Medições ativas na Internet: algoritmos baseados em retardo fim-a-fim e experimentos

por

Antonio Augusto de Aragão Rocha

UFRJ

Tese submetida para a obtenção do título de

Mestre em Engenharia de Sistemas e Computação

ao Programa de Pós-Graduação de Engenharia de Sistemas e Computação

Coordenação de Pesquisas e Pós-graduação em Engenharia da UFRJ

por

Antonio Augusto de Aragão Rocha

Agosto de 2003

MEDIÇÕES ATIVAS NA INTERNET:
ALGORITMOS BASEADOS EM RETARDO FIM-A-FIM E EXPERIMENTOS

Antonio Augusto de Aragão Rocha

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:

Prof. Edmundo Albuquerque de Souza e Silva, Ph.D.

Prof. Rosa Maria Meri Leão, Dr.

Prof. José Augusto Suruagy Monteiro, Ph.D.

Prof. Otto Carlos Muniz Bandeira Duarte, Dr.Ing.

RIO DE JANEIRO, RJ - BRASIL

AGOSTO DE 2003

ROCHA, ANTONIO AUGUSTO DE ARAGÃO

Medições ativas na Internet: algoritmos baseados em retardo fim-a-fim e experimentos [Rio de Janeiro] 2003

XVIII, 92 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas e Computação, 2003)

Tese - Universidade Federal do Rio de Janeiro, Pesc/COPPE

1. Medição Ativa
2. Métricas Fim-a-fim
3. Sincronização de Relógios
4. Ferramenta TANGRAM-II Traffic Generator

I. COPPE/UFRJ II. Título (Série)

Dedico este trabalho a minha esposa, meus familiares e amigos, mas certamente ele não existiria sem a ajuda de duas pessoas muito especias.

Em memória de Raimundo e Marysses Aragão

Agradecimentos

A conclusão deste trabalho, certamente, não seria possível se não fosse pela ajuda e pelo apoio que recebi ao longo desta minha caminhada. Por isso, gostaria de agradecer a todos que direta ou indiretamente me auxiliaram na conclusão do trabalho.

Agradeço a Deus não só pela força e saúde, mas, principalmente, por ter me cercado de pessoas generosas durante este tempo.

A toda minha família, principalmente aos meus pais e irmãos, pelo amor, carinho e pela compreensão que sempre me proporcionaram. Em especial, aos meus avós Raimundo e Marysses que nunca mediram esforços para possibilitar que eu lutasse por meus objetivos, tudo que sou e que conquistei até hoje é graças a vocês. Seja onde estiverem, muito obrigado por tudo, serei eternamente grato a vocês!

A minha esposa também pelo amor e carinho, em especial pelo apoio e pela dedicação que tem sempre me dado ao longo desses anos.

Aos meus orientadores, Edmundo e Rosa, pela oportunidade de trabalho, pela confiança que depositaram em mim, e por toda sabedoria que me passaram. Ao professor José Augusto Suruagy pela ajuda desde a graduação e pelo interesse em meu trabalho.

Por fim, agradeço a toda família LAND e aos colegas de mestrado que me ajudaram durante todo esse tempo, em especial: Bernardo, Beto, Carol, Bruno, Ana, Guilherme, Allyson, Melba, Kelvin, Drika, Flavio, Daniel, Felipe, Isabela, Hugo, Ed, Sidney, Carolina, Carlo K, Fernando, Saulo, Kleber, Aline, e outros.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

MEDIÇÕES ATIVAS NA INTERNET:

ALGORITMOS BASEADOS EM RETARDO FIM-A-FIM E EXPERIMENTOS

Antonio Augusto de Aragão Rocha

Agosto/2003

Orientadores: Edmundo Albuquerque de Souza e Silva
Rosa Maria Meri Leão

Programa: Engenharia de Sistemas e Computação

A crescente demanda por aplicações multimídia torna o conhecimento das características da rede fundamental para o uso dessas aplicações. Algumas características podem ser facilmente obtidas. Porém, outras também de grande importância são mais difíceis de serem estimadas, como o atraso dos pacotes em um sentido. Sem o uso de dispositivos específicos, o cálculo dessa medida requer um tratamento especial às diferenças existentes entre os relógios envolvidos na medição.

Algoritmos foram propostos para solucionar os problemas e possibilitar a estimativa de métricas baseadas no retardo fim-a-fim de pacotes. Neste trabalho, foram avaliadas e implementadas algumas dessas propostas, e ainda desenvolvidas novas propostas que consideram eventuais atualizações dos relógios e congelamento dos sistemas durante a coleta. Mecanismos foram utilizados para validar as implementações da ferramenta.

Uma série de experimentos foram feitos. As coletas foram estudadas na tentativa de identificar a distribuição que melhor caracterize a medida atraso em um sentido. Assim, modelos de simulação ou analíticos podem ser criados para representar situações reais de perda e atraso de pacotes na rede. A métrica atraso em um sentido é também parâmetro em uma nova proposta para estimar a capacidade de transmissão do gargalo. Essa proposta prevê a seleção dos pares a serem utilizados no cálculo da métrica.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

INTERNET ACTIVE MEASUREMENTS:

ONE WAY DELAY BASED ALGORITHMS AND EXPERIMENTS

Antonio Augusto de Aragão Rocha

August/2002

Advisors: Edmundo Albuquerque de Souza e Silva
Rosa Maria Meri Leão

Department: Computer and System Engineering

With the increasing demand for multimedia applications, to know the internal network characteristics became essential for improving the quality of these applications.Some of those network characteristics can be easily measured. However, other important characteristics are not so easy to be obtained., e.g. one-way delay. To estimate one-way delay, without using special devices, is necessary to consider clocks not synchronized.

Algorithms have been proposed to solve these issues and to allow estimating metrics based on one-way delay of packets. In this work, these algorithms were avaliated and implemented. There was also developed new proposes considering clock atualizations and operational system halt. Mechanisms were used to validate the implemented tool.

Several experiments were applied. The collected results were studied trying to identify the best distribution to characterize the one-way delay. Those estimated distributions may be used to represent losses and delays in simulation and analitical models. The metric one-way delay alse is used as parameter of a new proposal for estimating the bottleneck transmission capacity. The new proposal selects what packet pairs should be used for calculating the metric.

Palavras-chave

1. Medição Ativa
2. Métricas Fim-a-fim
3. Sincronização de Relógios
4. Ferramenta TANGRAM-II Traffic Generator

Glossário

Skew:	Diferença na taxa de crescimento dos relógios;
Offset:	Diferença entre os instantes de tempo de dois relógios;
Sondas:	Pacotes usados em medições ativas (*Probes*);
QoS :	Qualidade de Serviço (*Quality of Service*);
PMF :	*Probability Mass Function*;
PDF :	*Probability Density Function*;
CDF :	*Cumulative Distribution Function*;
OWD :	Atraso em um sentido (*One way delay*);
Jitter:	Variação do atraso;
RTT :	Atraso de ida e volta (*Round Trip Time*);
LAN :	Rede Local (*Local Area Network*);
AS:	Sistema Autônomo (*Autonomous System*);
Broadcast:	Difusão;
TTL:	*Time to live*;
CBR :	Taxa Constante de Bits (*Constant Bit Rate*);
ICMP:	*Internet Control Message Protocol*;
SNMP:	*Simple Network Management Protocol*;
RFC:	*Request for comments*;
IETF:	*Internet Engineering Task Force*;
IPFIX:	*IP Flow Information Export*;
UFRJ:	Universidade Federal do Rio de Janeiro;
UNIFACS:	Universidade Salvador;
UMASS:	*University of Massachusetts*;

Sumário

Resumo v

Abstract vi

Glossário viii

1 Introdução 1

1.1 Motivação 1

1.2 Objetivos 3

1.3 Resumo das Contribuições 4

1.4 Organização do Trabalho 6

2 Medição de Tráfego 7

2.1 Medição Passiva 8

2.1.1 Técnicas de Coleta 8

Monitor de Pacotes 8

Monitor de Fluxo 10

Monitor de Agentes 11

2.1.2 Matriz de Resultado 11

Matriz de Caminho 12

Matriz de Tráfego 12

Matriz de Demanda 12

2.2 Medição Ativa 13

2.2.1 Métodos de Geração do Tráfego 14

2.2.2 Modelo de Geração de Sondas 15

3 Medidas de Desempenho 17

3.1 Medidas de Desempenho Obtidas Através de Medição Ativa 17

3.1.1 Caminho entre Dois Pontos 18

3.1.2 Perda 18

3.1.3 Atraso 19

3.1.4 Capacidade de Transmissão 20

3.1.5 Disponibilidade e Confiabilidade 22

3.2 *TANGRAM-II Traffic Generator v2.0* 22

3.2.1 TANGRAM-II 23

3.3 Resumo e Ferramentas de Medição Ativa 24

4 Cálculo do Atraso em Um Sentido 28

4.1 Definição dos Problemas para Cálculo da Métrica 29

4.2 Avaliação das Soluções Propostas 30

4.2.1 Algoritmos para Remoção do *Offset* 31

4.2.2 Algoritmos para Remoção do *Skew* 33

4.2.3 Implementações na Ferramenta e Resultados Obtidos 36

Geração e Coleta das Sondas na Ferramenta 36

Algoritmos Implementados na Ferramenta 37

Resultados com a Ferramenta 38

4.3 Outros Problemas no Cálculo do Atraso 38

4.3.1 Descrições dos Problemas de Atualizações e Ruídos 39

Atualização nos Relógios 39

Congelamento do Sistema 40

4.3.2 Proposta de Um Método para Resolução dos Problemas Acima 41

Remoção de Atualização nos Relógios 42

Remoção de Ruídos 45

4.3.3 Resultados com o Algoritmo Proposto 47

4.4 Cálculo do Atraso para Aplicações 48

5 Validação, Experimentos e Resultados **52**

5.1 Validação das Implementações na Ferramenta 52

5.2 Caracterização do Atraso em Um Sentido 55

5.2.1 Cálculo dos Parâmetros pelo Método de Momentos 56

Exponencial 57

Gaussiana (normal) 57

Lognormal 57

Pareto 59

Gamma 59

Weibull 60

5.2.2 Estimando a Distribuição que Melhor se Adequa aos Dados Coletados para o Atraso Fim-a-Fim 61

5.3 Avaliação dos Resultados dos Experimentos 62

5.4 Uso dos Resultados na Criação de Modelos 65

6 Nova Proposta para Estimar Métricas Baseadas em Pares de Pacotes 69

6.1 Método "Pares de Pacotes" 69

6.1.1 Usando o Método "Pares de Pacotes" para Estimar Capacidade de Transmissão do Enlace no Gargalo 70

6.1.2 Problemas com o Método de Keshav 72

6.2 Uma Nova Proposta para Estimar a Capacidade do Gargalo 73

6.3 Trabalho Futuro com Método *Loss Pair* 75

Estimando "Tamanho do *Buffer*" 77

Estimando "Taxa de Descarte" e Parâmetros de Políticas de Descarte 77

7 Conclusão 81

7.1 Contribuições 82

7.2 Trabalhos Futuros 83

Referências Bibliográficas 85

A *TANGRAM-II Traffic Generator Nova Versão* 89

A.1 Usando o *TANGRAM-II Traffic Generator* 89

Direção da Geração das Sondas 91

Modelo de Geração das Sondas 92

Lista de Figuras

2.1 Funcionamento da Técnica de Medição Passiva. 8

2.2 Coletor em redes locais. 9

2.3 Coletor em redes de alta velocidade. 9

2.4 Tabela de fluxos. 10

2.5 Funcionamento da Medição Ativa 13

3.1 Interface principal do Tangram-II. 23

3.2 Interfaces do Modelenv e TGIF. 24

4.1 Atraso de pacotes entre máquinas com relógios diferentes. 29

4.2 Decrescimento linear gerado pela variação do tamanho dos pacotes: (A) Decrescimento obtido por um experimento real; (B) Gráfico ilustrativo da proposta do algoritmo. 32

4.3 Funcionamento dos algoritmos para remoção do Skew: (A) Identificação do fecho convexo entre os pontos da amostra; (B) Identificação da reta f(x) para o caso (i); (C) Identificação da reta f(x) o caso (ii). . 35

4.4 Resultado dos experimentos de medição do Atraso fim-a-fim: (A) Antes da aplicação dos algoritmos; (B) Valores reais do atraso, após aplicação dos algoritmos; (C) PMF's dos valores reais do atraso. . . . 37

4.5 Resultados obtidos com a ferramenta *TANGRAM-II Traffic Generator* em experimentos entre a UFRJ e a UMASS. 38

4.6 Atraso de pacotes entre máquinas com relógios diferentes, sofrendo atualização nos relógios e congelamento do sistema operacional durante as coletas. 39

4.7 Remoção incorreta da tendência de crescimento, causada pelo erro na identificação da função convexa dos pontos. 40

4.8 Crescimento linear da reta de geração das sondas. 42

4.9 Alteração na taxa de envio das sondas gerada pela atualização do relógio. 44

4.10 Alteração na taxa de envio das sondas gerada pelo congelamento do sistema operacional. 46

4.11 Resultado obtido após aplicação dos algoritmos de remoção de atualização dos relógios e ruídos. 47

4.12 Resultado obtido após aplicação dos algoritmos de remoção de *Skew* e de *Offset*: (A) Resultado após aplicação do algoritmo para remoção do *Skew*; (B) Resultado após aplicação do algoritmo para remoção do *Offset*; (C) PMF do atraso. 48

4.13 Atraso calculado para um tráfego em um sentido, paralelo às sondas de medição nos dois sentidos:(A) com *Skew* e *Offset*; (B) após remoção de *Skew* e de *Offset*. 49

4.14 Medição obtida de um tráfego de vídeo: (A) com *Skew* e *Offset*; (B) após remoção de *Skew* e de *Offset*; (C) PMF do Atraso. 50

5.1 Atraso de “ida e volta” das sondas. 53

5.2 PMF’s estimadas nas medições: (A) PMF’s estimadas dos atrasos das sondas, em cada um dos sentidos; (B) PMF do atraso de ida e volta das sondas. 54

5.3 Comparação entre os cálculos dos atrasos: (A) PMF's: "Atraso de ida e volta" e Soma dos "Atrasos em um sentido"; (B) CDF's: "Atraso de ida e volta" e Soma dos "Atrasos em um sentido" (C) CDF's em escala LOG no eixo "Y": "Atraso de ida e volta" e Soma dos "Atrasos em um sentido". 55

5.4 Processo de escolha do valor de α, a partir do valor do coeficiente de variação 61

5.5 Estimativa das distribuições: (A) *CDF*; (B) Distribuição Complementar; (C) Distribuição Complementar em escala logaritmica. 62

5.6 Resultados de experimentos entre a UFRJ e a UMASS. 64

5.7 Resultado de uma estimativa de atraso para o tráfego de vídeo entre a UFRJ e a UMASS. 65

5.8 Resultados de experimentos entre a UFRJ e UNIFACS. 66

5.9 Exemplo de modelo do ambiente TANGRAM-II Modelenv. 67

5.10 Objeto canal. 68

5.11 Exemplo de um modelo criado no ambiente *TANGRAM-II Modelenv* onde o objeto "Canal" foi gerado com parâmetros estimados de uma coleta. 68

6.1 Ilustração do funcionamento do método Pares de Pacotes com a dispersão imposta pelo enlace de menor capacidade. 70

6.2 Estimativa da capacidade de transmissão do gargalo em uma rede local. 72

6.3 Estimativa errada da capacidade de transmissão do gargalo no caminho entre a UFRJ e a UMASS, causada por um alto tráfego concorrente. 73

6.4 Estimativa correta da capacidade de transmissão do gargalo no caminho entre a UFRJ e a UMASS, mesmo com um alto tráfego concorrente. 75

6.5 Estimativa da capacidade de transmissão do gargalo no caminho entre a UFRJ e a UMASS. 80

A.1 Estrutura da nova versão da ferramenta. 90

A.2 Interfaces da nova versão da ferramenta. 91

Lista de Tabelas

3.1 Ferramentas de Medição Ativa. 26

5.1 Resultados do *MSE* para as distribuições indicadas. 63

5.2 Freqüência obtida das distribuições nos experimentos entre a UFRJ e UMASS. 65

5.3 Freqüência obtida das distribuições nos experimentos entre a UFRJ e UNIFACS. 65

Capítulo 1

Introdução

NESTE trabalho é feito um estudo detalhado das técnicas de medição em redes de computadores, com a aplicação de algumas dessas técnicas para coleta de determinadas métricas de desempenho. Além dos métodos de medição já conhecidos e amplamente usados nos dias de hoje, foram também avaliados e implementados alguns outros métodos, recentemente propostos na literatura, para a obtenção de métricas ainda pouco exploradas em medições de desempenho de redes e algumas extensões para estas propostas. O objetivo é fazer uma avaliação de alguns parâmetros de qualidade de serviço oferecidos.

As motivações e os objetivos deste trabalho, assim como as contribuições obtidas, e ainda o roteiro da organização dessa dissertação, encontram-se descritas nas próximas seções.

1.1 Motivação

A busca pelo conhecimento e estudo das características das redes de computadores através de medições tem sido uma tarefa constante desde o início da Internet, até então conhecida como *ARPANET*. Isso se dá, principalmente, pela forma como esta rede foi concebida, seguindo um padrão de comutação por pacotes. Devido a aleatoriedade das transmissões, requisições por serviços da rede e a filosofia de "melhor esforço"(*best effort*) da Internet, não é possível garantir a qualidade dos serviços (QoS) oferecidos pela rede. Em conseqüência, é importante entender os processos aleatórios da rede, como por exemplo o retardo entre dois pontos e o processo da perda, de forma a tentar prever a QoS que as aplicações podem obter. Com isso, tem havido um crescente interesse por técnicas e ferramentas para coleta de estatísticas na Internet. Atualmente, existe um grande número de trabalhos nesta área de medição motivado, principalmente, pela crescente demanda por aplicações com requisitos mínimos de qualidade de serviço.

