royal Statistical Society,** v. 34, n. Ser. B, p. 350-363, 1972.

[47] YAACOB, A. H.; TAN, I. K. T.; SU FONG, C.; HON KHI, T. **ARIMA Based Network Anomaly Detection**. Communication Software and Networks, 2010. ICCSN '10. Second International Conference on. 26-28 Feb. 2010, 2010. 205-209 p.

[48] ZHU, B.; SASTRY, S. **Revisit Dynamic ARIMA Based Anomaly Detection**. Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom). 9-11 Oct. 2011, 2011. 1263-1268 p.

[49] DE ASSIS, M. V. O.; CARVALHO, L. F.; RODRIGUES, J. J. P. C.; PROENÇA, M. L., JR. **Holt-Winters statistical forecasting and ACO metaheuristic for traffic characterization**. Communications (ICC), 2013 IEEE International Conference on. 9-13 June 2013, 2013. 2524-2528 p.

[50] DE ASSIS, M. V. O.; RODRIGUES, J. J. P. C.; PROENÇA JR, M. L. A seven-dimensional flow analysis to help autonomous network management. **Information Sciences,** v. 278, n. 0, p. 900-913, 9/10/ 2014. ISSN 0020-0255. Disponível em: < http://www.sciencedirect.com/science/article/pii/S0020025514003995 >.

[51] DAINOTTI, A.; PESCAP, A.; #233; VENTRE, G. A cascade architecture for DoS attacks detection based on the wavelet transform. **J. Comput. Secur.,** v. 17, n. 6, p. 945-968, 2009. ISSN 0926-227X.

[52] YASAMI, Y.; KHORSANDI, S.; MOZAFFARI, S. P.; JALALIAN, A. **An unsupervised network anomaly detection approach by k-Means clustering & ID3 algorithms**. Computers and Communications, 2008. ISCC 2008. IEEE Symposium on. 6-9 July 2008, 2008. 398-403 p.

[53] ENCHEVA, S.; TUMIN, S.; KONDRATENKO, Y. **Application of paraconsistent annotated logic in intelligent systems**. Proceedings of the intelligent computing 3rd international conference on Advanced intelligent computing theories and applications. Qingdao, China: Springer-Verlag**:** 702-710 p. 2007.

[54] TORRES, C. R.; LAMBERT-TORRES, G.; BORGES DA SILVA, L. E.; ABE, J. M. **Intelligent System of Paraconsistent Logic to Control Autonomous Moving Robots**. IEEE Industrial Electronics, IECON 2006 - 32nd Annual Conference on. 6-10 Nov. 2006, 2006. 4009-4013 p.

[55] SILVA FILHO, J. I.; HURTADO, G. L.; GARCIA, D. Analysis and Diagnosis of Cardiovascular Diseases through the Paraconsistent Annotated Logic. In: (Ed.). **New Advances in Intelligent Decision Technologies**: Springer Berlin Heidelberg, v.199, 2009. cap. 29, p.295-303. (Studies in Computational Intelligence). ISBN 978-3-642-00908-2.

[56] GUIDO, R. C. et al. Introducing the Discriminative Paraconsistent Machine (DPM). **Inf. Sci.,** v. 221, p. 389-402, 2013. ISSN 0020-0255.

[57] DA COSTA, N. C. A.; KRAUSE, D.; BUENO, O. Paraconsistent Logics and Paraconsistency. In: JACQUETTE, D. (Ed.). **Philosophy of Logic**. Amsterdam: North-Holland, 2007. p.791-911. ISBN 18789846.

[58] KAVITHA, B.; KARTHIKEYAN, D. S.; MAYBELL, P. S. An ensemble design of intrusion detection system for handling uncertainty using Neutrosophic Logic Classifier. **Know.-Based Syst.,** v. 28, p. 88-96, 2012. ISSN 0950-7051.

[59] CHATZIGIANNAKIS, V.; ANDROULIDAKIS, G.; PELECHRINIS, K.; PAPAVASSILIOU, S.; MAGLARIS, V. **Data fusion algorithms for network anomaly detection: classification and evaluation**. Proceedings of the Third International Conference on Networking and Services: IEEE Computer Society**:** 50 p. 2007.

[60] PROENÇA, M. L., JR.; ZARPELAO, B. B.; MENDES, L. S. **Anomaly Detection for Network Servers Using Digital Signature of Network Segment**. Advanced Industrial Conference on Telecommunications: IEEE Computer Society**:** 290-295 p. 2005.