Estudar as características do caminho por onde trafegam os pacotes pode não só ajudar no dimensionamento e no conhecimento do estado atual das redes, mas também auxiliar na implementação de aplicações e na criação de modelos que a representem. Com isso é possível prever a capacidade da rede em fornecer um determinado serviço.

Apesar de já serem estudadas há bastante tempo, técnicas de medição continuam sendo temas de muitas pesquisas. Isso se dá pelo fato de que algumas métricas de interesse não são tão simples de serem estimadas. Várias dessas métricas, que possuem problemas para serem estimadas, são relacionadas a medidas de tempo. Quando o cálculo de um parâmetro depende do tratamento de instantes de tempo em diferentes relógios, como o retardo de pacotes em um único sentido, alguns problemas devem ser levados em consideração. Esses problemas ocorrem porque relógios distintos raramente encontram-se sincronizados. Ignorar a falta de sincronia deles, certamente causará erros nas estimativas dessas métricas. O problema citado pode ser resolvido se forem usados equipamentos que sincronizem os relógios envolvidos na medição, porém, nem sempre isso é possível. Conseqüentemente, métodos para resolver este problema, que possibilitem estimar as métricas sem a necessidade de uma infra-estrutura montada para isso, continua sendo uma motivação de estudo na área de medição de rede.

1.2 Objetivos

O objetivo macro deste trabalho é coletar e analisar os resultados de algumas métricas de interesse, obtidas através de medições ativas, em redes IP(*Internet Protocol*), principalmente na Internet. Apesar de terem sido estudadas e coletadas várias métricas, o interesse maior é por aquelas que têm alguma relação com instantes de tempo em relógios de diferentes máquinas, mas sem a necessidade do uso de equipamentos específicos de sincronização desses relógios.

Para alcançar o objetivo principal deste trabalho, foram necessários definir outros objetivos intermediários, mas fundamentais para a sua conclusão. O primeiro desses objetivos é avaliar as propostas existentes na literatura que tentam solucionar o problema da falta de sincronia entre dois relógios. Algumas dessas propostas foram implementadas, possibilitando a coleta da métrica retardo fim-a-fim entre duas máquinas, sem equipamentos de sincronia dos relógios. Com isso, a validação de toda essa implementação é também um objetivo deste trabalho.

Resolver a falta de sincronia entre dois relógios, em princípio, parece resolver o problema do cálculo do retardo sofrido por pacotes em um único sentido. Porém, alguns resultados obtidos nos experimentos revelam a existência de outros fatores que também devem ser considerados. A atualização dos relógios, ou mesmo o congelamento temporário (isto é, por um espaço curto de tempo) do sistema operacional, em qualquer uma das duas máquinas durante a medição, pode invalidar os resultados obtidos no cálculo das métricas. Uma vez que não existem propostas na literatura que se disponham a resolver os problemas citados, passa a fazer parte dos objetivos deste trabalho a proposta de uma extensão dos métodos estudados para tratar dos problemas causados por atualização dos relógios ou congelamento temporário do sistema operacional durante a medição.

Devido aos problemas encontrados para calcular o retardo fim-a-fim, as propostas existentes na literatura para resolver este problema prevêem restrições na forma como a medição deve ser aplicada. Estas restrições variam desde o modo de geração das sondas[1] até o sentido em que elas devem ser geradas. As imposições das propostas implicam em certas restrições na estimativa das métricas, conforme será descrito nas próximas seções. Serão propostas técnicas para contornar os problemas encontrados.

Após tratados os problemas encontrados para estimar o atraso em um único sentido, foram executados durante alguns meses uma série de experimentos entre a UFRJ (Universidade Federal do Rio de Janeiro) e a UMASS (*University of Massachusetts*). O objetivo foi estudar as métricas coletadas durante os experimentos, e fazer uma análise criteriosa desses resultados. Um dos objetivos foi verificar a existência de algum padrão nos valores de parâmetros estimados para retardo fim-a-fim encontrado durante os experimentos, de forma a poder obter, por exemplo, as distribuições que melhor representem esta métrica. Para tal, métodos foram implementados para fazer essa estimativa. Os resultados podem ser usados na criação de modelos de simulação ou analíticos mais reais, usando distribuições que melhor representem o evento "Retardo dos pacotes na rede".

Além da métrica retardo fim-a-fim, com a medição em um único sentido, é também possível estimar outras métricas de interesse, como por exemplo: perda, variação do atraso, vazão, capacidade de transmissão do enlace no gargalo, dentre outras. Porém, em medições onde exista um alto tráfego concorrente, a estimativa de métricas como a capacidade de transmissão do gargalo pode sofrer influência. Neste trabalho, é ainda proposto um novo método que usa a métrica retardo fim-a-fim para aumentar a precisão do cálculo da capacidade de transmissão do enlace de um gargalo no caminho, em situações de alta carga na rede.

1.3 Resumo das Contribuições

As principais contribuições deste trabalho são: a avaliação de resultados obtidos, através de medições executadas na Internet; e a criação de uma ferramenta de medição ativa para redes IP, que possibilita a obtenção de métricas fim-a-fim,

[1] Pacotes gerados por ferramentas de medição ativa.

inclusive algumas relacionadas ao retardo sofrido por pacotes em um único sentido, sem o uso de equipamentos para sincronização dos relógios. Para a criação desta ferramenta, foi necessário levar em consideração alguns pontos importantes. Vários problemas existentes no cálculo desta métrica foram avaliados. Métodos já propostos para solucionar alguns dos problemas foram estudados, e novas propostas foram também desenvolvidas. De todos os pontos levantados durante a implementação da ferramenta, alguns devem ser destacados também como contribuição desse trabalho:

- Avaliação e implementação de algumas técnicas de medição existentes na literatura, e extensão destas técnicas, para estimar a métrica retardo em um único sentido sem uso de *hardware* de sincronização. Foi também elaborado um método que possibilita estimar a métrica retardo em um único sentido, a partir de modelos de geração de tráfego;

- Implementação de um módulo para a obtenção de métricas no sentido "ida e volta", e a partir deste fazer a validação da implementação do módulo que estima as métricas "em um único sentido";

- Implementação de métodos para obter a estimativa da melhor distribuição de probabilidade que represente o retardo obtido em experimentos;

- Criação de módulo para facilitar a construção de modelos de canais de comunicação no ambiente *TANGRAM-II*, usando medidas coletadas a partir de experimentos reais;

- Implementação de um método que possibilita estimar a capacidade de transmissão do enlace no gargalo;

- Desenvolvimento de uma nova proposta, e implementação do método, que possibilita estimar a capacidade de transmissão no enlace do gargalo em situações de alto tráfego concorrente;

- Implementação dos métodos para calcular diversas métricas de interesse baseadas no retardo, como por exemplo o tamanho do *Buffer* no gargalo proposto em [30].

1.4 Organização do Trabalho

No Capítulo 2, é feita uma comparação entre as duas técnicas existentes para medição de tráfego: Ativa e Passiva. Detalhes do funcionamento e da forma de aplicação dessas técnicas também serão abordados neste capítulo, porém esta tese trata de medições ativas.

No Capítulo 3, são estudadas as medidas de desempenho obtidas com a técnica de medição ativa, e algumas ferramentas disponíveis para esta finalidade. É ainda apresentada neste capítulo uma breve descrição da ferramenta *TANGRAM-II Traffic Generator*, antes da novas implementações disponibilizadas no final deste trabalho.

Uma introdução sobre os problemas existentes para calcular o retardo em um único sentido é abordada no Capítulo 4. São também abordados neste capítulo as soluções existentes para alguns dos problemas descritos, assim como a apresentação de uma nova proposta que tenta resolver outros problemas encontrados durante experimentos, na coleta desta métrica.

Todas as propostas descritas no capítulo anterior foram implementadas em uma ferramenta. A validação das implementações e os resultados coletados serão apresentados no Capítulo 5. Neste capítulo é feita, também, a análise de resultados obtidos com uma série de experimentos executados entre a UFRJ e a UMASS.

O Capítulo 6 descreve um método para estimar a capacidade de transmissão do gargalo usando o método de pares de pacotes, assim como outros parâmetros baseados no retardo e na capacidade de transmissão. Uma avaliação da proposta de pares de pacotes é apresentada, e um novo método é proposto, com a mesma finalidade. A proposta tenta estimar com maior precisão a capacidade do gargalo usando o retardo fim-a-fim. A comparação entre resultados obtidos com os dois métodos demonstra a maior eficiência da nova proposta.

As conclusões e os trabalhos futuros estão descritos no Capítulo 7, e um resumo das novas características da ferramenta implementada encontra-se no Apêndice.

Capítulo 2

Medição de Tráfego

DUAS são as técnicas de medição de redes atualmente usadas: ativa e passiva. A primeira consiste na injeção de pacotes de controle na rede, chamados de *Sondas* (ou *Probes*). Essas sondas atravessam todo o caminho da rede a ser medido e são coletadas em algum ponto. Através das informações coletadas, algumas medidas de desempenho podem ser extraídas. Dependendo da medida a ser extraída, as sondas devem carregar consigo informações que serão analisadas no receptor. Para algumas outras medidas, técnicas mais sofisticadas podem ser necessárias.

A segunda técnica, chamada de medição passiva, apenas observa os pacotes que trafegam pela rede. Ao contrário da medição ativa, as passivas não injetam pacotes. Monitores, que funcionam em modo promíscuo, são colocados no caminho dos pacotes para coletar as informações que, posteriormente, serão analisadas. Apesar de não sobrecarregar a rede com pacotes adicionais, equipamentos e *softwares* sofisticados podem ser necessários para fazer a coleta e a análise dos dados.

Ao longo desse capítulo é feito um estudo das duas técnicas de medição de tráfego citadas acima. Na seção que trata de medição passiva, o estudo descreve a diferença entre as técnicas de coleta do tráfego e a forma como os dados coletados podem ser interpretados. A outra seção, que trata de medição ativa, descreve os métodos existentes de aplicação desta técnica e os diferentes modelos de geração das sondas.

2.1 Medição Passiva

O conceito na aplicação desta técnica é, basicamente, sempre o mesmo. A coleta de informações do tráfego que passa por um determinado ponto da rede é feita a partir de equipamentos e *softwares* instalados para esta finalidade. Na metodologia de medição passiva não existem sondas e as informações são coletadas a partir dos pacotes que trafegam pela rede gerados por diversas aplicações. No entanto, a forma como estas informações são coletadas e interpretadas podem ser distintas. O funcionamento desta técnica é ilustrado na Figura 2.1.

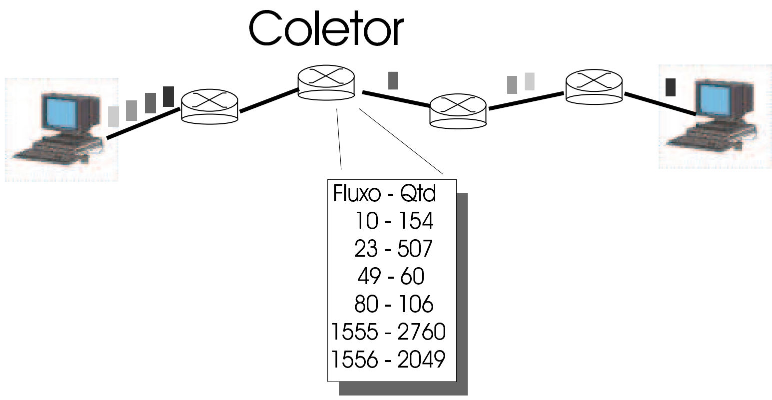

Figura 2.1: Funcionamento da Técnica de Medição Passiva.

2.1.1 Técnicas de Coleta

As técnicas definem como os "coletores" ou "monitores" retiram as informações dos pacotes que passam pela rede. O local onde esses coletores se instalam na rede e o tipo de informação que deve ser estimada dos pacotes podem variar. A seguir serão descritas e exemplificadas três técnicas definidas em [18].

Monitor de Pacotes

Esta técnica consiste na cópia de parte do conteúdo dos pacotes que passam pelo monitor. Com isso, informações de todas as camadas (Aplicação, Transporte, Rede, e Enlace) podem ser extraídas. Devido ao modelo de propagação dos pacotes usado pelas redes *Ethernet*, difusão (*broadcast*) no meio físico, este tipo de coleta pode facilmente ser aplicada em redes locais. Neste caso, a captura é feita pelo monitor que possui seu adaptador de rede configurado de forma promíscua, ficando apto a coletar todo o tráfego que passe pelo mesmo meio físico em que o coletor é instalado, como ilustra a Figura 2.2.

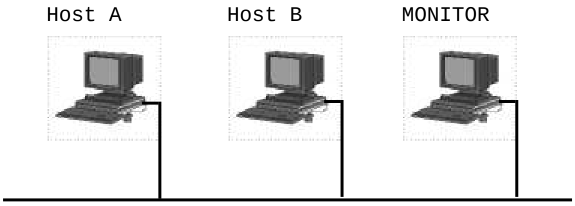

Figura 2.2: Coletor em redes locais.

Para redes locais o funcionamento é simples. No entanto, para as redes ponto-a-ponto de alta velocidade algumas dificuldades existem: primeiro, o volume de tráfego nas redes de alta velocidade é muito maior do que nas redes locais; segundo, existe o problema de instalação do coletor dos pacotes no meio físico. Nem sempre é possível dividir a rede e instalar o monitor entre os dois pontos, como mostra a Figura 2.3, porém, uma alternativa para este problema é a configuração de geração de *traces* nos roteadores, que então os envia aos coletores.

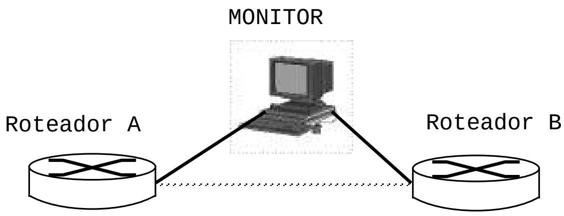

Figura 2.3: Coletor em redes de alta velocidade.

Para este tipo de monitoração, é sempre importante lembrar que o volume de informações tratado pode ser muito grande. Uma boa opção é aumentar a granularidade feita na filtragem dos pacotes. Ferramentas de domínio público, como tcpdump [45], que trabalham como monitores de redes locais são atualmente amplamente usadas.

Monitor de Fluxo

Este tipo de coleta é feita com base na medição por fluxo, onde a idéia é gerar estatísticas do tráfego agregado que passa pela rede. Para isso, um fluxo é definido como sendo uma agregação do tráfego que se enquadra em uma determinada regra, e elas devem descrever as características existentes naquele conjunto de pacotes que se deseja medir.

Um fluxo pode ser por exemplo: todo o tráfego que passe pelo coletor, cujo campo endereço IP de origem e de destino sejam semelhantes aos indicados na tabela; ou, além de serem de origem e destino iguais, utilizem uma porta específica. O intervalo de tempo entre pacotes pode também ser um outro parâmetro de diferenciação entre os fluxos, onde um pacote só pertence a um mesmo fluxo caso esteja separado por um intervalo máximo de tempo.

Na monitoração por fluxo, uma tabela é criada pelo "monitor" semelhante à ilustrada na Figura 2.4. Todo pacote que passe pelo coletor deve ser comparado às regras existentes, caso as suas características se enquadrem em alguma das classes definidas, a tabela deve ser incrementada. Caso contrário, uma nova classe pode ser criada.

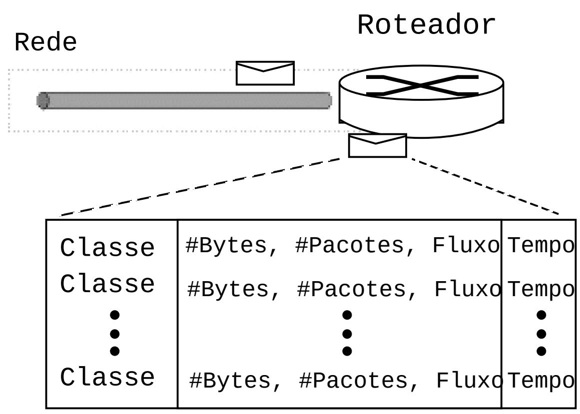

Figura 2.4: Tabela de fluxos.

A coleta por fluxo possui em geral uma granularidade maior em relação à existente nos monitores de pacotes, mas essas granularidades podem ser diminuídas de acordo com a necessidade da medição. Outro problema desta técnica é a necessidade da existência dos coletores nos equipamentos. No entanto, alguns já possuem essas ferramentas nativas de fábrica. Um exemplo são os equipamentos da Cisco com a ferramenta Netflow [12]. O Grupo de trabalho IPFIX *IP Flow Information Export* [1] do IETF (*Internet Engineering Task Force*) vem trabalhando na definição de um padrão de transferência de informações dessas ferramentas.

Monitor de Agentes

O "monitor de agentes" é baseado na medição feita por agentes localizados em equipamentos da rede. Esta técnica implementada por um protocolo SNMP (*Simple Network Management Protocol*), é um padrão criado pelo IETF para administração de rede. As estruturas administradas por este protocolo são definidas como MIB (*Management Information Base*). Nas MIB's são definidos os elementos que geram informações do estado atual da rede e essas estruturas definem agentes que são instalados nos equipamentos, como os roteadores que periodicamente responde a pedidos de seus gerentes/coletores enviando informações do seu estado.

Estruturas como MIB-II e RMON são padrões do protocolo SNMP que possuem os tais agentes definidos para coletar as informações da rede. As informações relacionadas ao tráfego, que passam por estes agentes, são coletadas e enviadas a um centralizador definido e configurado para isso. Os dados coletados podem ser então analisados, e medidas podem ser estimadas a partir deles.

2.1.2 Matriz de Resultado

A matriz de resultado é uma forma de se representar as coletas adquiridas com o monitoramento do tráfego, e de apresentar os resultados obtidos com as medições feitas na rede. A idéia desse tipo de representação é auxiliar na detecção de problemas existentes na rede, ou mesmo, na simplificação da visualização de suas características. Nas próximas seções serão descritas três formas de apresentação dessas matrizes, e em alguns casos a aplicabilidade de resultados. Nas seções seguintes, considera-se uma rede como sendo um AS(*Autonomous System*), e as origens e os destinos como pontos de entrada e saída da rede, respectivamente.

Matriz de Caminho

Na matriz de caminho, para toda origem (**O**) e destino (**D**), um caminho (**C**) é criado com um valor relacionado ao volume de tráfego (**V(C)**) experimentado neste trecho, de acordo com as medições executadas durante um determinado intervalo de tempo. Essa matriz captura o estado corrente e o comportamento da rede, proporcionando uma descrição detalhada do fluxo de tráfego que passa pela rede.

Este tipo de operação possibilita aos administradores detectar e diagnosticar problemas existentes em sua rede, por exemplo quando o volume de fluxo encontrado em um determinado caminho for muito grande, o administrador pode optar por fazer alterações nas tabelas de roteamento da AS, na tentativa de solucionar algum tipo de congestionamento.

Matriz de Tráfego

A matriz de tráfego especifica o volume de dados (**V(O,D)**) experimentado em cada par de origem (**O**) e destino (**D**) existentes na rede. Neste caso, um fluxo é definido de acordo com a sua origem e seu destino, independente do caminho usado pelos pacotes.

Algumas aplicações são possíveis com esses resultados, sendo uma delas a previsão em caso de uma possível alteração na rede, como por exemplo na rota. Isso possibilita ao operador prever o impacto que uma alteração pode causar à sua rede, baseado na idéia hipotética de que o fluxo não se altera com o tempo.

Matriz de Demanda

A matriz de demanda tem como característica representar o volume de tráfego ($\boldsymbol{V(O,\ \{D_1,\ \dots\ ,\ D_n\})}$) ocorrido entre uma origem **O** e vários destinos **D**. Esta matriz representa todo o fluxo que passa pelo AS, possibilitando que seja estimado todo o volume de tráfego inserido por um ponto de entrada desse AS. Com isso, é possível estimar o total do tráfego inserido na rede a partir de um ponto igual ao valor $V(O, \{D_1, \ldots, D_n\})$.

Se for considerado que o cliente de um AS equivale ao ponto de entrada O dessa rede, é possível gerar uma forma de tarifação baseada no fluxo injetado na rede pelo cliente.

2.2 Medição Ativa

Diferente da medição passiva, que apenas coleta informações do tráfego que passa pela rede, na medição ativa as métricas são estimadas a partir de informações coletadas de pacotes injetados por ferramentas apropriadas. A Figura 2.5 ilustra o funcionamento desta técnica de medição, que se resume no envio de pacotes por uma determinada máquina, e que serão coletados em algum ponto da rede, podendo este ponto ser a mesma máquina emissora ou uma outra máquina qualquer.

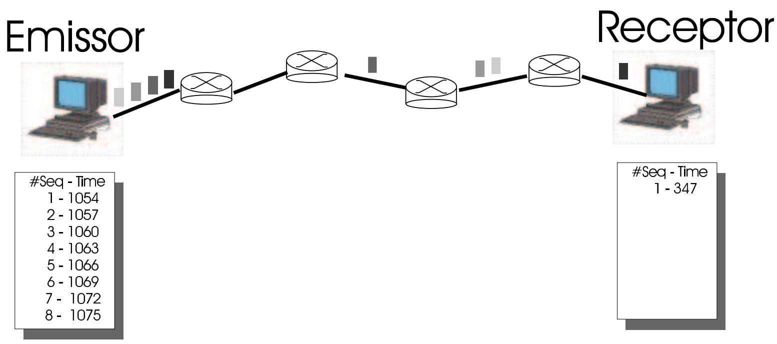

Figura 2.5: Funcionamento da Medição Ativa

Através das medições ativas é possível estimar uma série de medidas de desempenho, e essas medidas serão descritas com detalhe no Capítulo 3. No entanto, devido a vários problemas para estimar determinadas métricas, em alguns casos, uma técnica específica de medição ativa pode ser exigida.

A solução para alguns desses problemas é a variação na forma como as sondas são enviadas pela rede. Para cada métrica a ser coletada, pode ser exigido um método diferente de geração das sondas, assim como diferentes modelos de geração.

2.2.1 Métodos de Geração do Tráfego

A escolha do "método de geração do tráfego" para uma medição ativa consiste em decidir que formato de geração de sondas será usado durante a medição. Por exemplo o método onde as sondas são enviadas em uma única direção, ou replicadas pelos receptores, ou ainda geradas nas duas direções simultâneamente. A forma de enviar as sondas deve ser decidida de acordo com o método a ser usado na medição. O "método de geração do tráfego" define quais tipos de métricas podem ser estimadas a partir da coleta das sondas geradas. Portanto, a escolha do método é determinada, muitas vezes, por restrições impostas pelas próprias métricas a serem estimadas.

Dependendo da medida de interesse, diferentes métodos, quanto à geração das sondas, podem ser necessários. As três formas de variação estudadas neste trabalho estão descritas a seguir.

- *Um Sentido*: Neste formato de geração de tráfego, as sondas são enviadas a partir de uma origem, passam pelo caminho a ser estudado, sendo coletadas por um receptor em algum outro ponto da rede. A partir desta coleta, apenas algumas métricas podem ser estimadas. É importante perceber o problema para o cálculo de determinadas medidas de desempenho, usando este método. Com esta coleta em um único sentido, não é possível calcular com precisão medidas como o retardo, sem uso de equipamentos especiais. Este problema será detalhado nos Capítulos 3 e 4.

- *Dois Sentidos*: Este método de geração de sondas consiste na inicialização "simultânea" de gerações em um único sentido, entre dois pontos. Cada um dos dois pontos inicializa uma geração em um único sentido, para o outro, paralelamente. Essa variação do método "Um sentido", somada a uma variação no tamanho dos pacotes permite resolver problemas relacionados à falta de sincronia dos relógios. As soluções propostas para este problema estão descritas no Capítulo 4.

- *Ida e Volta*: A geração de sondas no método "Ida e volta" funciona como a ferramenta *PING* [35]. Os pacotes de medição são gerados a partir de uma origem, na direção do destinatário. Porém, ao contrário do método "Um sentido", os pacotes não são coletados no destino, eles são replicados e reenviados à origem. Desta forma, consegue-se estimar as características dos caminhos de ida e volta daquele pacote. Estas características podem, inclusive, ser referentes aos instantes de tempo de envio e recebimento dos pacotes, uma vez que os "carimbos de tempo" são tirados de um único relógio, localizado em uma mesma máquina.

2.2.2 Modelo de Geração de Sondas

O "intervalo entre sondas" é, em alguns casos, relacionado à métrica a ser estimada. As sondas podem ser transmitidas em rajadas de um ou mais pacotes e o intervalo entre as rajadas são parâmetros definidos pelas técnicas escolhidas. Os intervalos de geração mais comuns estão descritos a seguir.

- *CBR*: Na geração de tráfego CBR (*Constant Bit Rate*), as sondas são geradas seguindo um modelo determinístico, onde os intervalos entre os envios são constantes.

- *Modulado por Markov*: No Modulado por Markov (ou *Markov Modulated*), o tráfego gerado segue um modelo de Markov, de tempo contínuo, com número finito de estados. Uma taxa de envio do tráfego associada a cada estado da cadeia deve ser definida.

- *Trace*: O modelo de geração "*Trace*" prevê que o envio das sondas de medição são determinadas por um arquivo de entrada. Este modelo possibilita a representação de situações reais de uso das aplicações na rede, como por exemplo a transmissão de um vídeo em tempo real, onde no *trace* constam o tamanho dos quadros (*frames*) e os instantes do tempo de geração dos pacotes. Desta forma, é possível simular a transmissão desse vídeo em tempo real, e as métricas estimadas representam medidas de desempenho semelhantes às que seriam obtidas com a real utilização desse tipo de aplicação na rede.

- *Pares de Pacotes*: Neste caso são geradas rajadas de sondas consecutivas. O uso desta forma de geração de sondas, assim como as métricas estimadas, está descrito no Capítulo 6.

Capítulo 3

Medidas de Desempenho

NESTE capítulo serão estudadas as principais medidas de desempenho que podem ser estimadas em uma rede, através de medições ativas. O objetivo não é só descrever a definição das métricas estudadas durante o trabalho, mas também fazer um resumo das ferramentas disponíveis para a obtenção de algumas dessas medidas definidas. É ainda apresentado neste capítulo uma descrição mais detalhada da ferramenta *TANGRAM-II Traffic Generator*, antes de serem acrescentadas as implementações desenvolvidas durante este trabalho.

3.1 Medidas de Desempenho Obtidas Através de Medição Ativa

As medidas de desempenho descrevem as características do estado da rede encontradas durante experimentos. Essas medidas podem ser obtidas através de técnicas de medições ativas ou passivas. A seguir serão descritas as medidas de desempenho estimadas a partir das técnicas de medição ativa. No entanto, algumas dessas métricas podem também ser estimadas na forma passiva de medição, mas, por não fazer parte do escopo desta tese, essas técnicas não serão descritas.

3.1.1 Caminho entre Dois Pontos

O objetivo desta métrica é descobrir o caminho entre dois pontos da rede. Para esta métrica, uma forma especial do método "ida e volta" é usado, e cada um dos saltos do caminho, entre uma origem e um destino, é descoberto. Para isso, são enviadas sondas a um destino, cada uma com o valor do campo *TTL(Time-to-live)* variando de 1 ao número de saltos totais até o destino. Um roteador, ao receber pacotes com o campo *TTL* igual a zero, retorna à origem uma mensagem de informação de expiração daquele pacote. Portanto, a sonda com valor de *TTL* inicial igual a n, terá esse valor zerado no n-ésimo roteador, que então gera uma mensagem de controle para a fonte. O gerador da sonda descobre então o roteador que enviou o controle.

A ferramenta *TRACEROUTE* [28] implementa esta técnica. Porém, a medida pode ser obtida por qualquer outra ferramenta baseada na técnica de variação do campo *TTL*.

A aplicação de vários experimentos de descoberta dos caminhos, entre "pares de pontos" de uma rede, possibilita estimar a topologia existente naquela infra-estrutura. Estimar a topologia de uma rede é um tópico bastante estudado na área de medição, tanto usando técnicas ativas quanto passivas, e um dos métodos usados para auxiliar essas estimativas é exatamente o da descoberta de caminhos entre vários pontos da rede.

3.1.2 Perda

A sensibilidade das aplicações em relação ao número de pacotes perdidos motiva o estudo dessa métrica. Obviamente todas as aplicações são sensíveis à perda de pacotes. Estes são recuperados via retransmissão em geral. Entretanto, aplicações, como as de transmissão de vídeo e voz em tempo real, não permitem que haja re-transmissão, tornando essas aplicações particularmente sensíveis a perdas. Portanto, conhecer as características de perda e estudar o desempenho dos algoritmos de re-cuperação de pacotes auxilia. Dependendo do processo de perda presente na rede, mecanismos como a redução na qualidade do vídeo ou algoritmos de recuperação de perdas podem ser aplicados para melhorar a qualidade do serviço. A escolha dos mecanismos de recuperação depende do processo de perda na rede. Este processo pode ser estimado com as seguintes medidas de desempenho:

- **Fração de Perda dos Pacotes:** Número total de pacotes perdidos dividido pelo total de pacotes enviados. Esta métrica pode ser estimada a partir de várias ferramentas de medição ativa. Porém, não só a média, mas também outros detalhes sobre o processo de perda devem ser levados em consideração.

- **Perda Consecutiva:** A métrica "Fração de Perda dos Pacotes" não fornece ao avaliador maiores detalhes do comportamento das perdas dos pacotes. O impacto nas aplicações quando perdas ocorrem em rajadas e quando ocorrem distribuídas ao longo da coleta é em geral bem diferente. Esta diferença influencia também na escolha de uma solução para o problema. A métrica "Perda Consecutiva" calcula a distribuição do número total de pacotes perdidos em seqüência.

- **Sucesso Consecutivo:** De forma semelhante à métrica "Perda Consecutiva", esta medida fornece uma estimativa do número de pacotes transmitidos corretamente entre duas "rajadas de perda". A métrica calcula a distribuição do número de pacotes entregues corretamente em seqüência ao destino, entre duas rajadas de perda.

3.1.3 Atraso

As métricas relacionadas ao atraso em redes estimam o tempo que leva para um pacote sair de sua origem e chegar ao seu destino. No entanto, diversos problemas devem ser considerados para se estimar esses parâmetros, como falta de sincronia e diferentes taxas de crescimento dos relógios do transmissor e receptor. As soluções para estes e outros problemas serão discutidos no Capítulo 4. A seguir, serão definidas algumas métricas que avaliam o retardo sofrido por pacotes na rede.

- **Atraso fim-a-fim:** O tempo que um pacote leva do emissor ao receptor. Esta métrica possui vários problemas para ser estimada devido às diferenças dos relógios do transmissor e receptor. O problema pode ser solucionado se forem usados equipamentos de sincronização dos relógios. Porém, esses equipamentos nem sempre estão disponíveis. As propostas para solucionar estes e outros problemas serão tratadas no próximo capítulo.

- **Atraso de ida e volta:** É o tempo que um pacote leva para ser enviado a um receptor e devolvido ao emissor (*Round Trip Time*). Neste caso, sem o problema de sincronia dos relógios, o parâmetro fica muito mais simples de ser obtido. Algumas considerações devem ser feitas, por exemplo, a precisão mínima de leitura do relógio no sistema operacional. O atraso sofrido pelos pacotes na ida pode ser completamente distinto do retardo durante o retorno, e a existência desta diferença não pode ser estimada através dessa medida. Ferramentas, como *PING*, estão disponíveis para estimar esta métrica.

- **Variação do atraso entre pacotes(*Jitter*):** O *Jitter* é o intervalo entre a chegada de dois pacotes consecutivos em relação ao intervalo de sua transmissão. Diferente do "Atraso fim-a-fim", se os instantes de envio forem conhecidos ou o intervalo entre eles for constante, essa métrica não possui problemas para ser estimada entre máquinas com relógios não sincronizados.

3.1.4 Capacidade de Transmissão

Várias são as métricas de interesse relativas à capacidade de transmissão, e diferentes definições existem para esses tipos de métricas: capacidade de transmissão de cada um dos enlaces ao longo da rede, ou apenas do gargalo; largura de banda disponível ou alcançável em um caminho; vazão; utilização; dentre outras, sem citar as variações dessas métricas. Neste trabalho estão descritas apenas algumas dessas medidas de desempenho.

- **Capacidade de Transmissão dos Enlaces:** Esta medida possibilita descobrir a capacidade de transmissão de todos os enlaces ao longo de um caminho.

Esta métrica pode ser obtida de forma simples, via SNMP. Porém, o acesso a esse tipo de informação, em geral, não é possível, principalmente quando envolve redes públicas como a Internet. Por isso, métodos de medição ativa, que estimem essa métrica, vêm sendo estudados, e ferramentas desenvolvidas. Estas ferramentas (vide 3.3) usam a coleta de medidas do atraso de ida e volta, com a variação do campo TTL das sondas, para estimar a capacidade de transmissão dos enlaces.

- **Capacidade de Transmissão do Enlace no Gargalo:** Neste caso, a métrica estima apenas a capacidade de transmissão do enlace no gargalo do caminho entre uma fonte e um destino. O método "Pares de Pacotes" é o mais usado para esta finalidade. Porém, outros métodos vêm sendo desenvolvidos para calcular esta métrica. Uma das técnicas é o "Trem de Pacotes", onde no lugar de apenas duas, uma série de sondas são enviadas de um ponto a outro da rede. As duas propostas supõem que a utilização dos enlaces do caminho estudado é baixa, o que pode gerar erros na estimativa, quando esta suposição é violada. No Capítulo 6 a técnica de "Pares de Pacotes" será descrita em detalhes, assim como os problemas da sua utilização. Ainda no Capítulo 6, é apresentada uma proposta para minorar o erro desta técnica, quando usada em canais de alta utilização, com grande volume de tráfego concorrente.

- **Utilização do Enlace:** A métrica "utilização" estima a fração do total da capacidade de transmissão do enlace que é usada pelos fluxos que por ele passam. Esta métrica pode simplesmente ser obtida através da razão entre o total de banda usada pelo fluxo dividido pela capacidade de transmissão do enlace.

- **Capacidade Disponível:** É a capacidade de um enlace multiplicada pela fração não utilizada dele. Obviamente a métrica representa a capacidade não utilizada por um enlace.

- **Vazão:** Esta métrica é obtida através da razão do número total de pacotes enviados pelo seu tempo total de envio.

3.1.5 Disponibilidade e Confiabilidade

As métricas Disponibilidade e Confiabilidade estimam o montante de tempo que os enlaces ficam ativos ou inativos. Essas informações permitem, por exemplo, que sejam verificados se os acordos de serviços estão sendo cumpridos por parte do provedor.

Para se estimar essas métricas, é necessário definir o que deve ser considerado como "ativo" ou "inativo" em um enlace. Para exemplificar o problema, a simples falha de um pacote pode ser considerada como inatividade de um enlace, porém, nem sempre esse evento acontece em conseqüência de uma falha. Por isso, normalmente, um número "x" de perdas consecutivas deve ser atingido para que seja constatada uma queda no serviço.