[61] PROENÇA, M. L., JR. **Baseline Aplicado a Gerência de Redes.** 2005. Tese de doutorado Faculdade de Engenharia Elétrica e de Computação (FEEC), Universidade Estadual de Campinas (UNICAMP), Campinas.

[62] FERNANDES, G.; ZACARON, A. M.; RODRIGUES, J. J. P. C.; PROENÇA, M. L., JR. **Digital signature to help network management using principal component analysis and K-means clustering**. Communications (ICC), 2013 IEEE International Conference on. 9-13 June 2013, 2013. 2519-2523 p.

[63] CARVALHO, L. F.; RODRIGUES, J. J. P. C.; BARBON, S.; LEMES PROENCA, M. **Using Ant Colony Optimization metaheuristic and Dynamic Time Warping for anomaly detection**. Software, Telecommunications and Computer Networks (SoftCOM), 2013 21st International Conference on. 18-20 Sept. 2013, 2013. 1-5 p.

[64] SHANNON, C. E. A mathematical theory of communication. **SIGMOBILE Mob. Comput. Commun. Rev.,** v. 5, n. 1, p. 3-55, 2001. ISSN 1559-1662.

[65] JANACEK, G. J. **Practical Time Series**. Arnold, 2001. ISBN 9780340719992. Disponível em: < http://books.google.com.br/books?id=44gafXIFTMgC >.

[66] BROCKWELL, P. J.; DAVIS, R. A. **Introduction to Time Series and Forecasting**. Springer, 2002. ISBN 9780387953519. Disponível em: < http://books.google.com.br/books?id=VHB4OSAmwcUC >.

[67] MAKRIDAKIS, S.; HIBON, M. ARMA Models and the Box–Jenkins Methodology. **Journal of Forecasting,** v. 16, n. 3, p. 147-163, 1997. ISSN 1099-131X. Disponível em: < http://dx.doi.org/10.1002/(SICI)1099-131X(199705)16:3<147::AID-FOR652>3.0.CO;2-X >.

[68] HÖGLUND, R.; ÖSTERMARK, R. Automatic arima modelling by the cartesian search algorithm. **Journal of Forecasting,** v. 10, n. 5, p. 465-476, 1991. ISSN 1099-131X. Disponível em: < http://dx.doi.org/10.1002/for.3980100503 >.

[69] JENKINS, G. M.; WATTS, D. G. **Spectral analysis and its applications**. Holden-Day, 1969. Disponível em: < http://books.google.com.br/books?id=1kBMAAAAMAAJ >.

[70] FILHO, J. I. D. S.; LAMBERT-TORRES, G.; ABE, J. M. **Uncertainty treatment using paraconsistent logic: introducing paraconsistent artificial neural networks**. IOS Press, 2010. 308 ISBN 978-1-60750-557-0.

[71] STONE, J. V. **Bayes' Rule: A Tutorial Introduction to Bayesian Analysis**. Sebtel Press, 2013. 180 ISBN 0956372848, 9780956372840.

[72] HAWTHORNE, J. Confirmation Theory. In: BANDYOPADHYAY, P. S. e FORSTER, M. R. (Ed.). **Philosophy of Statistics**. Amsterdam: North-Holland, v.7, 2011. p.333-389. ISBN 18789846.

[73] LEPSKIY, A. Estimation of Conflict and Decreasing of Ignorance in Dempster-Shafer Theory. **Procedia Computer Science,** v. 17, n. 0, p. 1113-1120, // 2013. ISSN 1877-0509. Disponível em: < http://www.sciencedirect.com/science/article/pii/S1877050913002755 >.

[74] ABE, J. M.; FILHO, J. I. D. S.; PRADO, J. C. A. **Alguns aspectis práticos da lógica paraconsistente aplicados à rede neural artificial**. III Congresso Brasileiro de Computação – CBComp 2003. Itajaí (SC) 2003. p.

[75] SHUYING, C.; XUESONG, Q.; ZHIPENG, G.; KE, L.; FENG, Q. **A flow-based anomaly detection method using sketch and combinations of traffic features**.

Network and Service Management (CNSM), 2010 International Conference on. 25-29 Oct. 2010, 2010. 302-305 p.