- **Disponibilidade:** Esta medida estima o percentual de tempo que um enlace fica disponível em relação ao tempo total de observação.
- **Confiabilidade:** Esta métrica calcula a probabilidade de um enlace ficar em funcionamento durante todo o período de observação.

3.2 *TANGRAM-II Traffic Generator v2.0*

A ferramenta *TANGRAM-II Traffic Generator* [26] foi idealizada para ser o módulo de medição ativa do ambiente *TANGRAM-II* [27]. O gerador de tráfego implementado na versão 2.0 (anterior a este trabalho) foi concebido para dar ao usuário grande flexibilidade na geração de tráfego *IP* ou *ATM* nativo. Em mais detalhes, ao usuário é permitida a especificação das distribuições dos tempos entre rajadas de pacotes, e a forma com que essas rajadas são enviadas. O tráfego gerado pode ser especificado através de um modelo markoviano (por exemplo, *MMPP*[1]) ou de simulação, ou até mesmo a partir de arquivo de *traces*. Com gerações de sondas a partir de modelos definidos pelo usuário (*CBR*, *Markov Modulated* ou *Trace*), nos

[1]Markov Modulated Poisson Process.

formatos *unicast* e *multicast*, a ferramenta possibilita estimar algumas métricas de interesse como: *jitter*, perda, sucesso e vazão.

Uma das grandes vantagens do *TANGRAM-II Traffic Generator* é a possibilidade de integração com os demais módulos do *TANGRAM-II*, permitindo, por exemplo, que sejam definidos e resolvidos modelos de tráfego que serão mais tarde utilizados como entrada para a geração das sondas. Na seção seguinte, é feita uma breve descrição do ambiente *TANGRAM-II* e de um de seus principais módulos, o *TANGRAM-II Modelenv*. Nos capítulos que se seguem serão indicadas as novas funcionalidades da ferramenta, obtidas pela inclusão de algoritmos recentes de medição ativa e as extensões propostas.

3.2.1 TANGRAM-II

O *TANGRAM-II* [7, 38, 27, 15] é um ambiente de modelagem e experimentação de sistemas de computação e comunicação, desenvolvidos com propósito educacional e de pesquisa. O ambiente da ferramenta é composto por vários módulos, com diferentes finalidades. A Figura 3.1 ilustra a tela principal deste ambiente, que inclui módulos para experimentos de rede, modelagem de tráfego e aplicações multimídia.

Figura 3.1: Interface principal do Tangram-II.

Um dos módulos do *TANGRAM-II*, o *Modelenv (Modeling Environment)*, é um ambiente de modelagem que permite ao usuário a criação de modelos, analíticos ou de simulação, com o uso de uma interface gráfica, *Tangram User Interface* [10]. Um modelo é representado por uma coleção de objetos que interagem entre eles através do envio e recebimento de mensagens. Uma série de métodos de solução estão disponíveis para possibilitar a coleta das medidas de interesse dos modelos criados. As interfaces do *Modelenv* e do *TGIF* estão ilustradas na Figura 3.2.

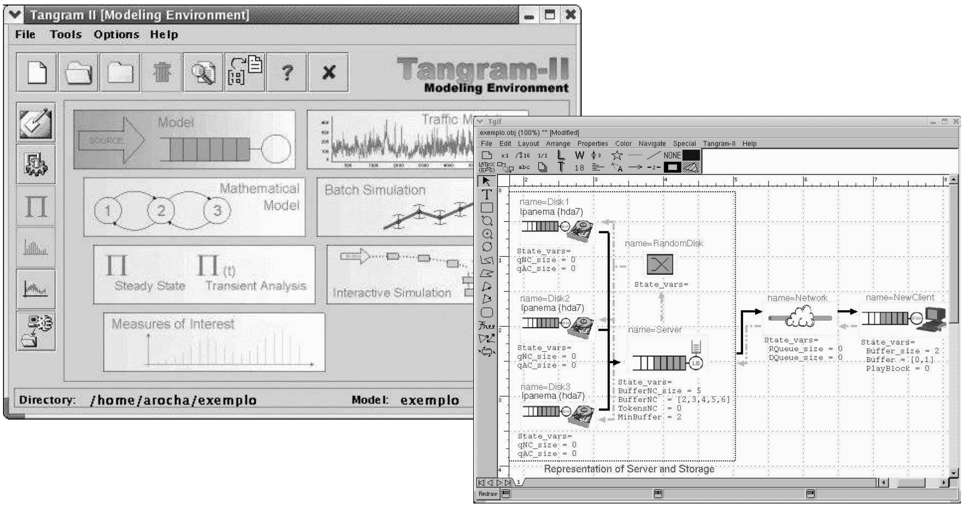

Figura 3.2: Interfaces do Modelenv e TGIF.

3.3 Resumo e Ferramentas de Medição Ativa

Nesta seção, é feito um resumo de algumas ferramentas de medição ativa. As principais medidas de desempenho estimadas por elas também estão listadas neste resumo. A partir do resumo das ferramentas de medição ativa disponíveis será possível comparar o trabalho desta dissertação com o estado da arte de ferramentas ativas, e com isso avaliar nossa contribuição e a nova versão do *TANGRAM-II Traffic Generator*.

São dezesseis as principais ferramentas de medição ativa disponíveis em domínio público. A ferramenta Traceroute usa pacotes *ICMP* para medir o atraso de ida e volta e o caminho percorrido pelas sondas. Pacotes *ICMP* são usados também pelas ferramentas Ping e Bing para estimar o atraso de ida e volta e a fração de perda das sondas.

As ferramentas Pchar e Pathchar usam, além de sondas *ICMP*, sondas *UDP* para estimar a capacidade de transmissão dos enlaces ao longo do caminho, vazão, atraso de ida e volta e fração de perda. Para isso, pacotes de tamanhos variados são gerados pelas ferramentas. A ferramenta Clink usa a mesma técnica para estimar a capacidade de transmissão dos enlaces ao longo do caminho e o atraso de ida e volta.

O conjunto de ferramentas Bprobe, Cprobe e Sprobe usa o conceito de pares de pacotes para calcular suas métricas. As ferramentas Bprobe e Cprobe usam pacotes *ICMP* para calcular a capacidade de transmissão do enlace no gargalo e a utilização respectivamente, enquanto que a ferramenta Sprobe usa pacotes *UDP* para estimar a capacidade de transmissão do gargalo. O método de pares de pacotes é também usado pelas ferramentas Netest e Nettimer para estimar a capacidade de transmissão do gargalo. O Netest estima ainda a capacidade disponível, atraso de ida e volta, e fração de perda.

A variação do método de pares de pacotes, o trem de pacotes, é usado pelas ferramentas Pathrate e Pipechar para estimar a capacidade de transmissão de gargalo. O Pipechar também estima a capacidade disponível, atraso de ida e volta, e fração de perda.

A ferramenta Treno usa pacotes *UDP* e *ICMP* para simular o funcionamento do *TCP* e estimar métricas como: capacidade de transmissão de gargalo e capacidade disponível.

A ferramenta Iperf faz uso de sondas *UDP* e *TCP* para estimar as métricas *jitter*, fração de perda e capacidade disponível.

Algumas ferramentas exigem equipamentos específicos para fazer as estimativas corretamente. Uma delas é a Owping, que para estimar o atraso das sondas em um sentido requer o uso de equipamentos para sincronização dos relógios (*GPS*). Esta ferramenta ainda estima a fração de perda das sondas experimentada na medição.

Além das ferramentas, infra-estruturas vêm sendo implementadas na tentativa de se criar um ambiente de medições ativas na Internet. RIPE [4], Surveyor [5], e IPMA [2] são alguns desses ambientes criados com esta finalidade. Essa infra-estrutura é formada por várias máquinas que executam medições ativas entre si, fazendo uso de

Ferramenta	Métricas
BING	Atraso de ida e volta, Fração de Perda
BPROBE	Capacidade de Transmissão do Gargalo
CPROBE	Utilização
CLINK	Capacidade de Transmissão dos Enlaces, Atraso de ida e volta
IPERF	Jitter, Fração de Perda, Capacidade Disponível
NETEST	Capacidade de Transmissão do Gargalo, Capacidade Disponível, Atraso de ida e volta, Fração de Perda
NETTIMER	Capacidade de Transmissão do Gargalo
OWPING	Atraso Fim-a-Fim (com GPS), Fração de Perda
PATHCHAR	Capacidade de Transmissão dos Enlaces, Vazão, Atraso de ida e volta, Fração de Perda
PATHRATE	Capacidade de Transmissão do Gargalo
PCHAR	Capacidade de Transmissão dos Enlaces, Vazão, Atraso de ida e volta, Fração de Perda
PIPECHAR	Capacidade de Transmissão do Gargalo, Capacidade Disponível, Atraso de ida e volta, Fração de Perda
PING	Atraso de ida e volta, Fração de Perda
SPROBE	Capacidade de Transmissão do Gargalo
TRACETOUTE	Caminho, Atraso de ida e volta
TRENO	Capacidade de Transmissão do Gargalo, Capacidade Disponível

Tabela 3.1: Ferramentas de Medição Ativa.

ferramentas como as citadas acima, coletando e avaliando as métricas desejadas.

A arquitetura NIMI (*National Infraestructure Measurement Internet*) [44, 43, 42, 3] tem como objetivo criar uma arquitetura de administração descentralizada, com fortes recursos de autenticação e segurança, onde direitos para execução das medições serão fornecidos aos usuários, independente dos direitos de acesso aos pontos de medição a serem testados. Dessa forma, é possível que qualquer máquina participante da arquitetura possa ser utilizada para fazer os testes desejados com as outras, independente de direitos de acesso ao sistema local.

Capítulo 4

Cálculo do Atraso em Um Sentido

O C álculo do atraso sofrido por pacotes em um único sentido sem o uso de equipamentos especiais não é trivial. Essa medida de desempenho é também chamada de atraso fim-a-fim, ou OWD (*one-way delay*) [17]. Alguns dos problemas existentes para estimar esta métrica foram definidos pelo IETF(*Internet Engineering Task Force*) em [17], e serão descritos na seção 4.1.

Apesar da RFC(*Request for Comments*) 2679[17] descrever vários dos problemas existentes para estimar o *OWD*, não existe indicação de um algoritmo que resolva todos esses problemas. Neste trabalho, é feita uma avaliação das propostas de soluções existentes na literatura [37, 40, 31, 32, 29, 36, 41, 34]. Com a implementação de alguns desses métodos, e com a coleta de resultados através de experimentos, foi constatado um novo problema para estimar a métrica *OWD*. Além de descrever este problema, será apresentada neste capítulo a proposta de um método para solucioná-lo.

4.1 Definição dos Problemas para Cálculo da Métrica

Quando existe a disponibilidade de equipamentos, baseados em sinais de satélites para sincronizar os relógios, ou quando a origem e o destino do pacote são a mesma máquina, calcular o atraso experimentado por um pacote na rede pode ser feito de forma simples, ou seja, $D = T_r - T_s$, onde T_s é o instante de envio do pacote e T_r o instante de chegada. O problema só existe quando os equipamentos de origem e destino desses pacotes possuem relógios que não estão perfeitamente sincronizados. Sem o uso de equipamentos especiais, os relógios das máquinas não estão sincronizados, e portanto, o valor absoluto dos relógios, assim como as taxas de incremento de cada um, são diferentes. Em conseqüência, é possível obter valores negativos de retardo e ainda valores com uma tendência de crescimento ou decrescimento.

Para ilustrar o problema, foi executada uma medição entre duas máquinas com sondas geradas nos dois sentidos. Para cada sentido, uma seqüência $\Omega := [v_i = (i, d_i) : i = 1, \ldots, N]$ foi coletada durante os experimentos e essas seqüências são mostradas na Figura 4.1, onde i e d_i equivalem ao número de seqüência do pacote e o retardo calculado respectivamente.

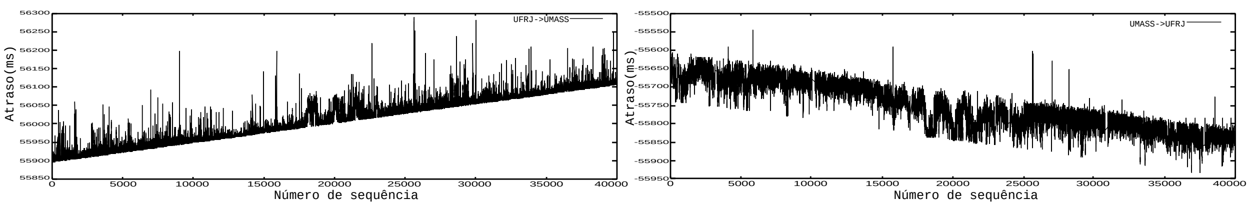

Figura 4.1: Atraso de pacotes entre máquinas com relógios diferentes.

Dois problemas ficam evidentes pela análise da Figura 4.1. O primeiro, chamado de *Offset*, aparece em consequência dos relógios nas máquinas envolvidas na medição possuírem valores distintos no início da medição. Os valores dessa diferença podem estar sendo somados ou diminuídos do valor total do atraso, resultando até mesmo em valores negativos para retardos estimados.

O segundo, chamado de *Skew*, é consequência das taxas dos relógios de máquinas diferentes possuírem taxas de crescimento distintas. Considerando relógios que não são atômicos, um microsegundo em uma máquina pode ser maior ou menor do que um microsegundo em outra. Com isso, o resultado do simples cálculo do retardo experimentado por pacotes entre duas máquinas sofre um crescimento ou decrescimento constante. Para um tempo "muito pequeno" de medição, esse fenômeno pode ser desconsiderado. No entanto, quando o experimento é executado por um tempo maior que poucos segundos, o erro causado pela diferença nas taxas de crescimento dos relógios não pode ser descartado.

4.2 Avaliação das Soluções Propostas

É importante ficar claro que os problemas de *Offset* e de *Skew* só são encontrados quando a máquina emissora difere da receptora pois, caso contrário, as medidas são calculadas com o mesmo relógio. Para possibilitar o cálculo de medidas com relógios distintos existem algumas propostas na literatura: a primeira é a instalação de equipamentos, como *GPS*, que garantem a sincronia dos relógios, assim como a sua taxa de crescimento [41]. O problema dessa solução é o custo e a dificuldade na instalação desses equipamentos, nas máquinas onde deseja-se realizar as medições; a segunda opção é a configuração de protocolos tipo NTP(*Network Time Protocol*) [34] para sincronizar os relógios. No entanto, além da configuração desses protocolos requererem acesso de administrador sobre as máquinas, e que essas máquinas estejam sincronizadas com um mesmo servidor NTP, a forma como o protocolo é definido não prevê uma atualização com a frequência necessária e ainda existe o problema de atraso para receber as informações de controle na rede. Dessa forma, a precisão dos relógios é pequena e, mesmo para uma rede local, diferenças maiores que milisegundos podem ocorrer; a terceira solução é a aplicação de algoritmos mais sofisticados que removem os valores falsos dos dados coletados. É exatamente esta solução que será avaliada nesta seção.

Uma série de algoritmos foram propostos para remover os valores inseridos pelo *Offset* e *Skew* [37, 40, 31, 32, 29, 36]. Em seguida, serão apresentadas algumas dessas propostas.

4.2.1 Algoritmos para Remoção do *Offset*

O método proposto em [31] e detalhado em [32], tem como objetivo estimar a diferença inicial entre os relógios, de máquinas distintas, envolvidos em uma medição. O método propõe que seja estimada a diferença entre os atrasos, decorrente dos tamanhos distintos dos pacotes usados na medição. Além disso baseia-se no fato do tempo de transmissão de um pacote em uma interface de rede ser proporcional ao seu tamanho, e ainda que o tempo de propagação não varia mesmo para pacotes de tamanhos distintos.

O algoritmo proposto prevê o envio de uma série de pacotes de diferentes tamanhos e em sentidos opostos, e supõe que não existe *Skew* nos valores coletados. Os pacotes que obtiverem o menor atraso, para cada tamanho usado, devem ser selecionados, como mostra a Figura 4.2A onde quatro pacotes, um de cada tamanho, foram escolhidos em cada direção. As amostras selecionadas equivalem a pacotes que supostamente não entraram em fila durante todo o caminho percorrido. Neste caso, o tempo do atraso de um pacote é igual ao tempo de propagação no caminho somado ao tempo de transmissão em cada uma das interfaces de rede encontradas.

Para pacotes que não encontraram nenhuma fila no caminho, e usando-se tamanhos distintos de pacotes obedecendo uma função linear, não é difícil mostrar que em [31] é possível estimar o *Offset* entre os relógios. O tempo de propagação na rede é estimado eliminando-se o tempo de transmissão dos pacotes. Este tempo é igual ao sofrido por um pacote de tamanho zero que não encontrou fila no caminho. Devido à linearidade dos tempos medidos, conforme mostra a Figura 4.2A, o tempo de propagação é obtido pelo ponto onde as retas cruzam o eixo das ordenadas.

- Seja, t_{AB} o tempo de transmissão de um pacote de tamanho nulo, enviado da máquina A para a B, e t_{BA} o tempo medido de B para A;
- Seja, "O" o valor do *Offset* entre os dois relógios das máquinas A e B;
- Seja, t_B o instante de recebimento do pacote de comprimento zero pela máquina B(medido no relógio de B), e t_A o instante de envio deste mesmo pacote

de tamanho zero pela máquina A (medido no relógio de A);

- Seja, t_{prop} o tempo de propagação entre as duas máquinas.

Então, $t_{AB} = t_B - t_A = t_A + O + t_{prop} - t_A = O + t_{prop}$

Da mesma forma, $t_{BA} = t'_A - t'_B = t'_B - O + t_{prop} - t'_B = -O + t_{prop}$

onde, t'_A e t'_B são os instantes de recebimento e envio do pacote enviado de B para A, medido nas máquinas A e B respectivamente.

Então, $t_{AB} - t_{BA} = 2O + t_{prop} - t_{prop} = 2O$.

Portanto, o *Offset* é obtido subtraindo-se os valores t_{AB} e t_{BA} estimados, e dividindo-se o resultado por dois.

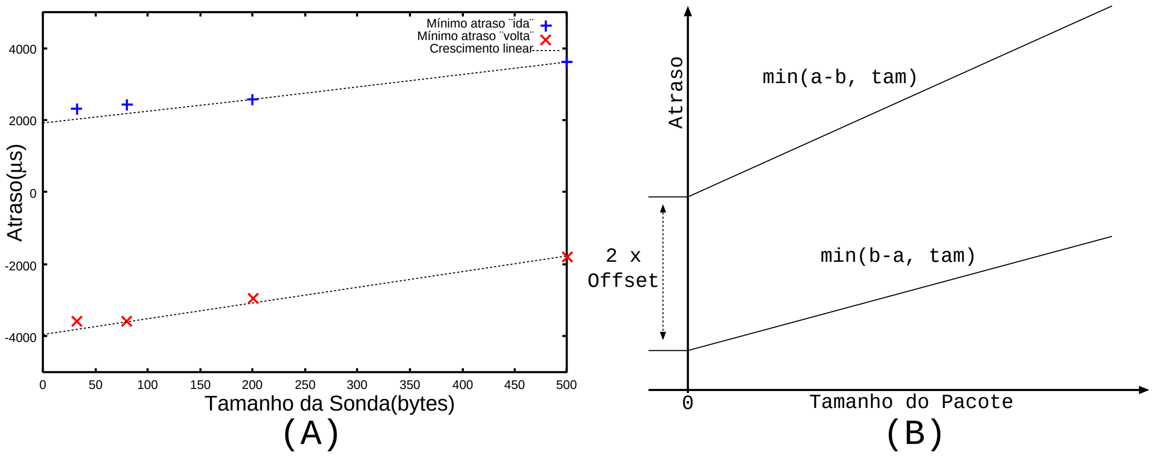

Figura 4.2: Decrescimento linear gerado pela variação do tamanho dos pacotes: (A) Decrescimento obtido por um experimento real; (B) Gráfico ilustrativo da proposta do algoritmo.

Além do algoritmo descrito acima, existem outras propostas para resolver o problema de medições entre relógios não sincronizados. No entanto, todas as outras propostas não prevêem caminhos assimétricos entre os pontos de medição. O problema dessas propostas é que as capacidades de transmissão das interfaces de rede, ao longo do "caminho de ida", têm que ser iguais às existentes no caminho de volta. O que, de fato, nem sempre pode ser considerado válido.

4.2.2 Algoritmos para Remoção do *Skew*

Várias são as propostas para solucionar o problema da diferença entre as taxas de crescimento dos relógios (*Skew*) [37, 29, 40]. O objetivo de todas as propostas é estimar uma função que represente a tendência de crescimento linear encontrada nos atrasos calculados. Para isso, é necessário encontrar uma reta que esteja imediatamente abaixo de todos os pontos da amostra.

Para resolver este problema, Moon e Towsley [40] propõem o uso de um algoritmo de programação linear para estimar a função que representa a diferença entre as taxas de crescimento nos relógios das máquinas. Em [40] os autores, além de proporem o novo método, fazem uma comparação entre esse e o proposto por Paxon, em [37]. Na avaliação dos algoritmos é demonstrado um fraco desempenho no quesito robustez por parte da proposta de Paxon, sendo verificado que em caso de altos valores do *Skew* o algoritmo falha na estimativa desse parâmetro.

Uma avaliação dos algoritmos é também apresentada em [29]. Neste caso, Liu e Zhang provam que sua proposta possui uma menor complexidade computacional do que a feita por Paxon, e menor ou igual do que a proposta de Moon e Towsley, além de mais intuitiva do que esta última.

Na proposta de Liu e Zhang é usado o conceito de fecho convexo para encontrar a melhor função que represente a diferença entre as taxas de crescimento dos relógios. Dado Ω para representar a coleta, onde $\Omega := [v_i = (i, d_i)|i = 1, \ldots, N]$ (conforme definido anteriormente, na Seção 4.1), o objetivo é estimar o limite inferior do fecho convexo definido para os pontos v_i da coleta Ω.

O fecho convexo, FC, de um conjunto de pontos consiste no menor polígono convexo formado por um subconjunto desses pontos, onde todos os outros pontos deste conjunto se encontram na parte interior do polígono. Os pontos pertencentes ao subconjunto equivalem aos vértices do polígono definido. O limite inferior de um fecho é formado pelos vértices inferiores do polígono entre o ponto de menor valor na dimensão "x" até o ponto de maior valor na dimensão "x".

A função do algoritmo proposto é identificar o limite inferior do fecho convexo de uma coleta, FC(Ω). Para isso, o algoritmo faz uma análise de todos os pontos (v_i) da coleta seguindo a ordem crescente, definida pelo número de seqüência das sondas. Para cada ponto o algoritmo determina: (i) se este v_i deve ser incluído na "pilha de vértices"; (ii) se este deve ser descartado; (iii) ou se outros vértices anteriormente empilhados devem agora ser retirados dessa pilha. Após analisado o n-ésimo ponto, todos os pontos presentes na pilha representam os vértices do limite inferior do fecho convexo definido para aquela coleta. O resultado após a aplicação do algoritmo em uma coleta é ilustrado na Figura 4.3A.

Algoritmo Fecho_ Conv_ Inf:

```
(1) Begin
       empilha V1;
       empilha V2;
(2) For i=3 to N
       if(Vi above line(topo_pilha, abaixo_topo_pilha))
          empilha Vi;
       else
          while(Vi below line(topo_pilha, abaixo_topo_pilha)and(tam_pilha >1))
               desempilha;
          empilha Vi;
(3) End;
```

Nota: *line*(w_1, w_2) representa uma reta que passa pelos pontos w_1 e w_2. A função empilha insere o vértice (V1) na pilha e atualiza as variáveis topo_pilha e abaixo_topo_pilha. A função desempilha remove o vértice que se encontra no topo da pilha, e atualiza as variáveis topo_pilha e abaixo_topo_pilha.

Dentre todos os pontos que, ao final do algoritmo, se encontram na pilha, é identificada a "melhor" reta para representar a tendência de crescimento existente na coleta. Para isso, os autores em [29] identificam o ponto médio da coleta (no caso de sondas *CBR* é dado por $N/2$, onde N é o maior número de seqüência gerado). Dois casos podem ocorrer: (i) o ponto médio é um dos vértices do polígono. Neste caso, a melhor reta é definida com origem no ponto médio e terminando no próximo vértice à direita do ponto médio (ver Figura 4.3B); (ii) o ponto médio não é um vértice do polígono. Neste caso, a melhor reta é aquela que contém o ponto médio (ver Figura 4.3C). Encontrada a reta e a sua função f(x) o valor de α equivale à inclinação desta reta em relação ao eixo das abcissas. Este α representa a taxa de crescimento de um relógio em relação ao outro.

Seja $y = f(x)$ a reta estimada pelo algoritmo do Fecho Convexo.

Sejam $V_{skew}(x_1, y_1)$ e $V'_{skew}(x_2, y_2)$ os vértices presentes na pilha do fecho convexo, e dois pontos distintos da reta $f(x)$.

A inclinação desta reta (ou coeficiente angular) é dada por:

$\alpha = \frac{y_2 - y_1}{x_2 - x_1}$

Esse valor de α equivale à diferença entre as taxas de crescimento dos relógios.

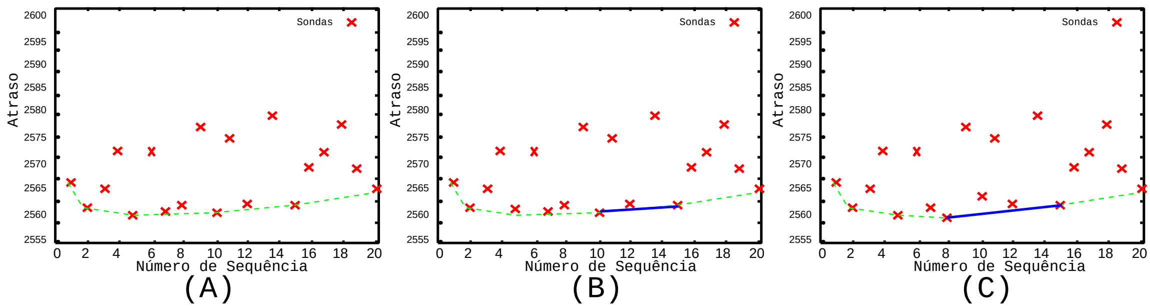

Figura 4.3: Funcionamento dos algoritmos para remoção do Skew: (A) Identificação do fecho convexo entre os pontos da amostra; (B) Identificação da reta f(x) para o caso (i); (C) Identificação da reta f(x) o caso (ii).

Com a estimativa da diferença entre a taxa de crescimento de um relógio em relação ao outro, remover o valor inserido pelo crescimento assimétrico dos relógios é relativamente simples. Para isso basta identificar o tempo que passou desde o início da coleta até a emissão da sonda que se deseja remover o *Skew* e multiplicar esse tempo pela taxa de crescimento estimada. O valor obtido desse cálculo é removido do retardo inicial calculado, gerando assim um novo valor para o atraso da sonda. Esse processo é repetido para as "n" sondas e resulta em um novo *trace*, porém, agora sem o valor do *Skew*. É importante perceber que, se no exato momento do início da medição os dois relógio não se encontrarem sincronizados, este "novo *trace*" contém o *Offset* inicial da coleta. No entanto, técnicas para remover o *Offset* já foram vistas e podem ser aplicadas.

Algoritmo Remove_Skew:

```
(1) Begin
     Instante_e_sonda_1  \\Instante de emissão da primeira sonda
     alfa                \\Diferença entre as taxas de crescimento
(2) For i=1 to N
     atraso_sem_skew_i=atraso_i-((Instante_e_sonda_i -Instante_e_sonda_1)*alfa)
(3) End;
```

4.2.3 Implementações na Ferramenta e Resultados Obtidos

As propostas [29], para remoção do *Skew*, e [31] para remoção do *Offset*, foram implementadas e integradas ao *TANGRAM-II Traffic Generator*. Com essa ferramenta é possível executar experimentos em redes IP e coletar medidas, até então não disponíveis do nosso conhecimento em qualquer outra ferramenta de medição ativa, que não requer dispositivos para sincronizar os relógios.

Geração e Coleta das Sondas na Ferramenta

Para possibilitar a estimativa da métrica, o tráfego gerado para fornecer os dados para os algoritmos implementados devem seguir os padrões definidos por esses algoritmos. Como foi visto na Seção 3.2, a ferramenta já possuía um gerador de tráfego flexível em sua primeira versão. Por isso, foi necessário apenas uma adaptação para se adequar às necessidades dos novos algoritmos implementados na ferramenta.

A geração do tráfego para estimar o atraso em cada sentido é feita nas duas direções e com tamanho de pacotes variando entre: 40, 80, 200, e 500 (*bytes*). Esses valores foram escolhidos para prevenir a segmentação dos pacotes, evitando a possibilidade de erros nos cálculos da métrica. Apesar da grande maioria dos pacotes nas redes IP's manterem o tamanho máximo de 1500 *bytes*, existem MTU's bem menores que podem chegar a 576 *bytes*.

Os pacotes são enviados em intervalos constantes de tempo para garantir maior precisão possível na obtenção da medida. Detalhes serão apresentados na Seção 4.3.

Algoritmos Implementados na Ferramenta

Após geradas e coletadas as sondas, os algoritmos são aplicados aos *traces* e, então, os valores referentes ao *Skew* e ao *Offset* são removidos. A Figura 4.4A ilustra a coleta inicial de um experimento que, além da tendência de crescimento, atribui valores bastante elevados para o retardo dos pacotes, tendo neste exemplo valores superiores a 25 segundos, muito maior do que os valores reais.

Com a aplicação dos algoritmos, um novo resultado é gerado, agora com os valores dos atrasos reais sofridos pelos pacotes. O resultado obtido, após remoção do *Skew* e do *Offset*, é mostrado na Figura 4.4B. É importante perceber que os valores atribuídos aos atrasos dos pacotes podem agora ser considerados aceitáveis para uma métrica como "Atraso em um sentido", diferente dos valores indicados antes da aplicação dos algoritmos. No entanto, isso não demonstra a validade deste resultado. A validação para esses resultados será discutida no Capítulo 5.

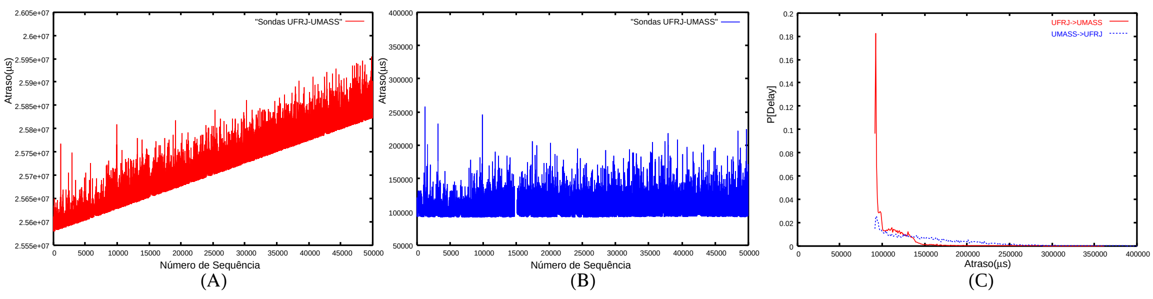

Figura 4.4: Resultado dos experimentos de medição do Atraso fim-a-fim: (A) Antes da aplicação dos algoritmos; (B) Valores reais do atraso, após aplicação dos algoritmos; (C) PMF's dos valores reais do atraso.

A Figura 4.4C ilustra as PMF's obtidas dos valores estimados com a métrica "Atraso em um sentido". Nesta figura, são mostradas as curvas dos histogramas, obtidas a partir dos resultados nos dois sentidos do experimento.

Resultados com a Ferramenta

Com a implementação da ferramenta, uma série de experimentos foram executados entre o LAND (Laboratório de Modelagem, Análise e Desenvolvimento de Redes e Sistemas de Computação), na UFRJ (Universidade Federal do Rio de Janeiro), e a UMASS (*University of Massachusetts*). Alguns desses resultados são mostrados na Figura 4.5. O objetivo desses experimentos é fazer um estudo do comportamento do atraso em um sentido de tráfego gerado na Internet. Este estudo será apresentado com mais detalhes no Capítulo 5.

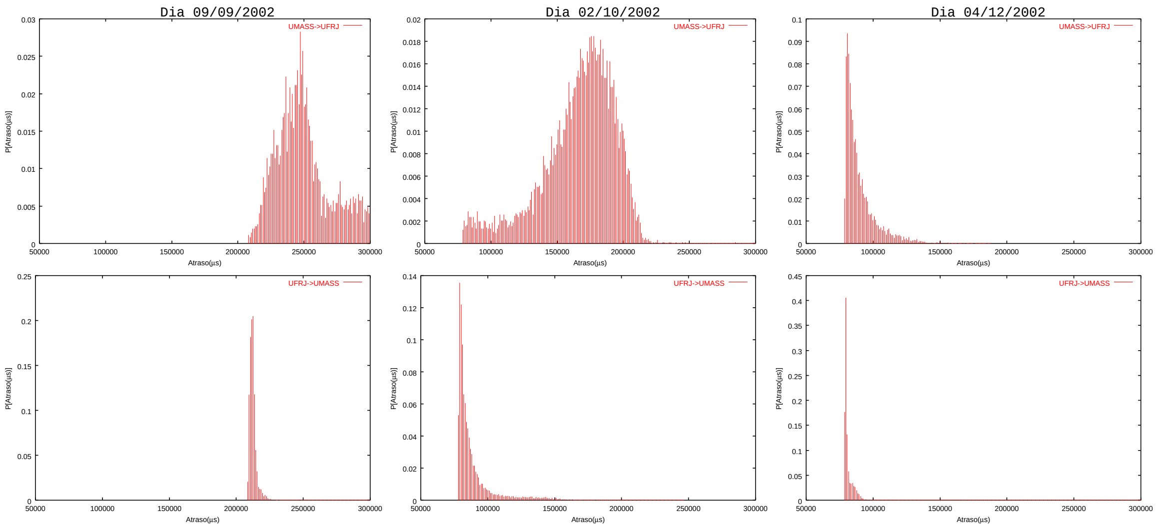

Figura 4.5: Resultados obtidos com a ferramenta *TANGRAM-II Traffic Generator* em experimentos entre a UFRJ e a UMASS.

4.3 Outros Problemas no Cálculo do Atraso

Durante a análise dos resultados obtidos com os experimentos, alguns outros problemas, além do *Skew* e do *Offset*, para estimar a métrica "Atraso em um sentido" foram detectados. Na próxima seção, esses problemas serão descritos, assim como o impacto deles sobre a estimativa da métrica. Um novo método, que propõe uma solução para este problema, será apresentado neste capítulo.

4.3.1 Descrições dos Problemas de Atualizações e Ruídos

Dentre os vários experimentos executados entre a UFRJ e a UMASS, pelo menos metade detectaram problemas que podem dificultar a estimativa da métrica de interesse. Esses problemas foram causados por atualizações nos relógios das máquinas, utilizadas para a medição, ou por eventuais congelamentos no sistema operacional desses equipamentos. Os dois problemas citados são ilustrados na Figura 4.6.