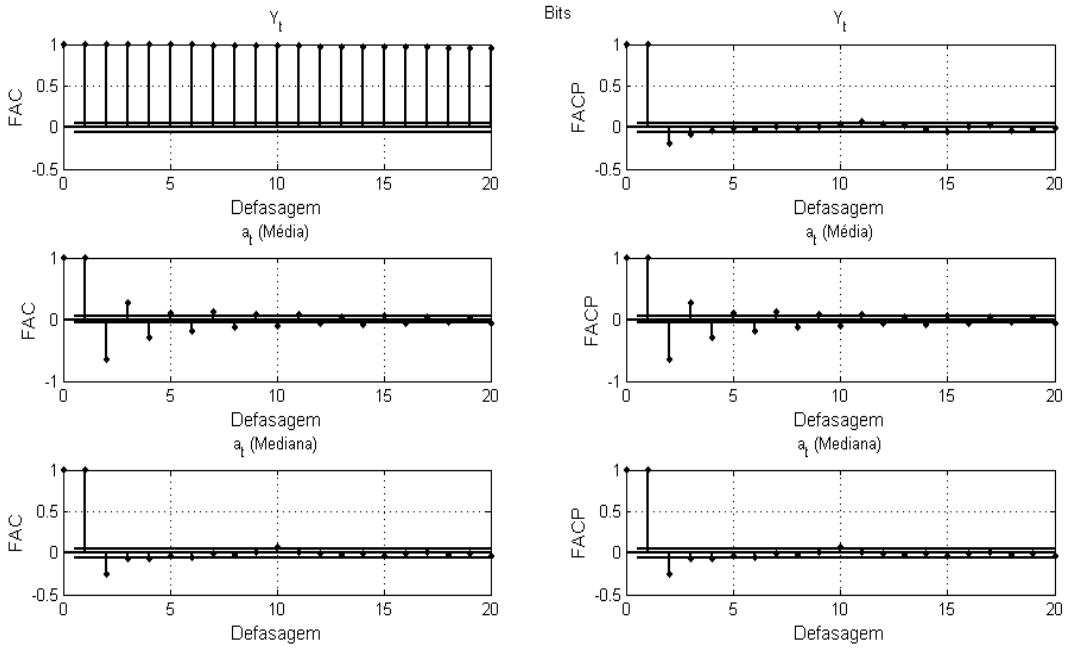
[76] ASSIS, M. V. O. sFlow Anomaly Simulator., 2013. Disponível em: < http://redes.dc.uel.br/scorpius/ >. Acesso em: 27/06/2014.

[77] BRUTLAG, J. D. **Aberrant Behavior Detection in Time Series for Network Monitoring**. Proceedings of the 14th USENIX conference on System administration. New Orleans, Louisiana: USENIX Association**:** 139-146 p. 2000.

APÊNDICES

APÊNDICE A - EXEMPLOS DE MODELOS UTILIZADOS

Atributo *Bits*

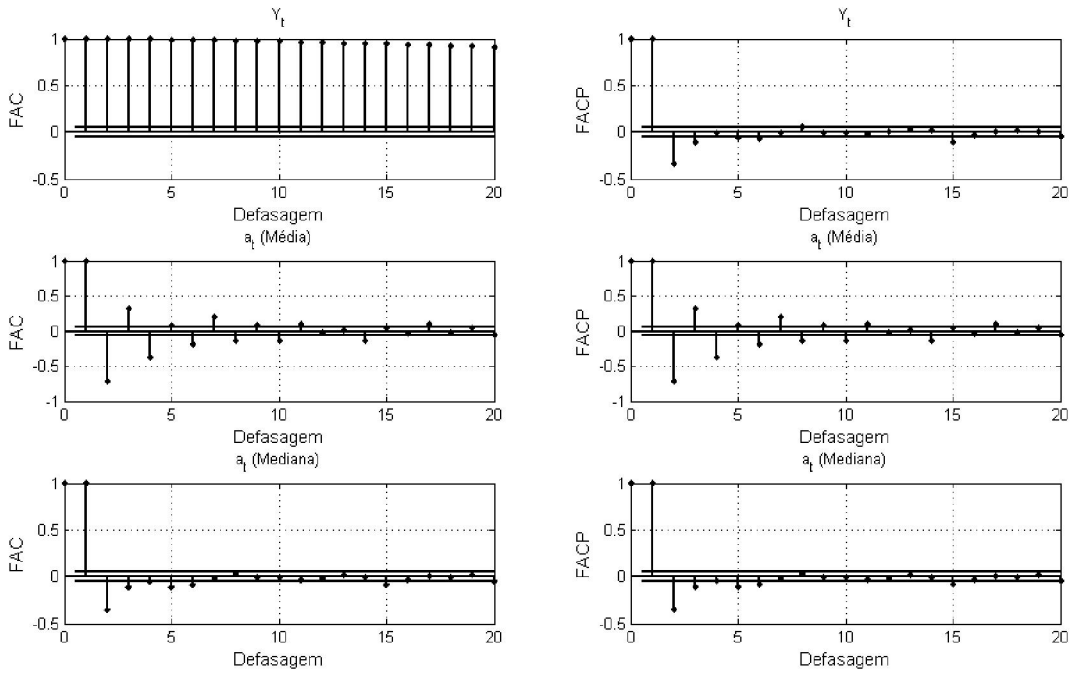


Correlogramas de exemplo para séries temporais do atributo Bits.

Modelo M1(a_t: Média) estimado: $\Delta Y_t = 0{,}2395\Delta Y_{t-1} + a_t - 0{,}1753\Delta a_{t-1}$

Modelo M2 (a_t: Mediana) estimado:$\Delta Y_t = 0{,}1095\Delta Y_{t-1} + 0{,}9592\Delta a_{t-1}$

Atributo Pacotes

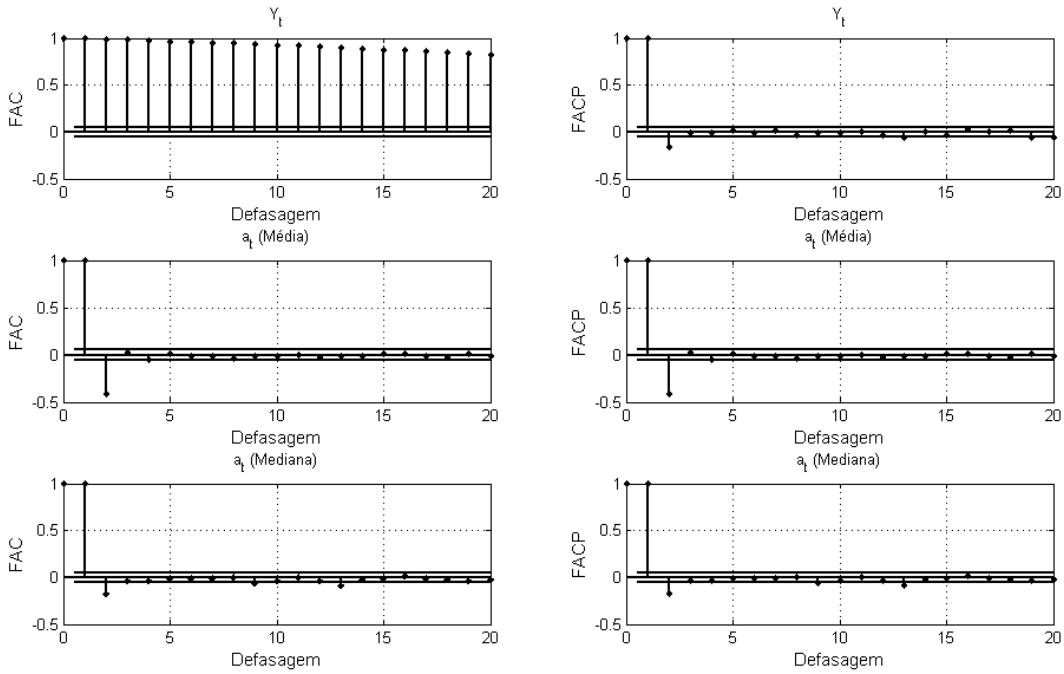


Correlogramas de exemplo para séries do atributo Pacotes.

Modelo M1(a_t: Média) estimado: $\Delta Y_t = 0{,}4161\Delta Y_{t-1} + a_t - 0{,}05020a_{t-1}$

Modelo M2 (a_t: Mediana) estimado:$\Delta Y_t = 0{,}5729\Delta Y_{t-1} + 0{,}4920\Delta a_{t-1}$

Atributo H(dstIP)