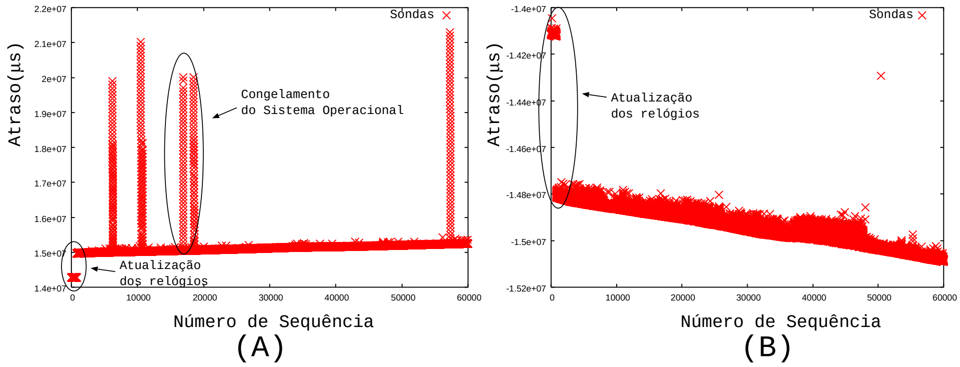

Figura 4.6: Atraso de pacotes entre máquinas com relógios diferentes, sofrendo atualização nos relógios e congelamento do sistema operacional durante as coletas.

Atualização nos Relógios

A atualização nos relógios causa, aos resultados calculados para atraso das sondas, uma alteração "permanente" nos valores estimados, isto é, que se propaga pelos valores subsequentes. Na Figura 4.6, é possível verificar que logo no início do experimento, próximo ao envio da sonda de número 800, ocorre uma atualização no relógio de uma das máquinas. Com isso, o resultado médio do atraso sofre uma alteração abrupta que permanece até o final da medição nas duas coletas.

Duas são as conseqüências causadas pelo problema de atualização dos relógios: a primeira é a possibilidade de erro na estimativa da função "*Skew*" para remoção dos valores inseridos pela diferença nas taxas de crescimento. A Figura 4.7 ilustra esse problema, onde a estimativa incorreta da "função convexa" na coleta ocasiona a remoção incorreta da tendência de crescimento pelo algoritmo. A segunda conseqüência é que, a partir do ponto da atualização de um dos relógios, a estimativa do *Offset* passa a ser inválida, pois a diferença inicial existente entre os relógios não é mais a mesma.

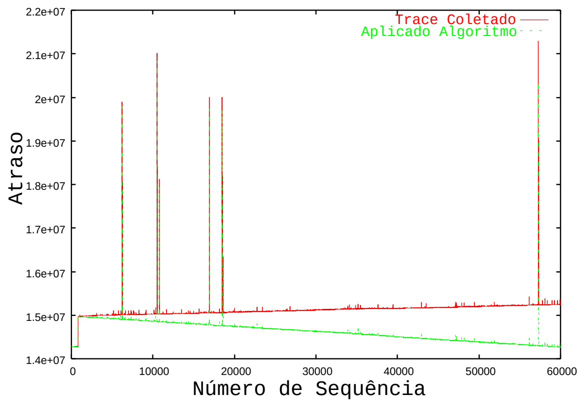

Figura 4.7: Remoção incorreta da tendência de crescimento, causada pelo erro na identificação da função convexa dos pontos.

Congelamento do Sistema

Diferente do problema causado pela atualização dos relógios, o congelamento do sistema operacional em uma das máquinas gera nos resultados grandes picos em apenas uma das direções. Na Figura 4.6A, são identificados, pelo menos, cinco grandes picos no atraso sofrido pelas sondas. Esses picos são considerados ruídos nos resultados do experimento.

A seguir descreveremos um cenário onde ocorreu o congelamento do relógio diversas vezes durante os experimentos realizados. Durante vários meses, aulas e seminários foram apresentados em parceria entre a UFRJ e a UMASS, usando uma ferramenta de vídeo em tempo real desenvolvida para educação a distância no LAND [14]. A ferramenta de vídeo em tempo real e o *TANGRAM-II Traffic Generator* foram executados na mesma máquina com o objetivo de possibilitar a remoção do *Skew* e *Offset* da sequência de vídeo a partir das amostras geradas pelo *TANGRAM-II Traffic Generator* (Maiores detalhes sobre estes experimentos podem ser encontrados na Seção 4.4).

Uma das funcionalidades desta aplicação de vídeo previa a sincronização entre a aplicação e a placa de captura de imagem. No entanto, toda vez que essa funcionalidade é executada, o sistema operacional é congelado por poucos microsegundos. Esse evento é pouco perceptível ao usuário, porém, significativo para uma ferramenta executando em paralelo. Neste caso, o *TANGRAM-II Traffic Generator* para de enviar suas sondas por alguns instantes, assim como para de recebê-las. A conseqüência do evento são os picos gerados nos resultados, pois no instante em que o sistema operacional volta a funcionar normalmente, várias sondas enviadas pelo módulo remoto durante o congelamento, já se encontram na memória da interface de rede para serem capturadas.

Os ruídos não geram qualquer conseqüência sobre o cálculo da métrica "Atraso fim-a-fim", no entanto influenciam nas análises dos resultados. Uma dessas análises, que será vista com mais detalhe no Capítulo 5, é a tentativa de identificar uma distribuição que represente bem o comportamento do atraso sofrido pelas sondas. Os ruídos, conseqüentes do congelamento do sistema, fazem com que no histograma resultante apareçam valores altos para o retardo com probabilidades baixas. Esses valores altos para o retardo do histograma podem indicar uma cauda longa da distribuição, sem que esse comportamento tenha de fato existido.

4.3.2 Proposta de Um Método para Resolução dos Problemas Acima

Fica claro que a presença dos eventos, atualização dos relógios e congelamento do sistema operacional, pode atrapalhar na obtenção da métrica "Atraso em um sentido". Neste trabalho, um método foi desenvolvido com o intuito de resolver os problemas gerados por esses eventos.

O método proposto neste trabalho requer que o intervalo entre a geração das sondas seja constante. A emissão determinística dos pacotes permite verificar em um gráfico gerado a partir do número de seqüência dos pacotes e seus instantes de envio, uma reta crescente e linear. A Figura 4.8 ilustra um exemplo deste gráfico, onde o eixo das abcissas representa o número de seqüência e o eixo das ordenadas representa o instante de envio desses pacotes, com o intervalo entre geração constante e igual a 100 milisegundos. Qualquer variação significativa na linearidade desta reta pode ser justificada por um dos dois eventos citados. Em seguida será exposto como o método se propõe a remover os valores inseridos pelas atualizações e os ruídos gerados por congelamentos no sistema operacional.

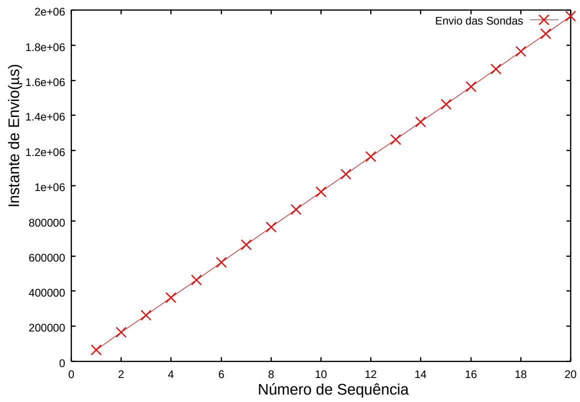

Figura 4.8: Crescimento linear da reta de geração das sondas.

Remoção de Atualização nos Relógios

O problema causado pela atualização dos relógios já foi tratado em dois outros trabalhos. Em [37], Paxon propõe um método para remover o erro devido a esta atualização. No método proposto em [29] os valores não são removidos, porém é definido um algoritmo para estimar o *Skew* em coletas que contenham atualizações dos relógios. Os métodos [29] e [37] são baseados na variação abrupta dos atrasos calculados no receptor, diferente do proposto neste trabalho, que avalia tanto a variação do atraso como a variação na taxa de envio das sondas.

Para identificar a atualização dos relógios, o algoritmo de Paxon[37] faz uma análise simultânea dos atrasos obtidos para cada uma das direções. O algoritmo supõe que houve uma atualização nos relógios quando são identificadas mudanças abruptas para os valores do atraso nos dois sentidos em instantes próximos de tempo (como mostrado na Figura 4.6). O problema desta abordagem é que: uma alteração significativa na média do atraso da rede, nos dois sentidos e em instantes próximos de tempo, pode ser considerada uma atualização nos relógios levando a resultados incorretos para a estimativa do atraso. Além disso, a proposta de Paxon assume que na coleta não existe *Skew*.

A idéia básica do algoritmo proposto em [29] para a remoção do *Skew* é estimar o fecho convexo não considerando os valores de retardo onde ocorreu uma atualização no relógio. Para isso os autores propõem que o intervalo de coleta seja dividido em sub-intervalos. Se os valores de retardo obtidos em um mesmo sub-intervalo diferirem de um valor superior a um certo limite, então este sub-intervalo não será considerado na estimativa do fecho convexo. O algoritmo de [29], assim como o de Paxon, baseia-se apenas nos resultados do atraso. A consequência disso é que as alterações abruptas no comportamento dos atrasos pode ser considerada uma atualização do relógio.

Com os problemas encontrados nas propostas [29] e [37], um novo paradigma é apresentado para fazer a remoção da atualização dos relógios. O método proposto se baseia na análise da reta gerada pelo envio dos pacotes. A atualização no relógio da máquina faz com que o "carimbo do instante" de emissão da sonda sofra uma variação proporcional à gerada pela atualização do relógio. Desta forma, uma variação na linearidade pode significar uma indicação de atualização no relógio. A Figura 4.9 ilustra a variação ocorrida pela atualização ilustrada anteriormente na Figura 4.6. Porém, nem sempre a alteração na reta caracteriza a atualização dos relógios. Caso um dos emissores fiquem sem enviar temporariamente, por exemplo por conseqüência de congelamento do sistema operacional, a reta de envio também apresentará uma variação na linearidade. Por isso os valores coletados para os atrasos na máquina receptora também devem ser verificados.

Detectada a variação na reta de envio das sondas superior a um fator estipulado no algoritmo (em nossos experimentos foi utilizado o valor $\Gamma = 0.1$), uma análise é feita para verificar se houve alteração significativa, também, no valor do atraso dessa sonda. Essa alteração é considerada significativa quando é superior à diferença relativa à atualização do relógio. Uma correção é aplicada às coletas nos dois sentidos, independente da direção em que foi identificada a diferença na taxa de envio. Um valor proporcional ao inserido pela atualização do relógio, medido na reta de envio, será removido dos atrasos calculados para aquelas sondas. A Figura 4.11 ilustra os

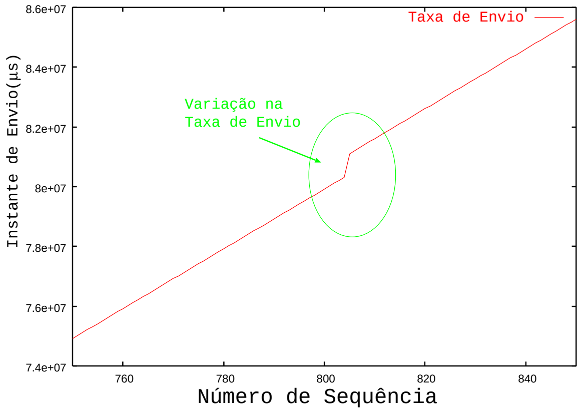

Figura 4.9: Alteração na taxa de envio das sondas gerada pela atualização do relógio.

resultados para o atraso em um sentido após a remoção dos valores da atualização dos relógios.

O algoritmo proposto se encontra detalhado a seguir.

Seja:

- "e" vetor que contém o instante de envio das sonda na máquina emissora;
- "re(l)" vetor que contém o retardo calculado na máquina emissora;
- "rr(i)" vetor que contém o retardo calculado na máquina receptora;
- "t" intervalo entre o envio de sondas que é constante;
- "gamma" variação percentual máxima permitida para o tempo entre envios;
- "N" número de amostras de retardo coletadas;
- "diff" diferença do instante real de envio da sonda em relação ao instante que a sonda deveria ter sido enviado;
- "Tre" vetor que contém o instante de recebimento dos pacotes na máquina emissora.

Algoritmo Remove_Atualização:

```
(1) Begin
      gamma
(2) For i=1 to N
      if (|e(i+1)-e(i)| >= t*(1+gamma))
         diff = |e(i+1)-e(i)| - t;
        if (rr(i+1) > diff)
             For j=i+1 to N
                rr(j) = rr(j) - diff;
             EndFor;
             For k=1 to N
               if ((Tre(k) > e(i))and(re(k) > diff))
                   re(k) = re(k) - diff;
               EndFor;
    EndFor;
(3) End;
```

Remoção de Ruídos

A idéia usada para remoção dos ruídos presentes em uma coleta é basicamente a mesma da proposta para remover o erro causado pela atualização dos relógios. A Figura 4.10 ilustra uma situação de congelamento do sistema operacional, vista anteriormente na Figura 4.6A. No entanto, diferentemente da atualização dos relógios, não existe uma variação na média do atraso dos pacotes nas duas direções, mas sim uma mudança abrupta nos atrasos das sondas recebidas pela máquina onde houve o congelamento. Este problema já foi descrito anteriormente e é causado pela parada temporária de envio de tráfego por uma das máquinas, enquanto a outra continua a emissão de suas sondas. São exatamente as sondas que chegaram ao receptor no instante em que ele se encontrava congelado, que possuem uma variação considerável no cálculo do seu atraso, mas em um curto período de tempo esse retardo volta a ficar bem próximo da média. É importante verificar que os dois eventos que indicam uma variação na taxa de envio identificados na Figura 4.10 estão relacionados a picos mostrados na Figura 4.6A, porém, pelo tamanho da escala em que a figura foi gerada esses picos aparecem quase que sobrepostos.

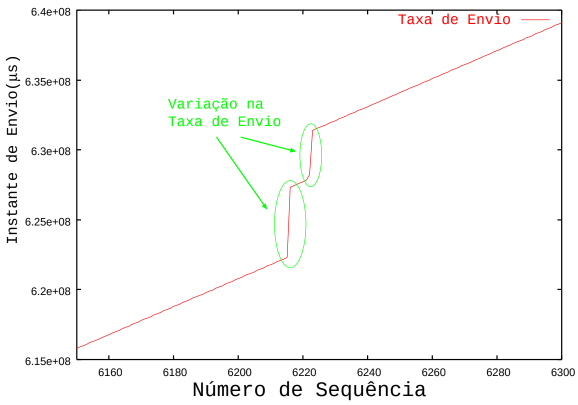

Figura 4.10: Alteração na taxa de envio das sondas gerada pelo congelamento do sistema operacional.

Após a identificação do instante onde ocorre o congelamento, é estimado o número de sondas que chegaram da máquina remota durante o período de congelamento. O atraso calculado para essas sondas são as amostras responsáveis pelos picos apresentados no *trace*. Para garantir que ruídos não serão considerados nos resultados, o dobro do número de sondas enviadas pela máquina remota, desde o inicio do congelamento serão descartadas. A Figura 4.11 ilustra o resultado da coleta após removidos os ruídos.

O algoritmo proposto se encontra detalhado a seguir.

Considerando as definições feitas para o algoritmo de remoção de atualização. Seja:

- "s(d)" Número de sondas enviadas em um intervalo de tempo "d";
- "amostra[m]" Pontero para os vetores de instante de envio e retardo da sonda "m";

Algoritmo Remove_Ruido:

```
(1) Begin
      gamma
```

```
(2) For i=1 to N
      if (|e(i+1)-e(i)| >= t*(1+gamma))
         diff = |e(i+1)-e(i)| - t;
        if (rr(i+1) > diff)
            "Executar Algoritmo Remove_Atualização"
         else
             For l=1 to N
               if (Tre(l) > e(i))
                   d = 2*diff;
                   For m=l to l+s(d)
                       Amostra[m] = NULL;
                   EndFor;
                   BreakFor;
             EndFor;
    EndFor;
(3) End;
```

4.3.3 Resultados com o Algoritmo Proposto

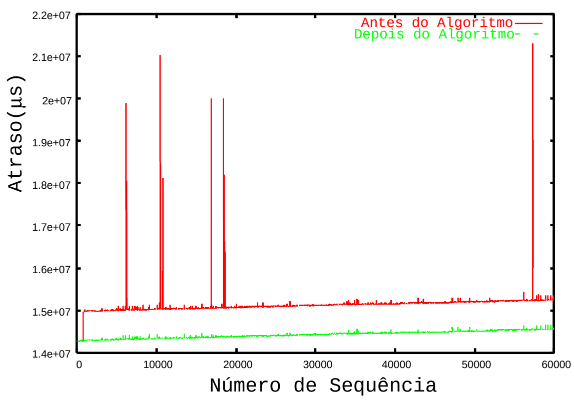

Figura 4.11: Resultado obtido após aplicação dos algoritmos de remoção de atualização dos relógios e ruídos.

Como é ilustrado na Figura 4.11, após a remoção dos valores inseridos pela atualização do relógio e dos ruídos gerados pelo congelamento do sistema, a coleta passa a ser semelhante a uma que não tenha sofrido os mesmos problemas. Os resultados, até então, não são os verdadeiros atrasos experimentados pelas sondas, nesses valores ainda constam os problemas inseridos pelo *Skew* e *Offset*. Porém, agora os algoritmos apropriados já podem ser aplicados para a obtenção dos resultados corretos do retardo sofrido pelo tráfego. Os resultados obtidos após a remoção do *Skew* e *Offset*, e a PMF dos atrasos estimados estão ilustrados nas Figuras 4.12 (A, B e C).