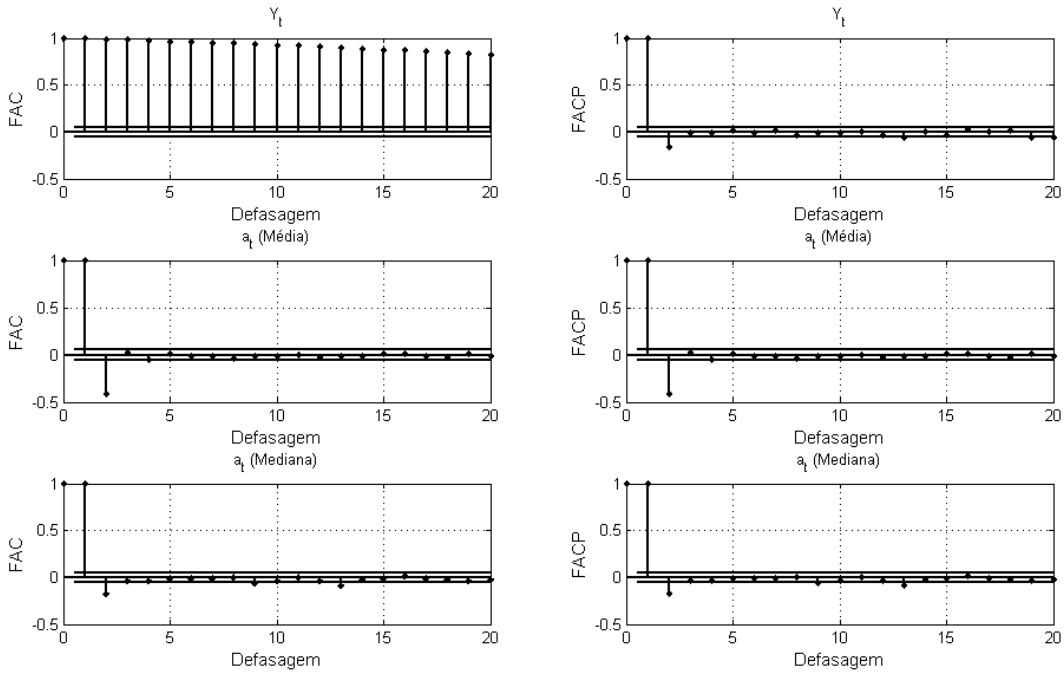


Correlogramas de exemplo para séries temporais do atributo H(dstIP).

Modelo M1(a_t: Média) estimado: $\Delta Y_t = 0{,}0148\Delta Y_{t-1} + a_t$

Modelo M2 (a_t: Mediana) estimado:$\Delta Y_t = 0{,}0130\Delta Y_{t-1} + a_t - 0{,}0400a_{t-1}$

Atributo H(orgIP)

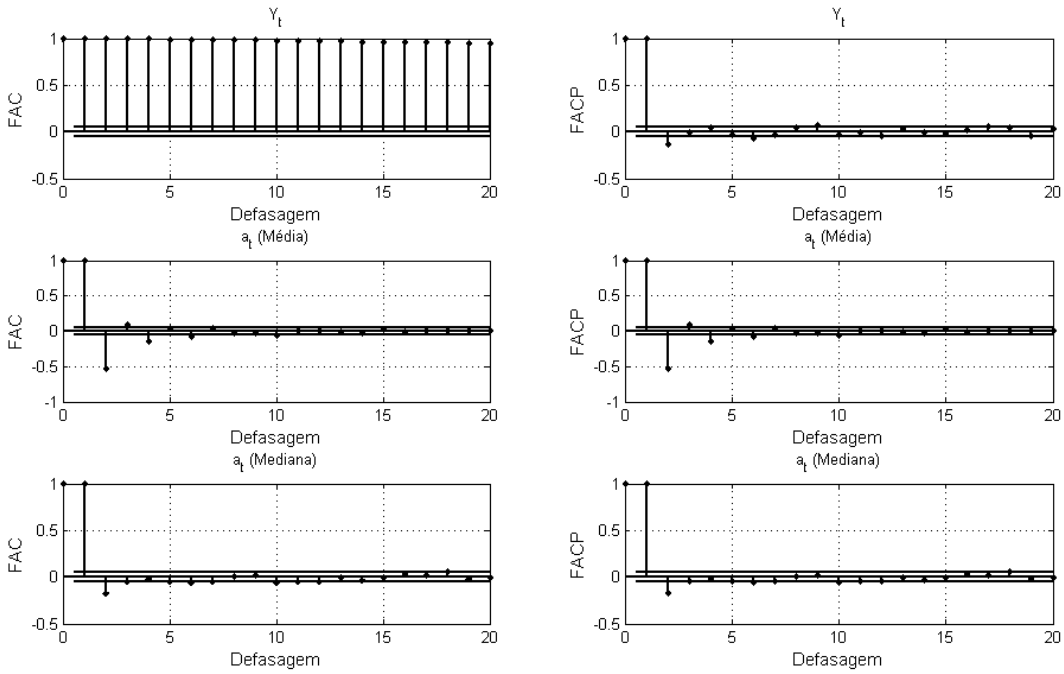


Correlogramas de exemplo para séries temporais do atributo H(orgIP).

Modelo M1(a_t: Média) estimado: $\Delta Y_t = -0.0175\Delta Y_{t-1} + a_t$

Modelo M2 (a_t: Mediana) estimado:$\Delta Y_t = 0.0731\Delta Y_{t-1}$+0,8743$a_{t-1}$

Atributo H(dstP)

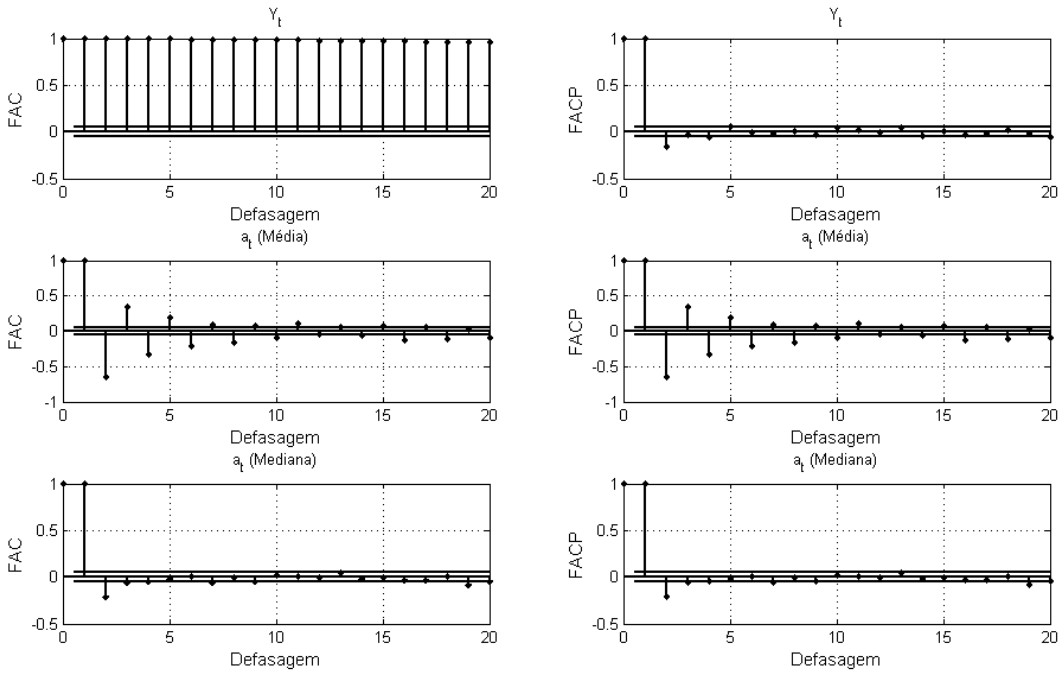


Correlogramas de exemplo para séries temporais do atributo H(dstP).

Modelo M1(a_t: Média) estimado: $\Delta Y_t = -0.7254\Delta Y_{t-1} + a_t + 0{,}8743a_{t-1}$

Modelo M2 (a_t: Mediana) estimado: $\Delta Y_t = -0.5002\Delta Y_{t-1} + a_t + 0{,}4023a_{t-1}$

Atributo H(orgP)



Correlogramas de exemplo para séries temporais do atributo H(orgP).

Modelo M1(a_t: Média) estimado: $\Delta Y_t = -0.0325\Delta Y_{t-1} + a_t$

Modelo M2 (a_t: Mediana) estimado:$\Delta Y_t = -0.2000\Delta Y_{t-1} + a_t + 0{,}1320a_{t-1}$

APÊNDICE B – ANT COLONY OPTIMIZATION FOR DIGITAL SIGNATURE

A metaheurística Ant Colony Optimization (ACO) foi apresentada por Marco Dorigo no início da década de 90. Tem como base os princípios dos Ant Algorithms, os quais são definidos como uma coleção de métodos computacionais fundamentados em observações de comportamentos das colônias de formigas, projetados para solucionar problemas complexos de otimização. Essa metaheurística pertence à categoria de métodos denominadas de inteligência de grupo, e é caracterizada pela auto-organização de agentes para realização de uma determinada tarefa. Como as demais metaheurísticas, o ACO é um conjunto de estratégias e algoritmos baseados em conjuntos de informações exclusivas para resolução do problema.

A criação da assinatura digital, ou DSNSF, utilizando a metaheurística Ant Colony Optimization apresenta uma série de modificações em relação ao modelo tradicional e visa otimizar a classificação e extração de padrões de tráfego propiciadas pelo processo clusterização. Tal procedimento é denominado Ant Colony Optimization for Digital Signature (ACODS) e apresenta propriedades de aprendizado não supervisionado, o que garante que os agentes possam avaliar as soluções ao longo da execução, encontrando aquelas que levam a resultados mais promissores e evitam ótimos locais.