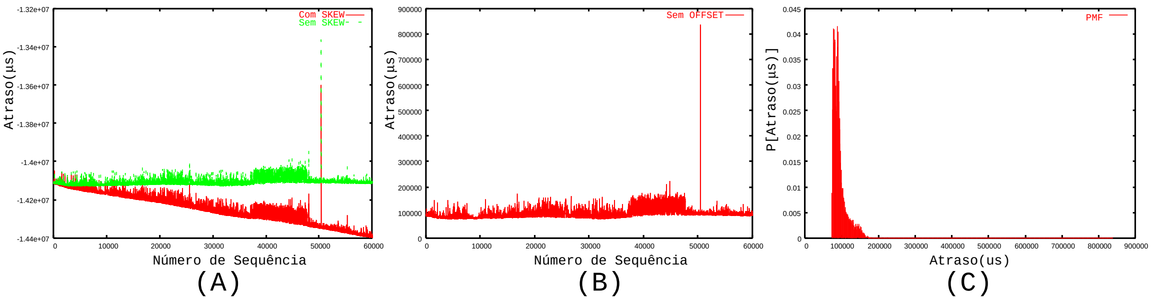

Figura 4.12: Resultado obtido após aplicação dos algoritmos de remoção de *Skew* e de *Offset*: (A) Resultado após aplicação do algoritmo para remoção do *Skew*; (B) Resultado após aplicação do algoritmo para remoção do *Offset*; (C) PMF do atraso.

4.4 Cálculo do Atraso para Aplicações

Os algoritmos apresentados nas seções anteriores exigem que as sondas sejam geradas de acordo com um padrão: geração das sondas em intervalos constantes, tráfego gerado nos dois sentidos, e pacotes com tamanhos pré-definidos e menores do que 500 *bytes*.

Muitas vezes pode ser bastante útil a obtenção do retardo fim-a-fim para uma determinada aplicação sendo executada. Suponha uma aplicação de transmissão de vídeo em tempo real que possua um mecanismo de adaptação para compensar o retardo na rede. Esse mecanismo pode usar estimativa de retardo coletadas dinamicamente ao longo da transmissão de vídeo, de forma a oferecer uma melhor *QoS* para os usuários.

São claros os problemas para estimar a medida de desempenho "Atraso em um sentido" para uma aplicação qualquer cujo padrão de geração de informações seja diferente do usado pelos algoritmos definidos nas seções anteriores. No entanto, pode-se gerar em paralelo à aplicação sondas obedecendo aos requisitos dos algoritmos e usar os resultados de *Skew* e *Offset* obtidos desta medição para estimar o atraso da outra coleta. Essa medição não necessariamente deve ser aplicada em paralelo, mas sim em um intervalo próximo de tempo (antes, durante ou depois) e por um tempo suficientemente grande (cerca de alguns segundos) para que se possa estimar os valores de *Skew* e *Offset*. A Figura 4.13A ilustra as coletas geradas em instantes próximos de tempo. As curvas vermelha e verde representam a medição aplicada "nos dois sentidos", que cumpre todas as regras necessárias para a estimativa da medida atraso em um sentido: geração nos dois sentidos, tamanhos definidos para as sondas e intervalos constantes de tempo. A curva azul representa a coleta do tráfego da aplicação de vídeo, gerada "em um sentido", que não se enquadra em qualquer uma das regras exigidas pelos algoritmos para estimar o atraso fim-a-fim.

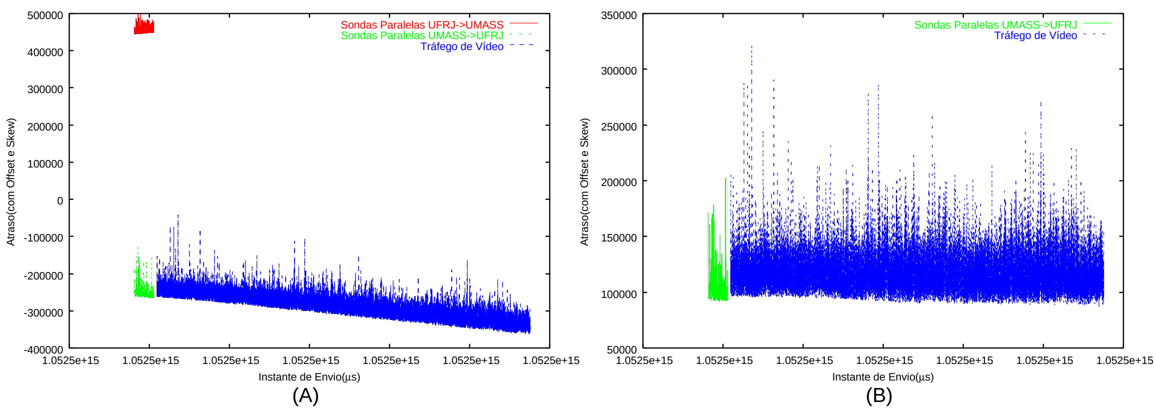

Figura 4.13: Atraso calculado para um tráfego em um sentido, paralelo às sondas de medição nos dois sentidos:(A) com *Skew* e *Offset*; (B) após remoção de *Skew* e de *Offset*.

Para chegar aos resultados verdadeiros da métrica atraso sofrido pelo tráfego em um sentido, o método utiliza parâmetros estimados a partir da medição "nas duas direções". Com isso, o algoritmo implementado deste método requer como entrada, além do *trace* gerado pelo tráfego em um sentido, parâmetros estimados da medição "nos dois sentidos", como: a diferença inicial entre os relógios (*Offset*); a taxa de crescimento de um relógio em relação ao outro (*Skew*); e o menor valor encontrado no *trace*, que toque a "função *Skew*" (o vértice da função) estimada pelo fecho convexo.

O algoritmo adapta a coleta "em um sentido" aos parâmetros estimados durante a medição nos dois sentidos. Para isso, ao invés de serem usados os parâmetros (*Skew* e *Offset*) gerados a partir desta amostra "em um sentido", são usados os parâmetros estimados pela coleta "nos dois sentidos". Assumindo que a diferença entre as taxas de crescimento dos relógios (*Skew*) não se alteraram durante todo o experimento (suposição razoável uma vez que as coletas são feitas em períodos próximos de tempo), o *Skew* calculado para a medição "nos dois sentidos" é utilizado para remover os valores causados pela diferença da tendência de crescimento, também, na coleta "em um sentido". Com a remoção do *Skew* dessa forma, o *Offset* inicial presente no experimento "em um sentido" passa a ser igual à estimada pelo experimento "nos dois sentidos". Então, para estimar o atraso sofrido pelos pacotes em um único sentido, basta remover o *Offset* calculado para a medição "nos dois sentidos". O resultado obtido, após a aplicação deste algoritmo na coleta do tráfego de vídeo, é ilustrado na Figura 4.13B.

A implementação deste módulo, na ferramenta *TANGRAM-II Traffic Generator*, possibilita que sejam estimadas métricas relacionadas ao atraso em um sentido, a partir de qualquer modelo de geração dos pacotes. O Apêndice A descreve com mais detalhe as características atuais da ferramenta. No entanto, com a utilização deste algoritmo é possível estimar o atraso sofrido pelos pacotes, enviados por qualquer modelo de geração de tráfego presente na ferramenta, ou qualquer outra ferramenta, desde que sejam fornecidos *traces* com as informações requeridas pelo algoritmo. Para isso, basta que a geração "nos dois sentidos" seja aplicada paralelamente.

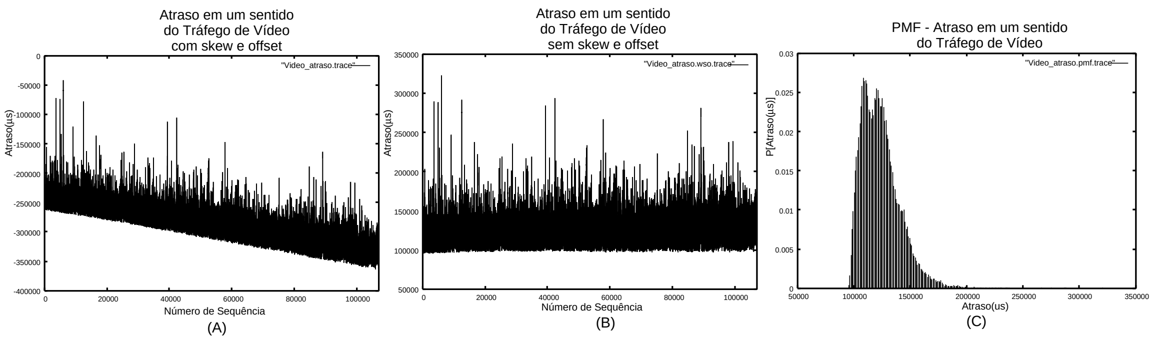

Figura 4.14: Medição obtida de um tráfego de vídeo: (A) com *Skew* e *Offset*; (B) após remoção de *Skew* e de *Offset*; (C) PMF do Atraso.

Durante os experimentos executados entre a UFRJ e a UMASS, a ferramenta utilizada para transmissão do vídeo estava configurada para gerar informações dos pacotes. Dessa forma, os algoritmos implementados no *TANGRAM-II Traffic Generator* foram utilizados para avaliar os atrasos sofridos pelos pacotes enviados pela ferramenta de vídeo. Um dos resultados obtidos nesses experimentos está ilustrado na Figura 4.14. Na figura é possível verificar o atraso sofrido por esses pacotes e os resultados obtidos após a aplicação dos algoritmos.

Capítulo 5

Validação, Experimentos e Resultados

NESTE capítulo é apresentada a validação dos algoritmos implementados na ferramenta *TANGRAM-II Traffic Generator*. É descrito um módulo que estima a distribuição que melhor se adequa ao Atraso fim-a-fim e também são descritos alguns dos experimentos realizados e seus respectivos resultados.

5.1 Validação das Implementações na Ferramenta

A idéia básica da validação é comparar a distribuição do retardo de "ida e volta" com a distribuição da soma dos retardos fim-a-fim obtidos em cada um dos sentidos da medição sendo realizada. Para isso, um módulo foi implementado para medir a distribuição e a média do "Atraso de ida e volta" *(RTT)* das sondas. Este módulo envia sondas a uma máquina de destino que os replica de volta para a origem. Semelhante à ferramenta PING [35], este módulo mede o tempo que as sondas levam para sair e voltar à máquina de origem. Após a coleta das amostras, a distribuição e a média do *RTT* são estimadas. A Figura 5.1 ilustra a *PMF* do *RTT* obtida a partir de experimentos realizados entre a UFRJ e a *UMASS*.

Várias foram as motivações para o desenvolvimento de uma ferramenta para

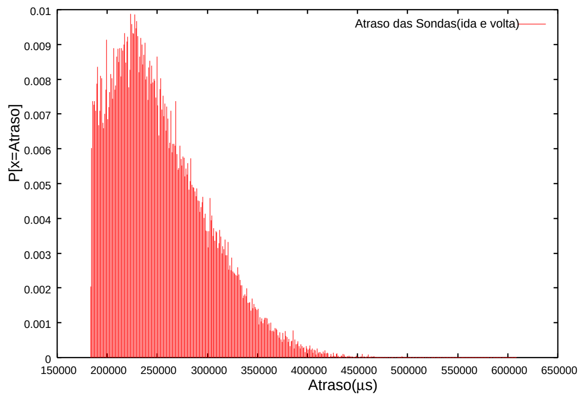

Figura 5.1: Atraso de "ida e volta" das sondas.

medir o retardo de ida e volta, ao invés de usar alguma já existente, como PING. Dentre as motivações principais estão: (i) a geração dos resultados no formato desejado e no padrão semelhante ao definido pela ferramenta; (ii) o cálculo de outras métricas, como distribuição e *PMF* do retardo, com as coletas; (iii) a possibilidade de enviar pacotes com intervalo de gerações muito curto, pois ferramentas com a mesma finalidade não fornecem alta precisão entre o envio dos pacotes (em geral os intervalos entre gerações não podem ser menores do que alguns milisegundos).

Foram realizados experimentos que consistiram na emissão das sondas para medição de atrasos "em um sentido" e "ida e volta". Como as sondas foram geradas ao mesmo tempo, o retardo obtido para a i-ésima sonda gerada para estimativa do atraso ida e volta deve ser semelhante à soma dos retardos obtidos para as i-ésimas sondas geradas para estimar o atraso em um sentido.

Na Figura 5.2 são mostrados os resultados obtidos com esses experimentos. As duas curvas da Figura 5.2A representam as funções probabilidade de massa dos atrasos dos pacotes em cada uma das direções. A curva ilustrada na Figura 5.2B representa a função probabilidade de massa do retardo de ida e volta das sondas.

Com os gráficos ilustrados na Figura 5.2 é possível verificar que a média do retardo de "ida e volta" dos pacotes está bem próxima da soma das médias obtidas pelas curvas em cada um dos sentidos. No entanto, as *PMF's* obtidas para o retardo em cada um dos sentidos são bem diferentes. Este resultado ilustra a importância da estimativa da distribuição do atraso em cada um dos sentidos da comunicação.

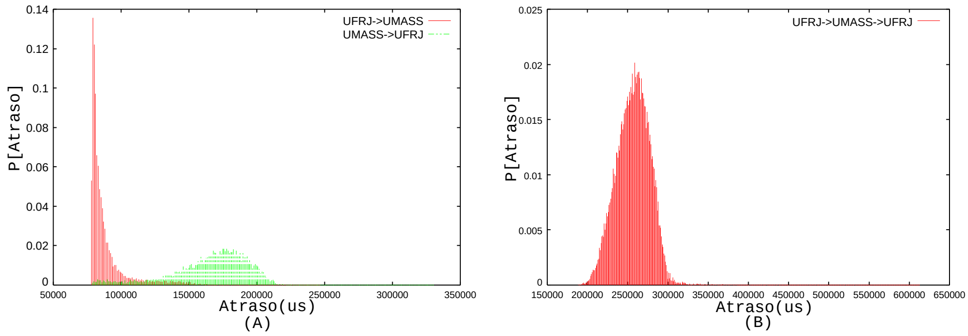

Figura 5.2: PMF's estimadas nas medições: (A) PMF's estimadas dos atrasos das sondas, em cada um dos sentidos; (B) PMF do atraso de ida e volta das sondas.

Obviamente o parâmetro média não é suficiente para demonstrar a validade das estimativas da métrica "Atraso em um sentido"AUS. Usaremos a distribuição da soma das variáveis aleatórias AUS.

Seja r_{ab} a variável aleatória que representa o AUS entre duas máquinas "A" e "B", no sentido de "A" para "B", e seja r_{ba} a variável aleatória que representa o atraso no sentido de "B" para "A". Obviamente as variáveis aleatórias r_{ab} e r_{ba} são contínuas, mas serão discretizadas no tempo para o cálculo da convolução. Desta forma, seja $p_{r_{ab}}(x) = P\left[x < r_{ab} \leq x + \Delta x\right]$ a função probabilidade de massa (PMF) de r_{ab} discretizada e $p_{r_{ba}}(x)$ a PMF de r_{ba}.

Para variáveis aleatórias discretas, a PMF da soma r_{ab} com r_{ba} é definida como 5.1:

$$p_{r_{AB}+r_{BA}}(t) = \sum_{x=0}^{t} p_x(x) p_y(t-x) \tag{5.1}$$

A distribuição é dada por 5.2:

$$F_{r_{AB}+r_{BA}}(t) = \sum_{x=0}^{t} F_x(x) F_y(t-x) \tag{5.2}$$

A PMF da soma dos "Atrasos em um sentido", assim como a PMF do "Atraso de ida e volta", estão ilustradas na Figura 5.3A. A precisão das estimativas pode ser avaliada pelas curvas mostradas na Figura 5.3B. As funções distribuição cumulativa geradas a partir das duas curvas, *PMF* da soma dos atrasos em um sentido e *PMF* do atraso de ida e volta, mostram-se praticamente sobrepostas nesta figura.

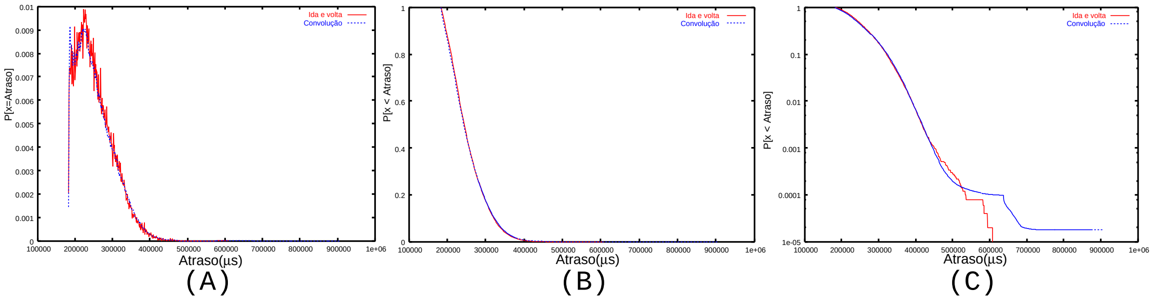

Figura 5.3: Comparação entre os cálculos dos atrasos: (A) PMF's: "Atraso de ida e volta" e Soma dos "Atrasos em um sentido"; (B) CDF's: "Atraso de ida e volta" e Soma dos "Atrasos em um sentido" (C) CDF's em escala LOG no eixo "Y": "Atraso de ida e volta" e Soma dos "Atrasos em um sentido".

Na Figura 5.3C são mostradas as mesmas curvas que as da Figura 5.3B, mas neste caso é usada a escala log no eixo das ordenadas. Nesta figura, é possível identificar uma diferença entre as curvas para atrasos superiores a 450 milisegundos. Porém, este erro é facilmente explicado. Isso se dá devido a alguns valores muito altos experimentados pelas sondas do "Atraso em um sentido". Esses valores, no cálculo da convolução das duas funções, são multiplicados por todos os outros valores da função relativa ao atraso no outro sentido. Durante o período de medição, suponha que uma sonda encontra um valor alto de atraso em um sentido, mas não nos dois sentidos quando o *RTT* é calculado. Neste caso a convolução fornecerá uma probabilidade não nula para um atraso grande nos dois sentidos. Esse valor, porém, pode não ter sido obtido nas medições do *RTT*.