Para a construção da assinatura digital de segmento de rede é utilizada a busca de padrões de comportamentos por meio da técnica de clusterização. Tal abordagem agrupa elementos que apresentam características semelhantes em grupos específicos. O principal objetivo é aumentar a similaridade entre elementos de um grupo enquanto a dissimilaridade entre grupos é maximizada. Um elemento da clusterização é uma 6-upla contendo os atributos do tráfego analisado.

O ACODS trabalho tem como objetivo otimizar a eficiência da clusterização, minimizando o valor da função objetivo, definida pela seguinte equação.

$$J = \sum_{i=1}^{E} \sum_{j=1}^{K} \sqrt{\sum_{a=1}^{A} \left(x_{ia} - c_{ja}\right)^2}$$

Em outras palavras, pretende-se criar soluções para o agrupamento de dados de uma maneira que permita a extração de padrões, comportamentos e características do tráfego. Assim, essa abordagem garante que cada elemento será agrupado ao melhor cluster. Além disso, o ACODS permite a construção de soluções que não são dadas por ótimos locais, que é um problema recorrente em muitos algoritmos de clusterização.

A função objetivo descrita é utilizada para mensurar quão boa é a solução criada por um agente. O valor correspondente ao seu resultado é a distância entre cada elemento em relação ao centro do cluster ao qual pertence. Quanto menor o valor de J, melhor é a solução. Nessa função, a variável E corresponde ao total de elementos a serem clusterizados e A indica a dimensionalidade dos dados, ou seja, a quantidade de atributos de fluxos que cada elemento contém. Para a construção do DSNSF os elementos adquiridos após o pré-processamento do tráfego são avaliados a cada intervalo de dez segundos, totalizando 8640 conjuntos de resultados ao longo do dia. A variável x_{ia} denota o valor do atributo a do elemento i, enquanto c_{ja} armazena o valor do centro do cluster j na dimensão a.

O algoritmo a seguir apresenta os procedimentos para criação do DSNSF.

Procedimento ACODS usado para criação do DSNSF

Entrada: Conjunto de atributos de volume e dos campos do cabeçalho do pacote coletados a partir de uma base de dados histórica, número de clusters, número máximo de iterações *I*.
Saída: Matriz compreendendo o DSNSF, na qual as linhas representam os intervalos de tempo (10 segundos) enquanto as colunas contêm os valores calculados para cada atributo no período de tempo correspondente.

1: **PARA** t = 1 **ATÉ** 8640 **FAÇA**
2: **PARA** *i* = 1 **ATÉ** ***I*** **OU** condição de parada não for satisfeita **FAÇA**
3: Criar solução
4: Avaliar solução utilizando a função objetivo
5: Atualizar a trilha de feromônio
6: **FIM-PARA**
7: Calcular o centro de cada cluster da melhor solução encontrada
8: **PARA** j = 1 **até** *K*
9: **SE** total de elementos no cluster $c_j < \gamma$ **ENTÃO**
10: descartar o cluster c_j
11: **FIM-SE**
12: **FIM-PARA**
13: **PARA** $\boldsymbol{a = 1}$ **até** ***A***
13: $X_{ta} \leftarrow$ média ponderada entre os clusters na dimensão a
14: **FIM-PARA**
15: **Retornar** *X*

De modo geral, as principais funções executadas pelo ACODS são a construção de soluções, a busca local e a atualização de feromônio.

A etapa de construção de soluções consiste no movimento simultâneo das formigas e de forma assíncrona por parte dos estados do problema.

A busca local destina-se a testar e avaliar soluções criadas pelas formigas através de uma busca local. Se a informação heurística sobre um domínio de problema particular não é facilmente alcançada, a busca local pode ajudar a construir boas soluções.

A atualização do feromônio é o processo em que a trilha de feromônio é modificada. O aumento da concentração de feromônio é um fator essencial para a implementação do algoritmo, uma vez que direciona as formigas a buscar novos locais mais propensos a adquirir uma solução ideal.

Após a clusterização, os grupos de dados são formados. Devido aos padrões de comportamento do tráfego de rede, muitas informações apresentam características semelhantes. Assim, os clusters formados por pequenas quantidades de dados que divergem muito do padrão devem ser rejeitados da construção da assinatura. Para tanto, é definido um limite γ, o qual determina a proporção mínima permitida de elementos agrupados em um cluster. Se algum grupo apresentar menos elementos associados a si do que o estipulado por γ, este é retirado da solução final, bem como elementos pertencentes a ele. Esta etapa compõe a estratégia de busca local, o que garante o mínimo envolvimento possível de tráfego anômalo na composição da assinatura.

O resultado desse algoritmo descreve a combinação dos grupos mais representativos do ambiente clusterizado. Para obter o valor de X_{ta}, a média ponderada é calculada entre os grupos no instante t, tal como mostrado na seguinte Equação.

$$X_{ta} = \frac{\sum_{j=1}^{K} c_{ja} p_j}{E}$$

Assim, o resultado estará mais próximo do cluster com o maior número de elementos, ou seja, daquele que melhor representar o comportamento dos dados coletados em intervalos de dez segundos. O termo c_{ja} é o valor central do cluster j na dimensão a e p_j é a quantidade de elementos associados ao conjunto de mesmo índice. Ao final da execução do algoritmo, uma matriz X de dimensões 8640×6 é retornada, ou seja, o DSNSF.

APÊNDICE C - LIMITES DE CONFIANÇA

Os limites de confiança indicam variações ou desvios entre previsão e observação considerados normais. Para cada perfil gerado para um atributo do DSNSF é calculado um limite superior e inferior que é considerado aceitável. Como as variáveis são continuas esse intervalo torna possível a aplicação de comparações.

Na metodologia adotada nesse trabalho, algumas assinaturas caracterizam o comportamento de determinadas anomalias. Por exemplo, uma anomalia *DDoS* afeta os atributos “Pacotes”, “H(dstIP)”, “H(dstP)” E H(orgP). Assim a comparação entre leituras reais de tráfego e perfis encontrados no DSNSF é feita apenas para esses atributos na identificação dessa anomalia. Caso o comportamento exceda os limites de confiança estipulados para cada perfil em todos os atributos descritos uma anomalia é aferida. Esse processo é feito da mesma maneira para os outros dois tipos de anomalia estudadas.

O cálculo dos limites de confiança (também chamados de *thresholds*) são estimados com as seguintes Equações:

$$LC_{superior} = (\alpha + 1)\bar{u}_{t-1}$$
$$LC_{inferior} = (\alpha + 1)\bar{u}_{t-1}$$

O parâmetro α especifica um intervalo aceitável acima do valor principal considerado. O termo $\bar{u}_{t-1}$ é relativo a um calculo baseado das medidas anteriores ao intervalo analisado, para $t = 1,2,3, \ldots \boldsymbol{T}$. O mesmo é feito com a seguinte Equação:

$$\bar{u}_t = \beta\bar{u}_{t-1} + (1 - \beta)x_t$$

Onde x_t é um valor do atributo do DSNSF no intervalo t é um fator de suavização. Baseado em testes numéricos e na literatura [77] os valores dos parâmetros são $\alpha = 0,2$ e β=0,7. De acordo com os testes conduzidos, outros valores resultaram em piora na taxa de detecção ou não obtiveram nenhuma melhora.

Trabalhos Publicados pelo autor

PENA, E.H.M.; ASSIS, M.V.O.; PROENÇA JR., M.L. Anomaly detection using forecasting methods ARIMA and HWDS. In: XXXII International Conference of the Chilean Computer Science Society (SCCC'2013), 2013, Temuco-CL. Proceedings of XXXII International Conference of the Chilean Computer Science Society (SCCC'2013), 2013. **(Qualis CC 2014 - B3)**

AMARAL, A. A.; MENDES, L. S.; **PENA, E.H.M.**; ZARPELÃO, B.B.; PROENÇA JR., M.L. Network anomaly detection by IP flow graph analysis: A DDoS attack case study. In: XXXII International Conference of the Chilean Computer Science Society (SCCC'2013), 2013, Temuco-CL. Proceedings of XXXII International Conference of the Chilean Computer Science Society (SCCC'2013), 2013. **(Qualis CC 2014 - B3)**

PENA, E.H.M.; BARBON JR., S.; RODRIGUES, J. J. P. C.; PROENÇA JR., M. L. Anomaly detection using digital signature of network segment with adaptive arima model and paraconsistent logic. Em: IEEE Symposium on Computers and Communications (ISCC 2014), 23-26 June, 2014, Madeira-Portugal. IEEE Symposium on Computers and Communications (ISCC 2014), 2014. **(Qualis CC 2014, A2)**

(Aceito para publicação) **PENA, E.H.M.**; CARVALHO, L. F.; BARBON JR., S.; RODRIGUES, J. J. P. C.; PROENÇA JR., M. L. Correlational Paraconsistent Machine for anomaly detection. Em: IEEE Global Communications Conference (GLOBECOM 2014), 8-12 December, 2014, Austin-USA. IEEE Gobal Communications Conference (GLOBECOM 2014), 2014. **(Qualis CC 2014 A1)**