5.2 Caracterização do Atraso em Um Sentido

Um dos objetivos definidos para este trabalho é a avalição das coletas obtidas através dos experimentos executados. Para tal, será feito um estudo dos resultados obtidos da métrica "Atraso em um sentido", identificando a distribuição que melhor caracteriza cada uma das coletas. Para se estimar a melhor distribuição que represente uma coleta, são usadas seis distribuições: Exponencial, Gaussiana, Lognormal, Pareto, Gamma e Weibull. Os parâmetros dessas distribuições são calculados através do método de momentos e, além da visualização gráfica, foi obtido o *Mean Square Error (MSE)* para auxiliar na estimativa da melhor distribuição.

5.2.1 Cálculo dos Parâmetros pelo Método de Momentos

O método de momentos consiste na obtenção de descritores de primeira e segunda ordem (por exemplo: média, variância, coeficiente de variação) para as amostras coletadas e cálculo dos parâmetros da distribuição através da solução de equações que relacionam os parâmetros com os descritores.

As seguintes equações foram usadas para obtenção dos descritores amostrais:

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, onde n é o número de amostras e X_i o valor relativo a cada uma delas;

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left|X_i - \overline{X}\right|$, onde n é o número de amostras, X_i o valor relativo a cada uma delas, e $\overline{X}$ a média calculada;

$CV = \frac{\sqrt{S^2}}{\overline{X}}$, onde S^2 é o coeficiente de variação e $\overline{X}$ a média calculada.

Estimar os parâmetros de algumas distribuições não é tão simples, pois nem sempre existe uma equação fechada para o cálculo dos parâmetros, usando o método de momentos. Em muitos casos, também, as funções distribuição cumulativa (CDF) não são fórmulas fechadas. Porém, o objetivo final não é a obtenção de valores de probabilidade especificamente, mas sim, de pontos da curva gerada pela função cumulativa. Para algumas distribuições foram usadas funções internas da ferramenta GNUPLOT [48], auxiliando na estimativa dos parâmetros e na geração dos pontos, além da visualização gráfica das curvas. A seguir será apresentado o cálculo para estimar os parâmetros baseados nos descritores, assim como a função de densidade, PDF e CDF das variáveis aleatórias usadas na caracterização dos resultados.

Exponencial

O parâmetro λ, da distribuição exponencial, tem relação inversa com a média. Assim, esse parâmetro pode ser estimado de forma bastante simples:

$\lambda = \frac{1}{\overline{X}}$

As funções são estimadas a partir de:

PDF: $f(x) = \lambda e^{-\lambda x}$

CDF: $F(x) = 1 - e^{-\lambda x}$

Gaussiana (normal)

Os parâmetros da distribuição Gaussiana são a média e o desvio padrão. A função distribuição da Normal não possui uma fórmula fechada, por isso a função "NORM" do GNUPLOT é usada para gerar a curva desta distribuição. A função $norm((x - \mu)/\sigma)$ no GNUPLOT retorna a curva da função distribuição do argumento. Os parâmetros são obtidos a partir de:

$\mu = \overline{X}; \sigma = \sqrt{S^2}$,

e as funções são estimadas da forma abaixo:

PDF: $f(x) = \frac{1}{\sigma\sqrt{2 \cdot \pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

CDF: $F(x) = norm[\frac{(x-\mu)}{\sigma}]$

Lognormal

Por definição [21], uma variável aleatória X é dita ter uma distribuição lognormal, se o logaritmo neperiano de X tem uma distribuição normal com média μ e desvio padrão σ. Para se estimar os parâmetros μ e σ da lognormal pelo método de momentos, são usados dois descritores (média e coeficiente de variação), pois as equações que relacionam estes descritores com os parâmetros são bastante simples. Após estimados os parâmetros, a função "NORM" do GNUPLOT é usada agora na forma $norm((\ln(x) - \mu)/\sigma)$.

A média e a variância da variável aleatória lognormal são definidas como:

$\mu' = e^{\mu + \frac{1}{2}\sigma^2} = e^{\mu} \cdot e^{\frac{1}{2}\sigma^2}$

$\sigma'^2 = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2} = e^{2\mu} \cdot e^{\sigma^2} \cdot \left[e^{\sigma^2} - 1\right]$

Definiremos $m = e^{\mu}$ e $w = e^{\sigma^2}$, logo podemos expressar a média e a variância da seguinte forma:

$\mu' = m \cdot \sqrt{w}$

$\sigma'^2 = m^2 \cdot w(w-1)$

O coeficiente de variação é dado por:

$\delta = \frac{\sqrt{\sigma'}}{\mu'} = \sqrt{w-1}$

Com a equação acima podemos obter o parâmetro w a partir do coeficiente de variação amostral, e σ^2 a partir de w:

$w = CV^2 + 1 \ \sigma^2 = \ln(cv^2 + 1)$.

O parâmetro m é obtido a partir da média amostral e do coeficiente de variação amostral:

$m = \frac{\overline{X}}{\sqrt{CV^2+1}}$

e μ é obtido a partir de m:

$\mu = \ln(m)$

As funções (PDF e CDF) são obtidas de:

PDF: $f(x) = \frac{e^{-\frac{1}{2}\left[\frac{\ln(x)-\mu}{\sigma}\right]^2}}{\sigma \cdot x \cdot \sqrt{2\pi}}$

CDF: $F(x) = norm[\frac{(\ln(x)-\mu)}{\sigma}]$

Pareto

Os parâmetros α e k, na distribuição Pareto caracterizam o formato e a inclinação da curva. Usaremos a abordagem de [39], que propõe usar o menor valor obtido, dentro de todas as amostras, como sendo igual ao parâmetro k da distribuição. A partir do valor de k, é fácil estimar o parâmetro α através da média amostral:

$\overline{X} = \frac{\alpha}{\alpha-1} * k$

As funções (PDF e CDF) são calculadas a partir de:

PDF: $f(x) = \frac{\alpha \cdot k^{\alpha}}{x^{-(\alpha-1)}}$

CDF: $F(x) = 1 - \left(\frac{k}{x}\right)^{\alpha}$

Gamma

Gamma é outra distribuição que não possui uma equação fechada para o cálculo da função cumulativa. Essa distribuição depende das funções GAMMA completa ($\Gamma(\alpha)$) e incompleta ($\Gamma(\alpha, \lambda x)$). O Gnuplot é também usado para auxiliar na obtenção desta função distribuição, pois as funções "Gamma" e "Igamma" estão disponíveis nesta ferramenta. Já os parâmetros, α e λ, podem ser obtidos a partir da média e da variância, usando o método de momentos. A média e variância da distribuição Gamma são definidas como:

$E[x] = \frac{\alpha}{\lambda}$

$var[x] = \frac{\alpha}{\lambda^2}$

Logo, podemos obter α e λ resolvendo as equações abaixo:

$\overline{X} = \frac{\alpha}{\lambda}$

$S^2 = \frac{\alpha}{\lambda^2}$

Para obter as funções (PDF e CDF) desta variável aleatória é necessário definir o conceito das funções Gamma completa e incompleta:

$\Gamma(\alpha) = \int x^{\alpha-1} e^{x} dx$

sendo: $u = \lambda x$

$\Gamma(\alpha, u) = \int u^{\alpha-1} e^{-u} du$

Assim, a *PDF* e a *CDF* são calculadas a partir de:

PDF: $f(x) = \frac{\lambda^{\alpha} \cdot x^{\alpha-1} \cdot e^{-\lambda x}}{\Gamma(\alpha)}$

CDF: $F(x) = \frac{\Gamma(\alpha, \lambda x)}{\Gamma(\alpha)}$

Weibull

Assim como para a distribuição Lognormal, serão usados os descritores média e coeficiente de variação para estimativa dos parâmetros da distribuição, pois as equações obtidas são simples.

O coeficiente de variação de uma variável aleatória Weibull é dado por:

$$\delta = \frac{\sqrt{\Gamma\left[1 + \frac{2}{\alpha}\right] - \Gamma^{2}\left[1 + \frac{1}{\alpha}\right]}}{\Gamma\left[1 + \frac{1}{\alpha}\right]} \tag{5.3}$$

Foi usada a função $\Gamma(x)$ do GNUPLOT para a obtenção do gráfico da função Weibull considerando diversos valores de α. Conhecendo-se o valor do coeficiente de variação amostral (CV), é então identificado no gráfico da função $\Gamma(x)$ qual o valor de α que corresponde ao CV.

A partir do valor escolhido para α, pode-se obter o valor de λ a partir da equação abaixo:

$\overline{X} = \left(\frac{1}{\lambda}\right)^{\frac{1}{\alpha}} \cdot \Gamma\left(1 + \frac{1}{\alpha}\right)$

A *PDF* e *CDF* da variável aleatória Weibull são dadas por:

PDF: $f(x) = \lambda \cdot \alpha \cdot x^{\alpha-1} \cdot e^{-\lambda x^{\alpha}}$

CDF:$F(x) = 1 - e^{-\lambda x^{\alpha}}$

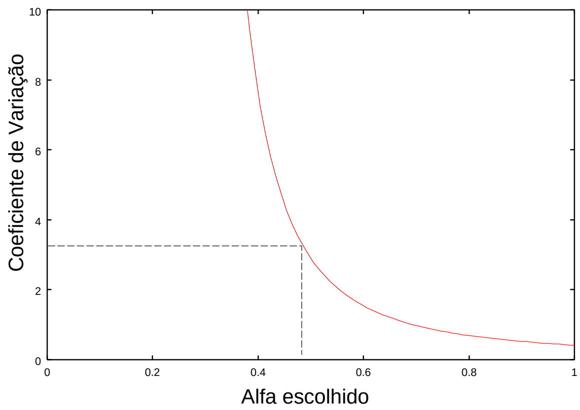

Figura 5.4: Processo de escolha do valor de α, a partir do valor do coeficiente de variação

5.2.2 Estimando a Distribuição que Melhor se Adequa aos Dados Coletados para o Atraso Fim-a-Fim

O método mais simples para escolha da distribuição é o visual. Primeiramente as distribuições são parametrizadas usando os descritores amostrais. Em seguida são traçadas as curvas das distribuições parametrizadas e a distribuição dos dados coletados (*trace*) em um mesmo gráfico.

A Figura 5.5A apresenta as distribuições parametrizadas através de dados coletados em um experimento. Na Figura 5.5B podem ser observadas as respectivas distribuições complementares.

A Figura 5.5C ilustra as distribuições complementares em escala logaritmica. A geração do gráfico neste formato permite que seja identificado um decaimento lento na distribuição complementar dos dados coletados. As distribuições que caracterizam de forma mais adequada o Atraso fim-a-fim deste experimento são Exponencial e Lognormal.

Além da visualização gráfica, o MSE (*Mean Square Error*) [33, 11] pode ser obtido para cada uma das distribuições. O *MSE* estima a média dos erros ao quadrado de uma função em relação à outra. Ele pode ser obtido a partir da equação:

$MSE = \frac{\sum_0^M \left|F(x) - \hat{F}(x)\right|^2}{M}$, onde M é o número de amostras coletadas no *trace*.

Para calcular o *MSE* das distribuições, amostras das curvas estimadas são gera-

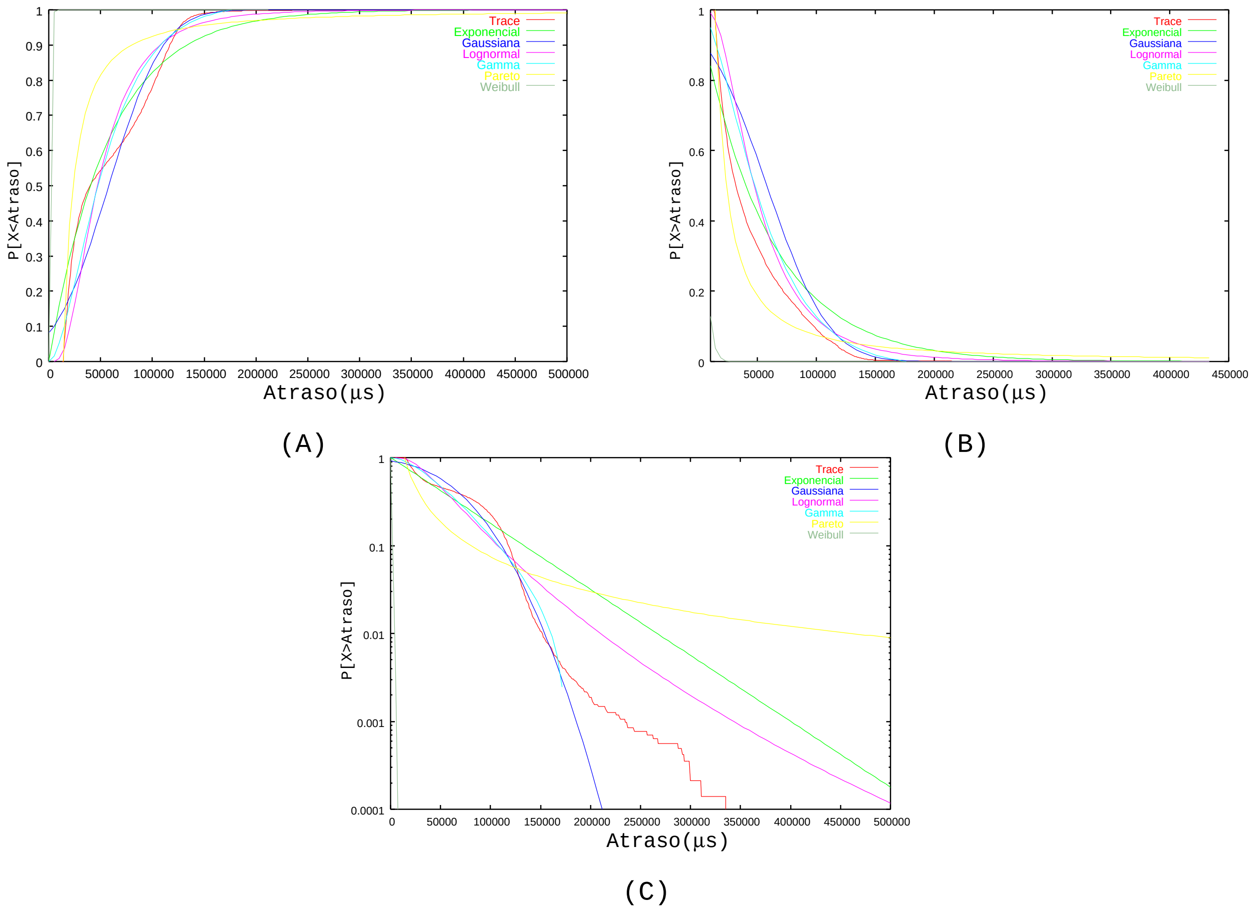

Figura 5.5: Estimativa das distribuições: (A) *CDF*; (B) Distribuição Complementar; (C) Distribuição Complementar em escala logaritmica.

das pelo GNUPLOT. A partir das amostras, o *MSE* de cada distribuição é calculado. O problema encontrado nesta implementação é a aproximação gerada pelo GNUPLOT, que é de aproximadamente 10^{-5}. Essa aproximação torna as estimativas feitas do *MSE*, para cada distribuição, pouco confiáveis.

A Tabela 5.1 apresenta os valores do *MSE* para cada uma das distribuições mostradas na Figura 5.5. Os resultados indicam que a melhor distribuição estimada para caracterizar o atraso em um sentido obtido no experimento é a exponencial.

5.3 Avaliação dos Resultados dos Experimentos

Durante seis meses foram executados experimentos entre a UFRJ e a UMASS. Estes testes foram realizados ao longo do segundo semestre de 2002, em paralelo

Distribuição	MSE
Exponencial	0.0000896541*
Gaussiana	0.0001109796
Lognormal	0.0001263688
Pareto	0.0004866717
Gamma	0.7155580298
Weibull	0.9820174230

Tabela 5.1: Resultados do *MSE* para as distribuições indicadas.

às sessões de videoconferência, onde aulas e seminários eram transmitidos entre os dois pontos. O resultado obtido desses experimentos é uma coleção significativa de *Traces* com amostras do "Atraso em um sentido", dos pacotes durante as sessões de videoconferência. Mais de 40 coletas obtidas dos experimentos (UFRJ/UMASS) foram analisadas. O objetivo da análise é estimar, para cada coleta, a distribuição que melhor caracteriza o atraso sofrido pelos pacotes.

Em todas as coletas foram aplicados os algoritmos de remoção do *Skew*, *Offset*, ruídos e atualizações de relógios. Foram ainda calculadas as funções de probabilidade de massa, as distribuições e as distribuições complementares das amostras coletadas. As curvas de diversas distribuições conhecidas, parametrizadas com os dados coletados, foram estimadas com a finalidade de se obter uma distribuição que caracterizasse o comportamento do atraso no caminho entre a UFRJ e *UMASS*. A Figura 5.6 mostra o resultado de três desses experimentos. Os gráficos da esquerda apresentam as distribuições e os da direita as distribuições em escala logaritmica.

Apesar das distribuições Gaussiana e Lognormal terem tido menor *MSE* em algumas das coletas, a distribuição Pareto foi a indicada na grande maioria dos casos como mostra a Tabela 5.2. Tanto no sentido UFRJ-UMASS, quanto no sentido oposto, grande parte dos atrasos são indicados pelo *MSE* como melhor caracterizado pela distribuição Pareto. O problema dessa estimativa é que, devido à aproximação gerada pelo GNUPLOT, pontos importantes estão sendo desconsiderados, como por exemplo o comportamento da cauda das distribuições estimadas.

Algumas das coletas geradas pela ferramenta de transmissão de vídeo também

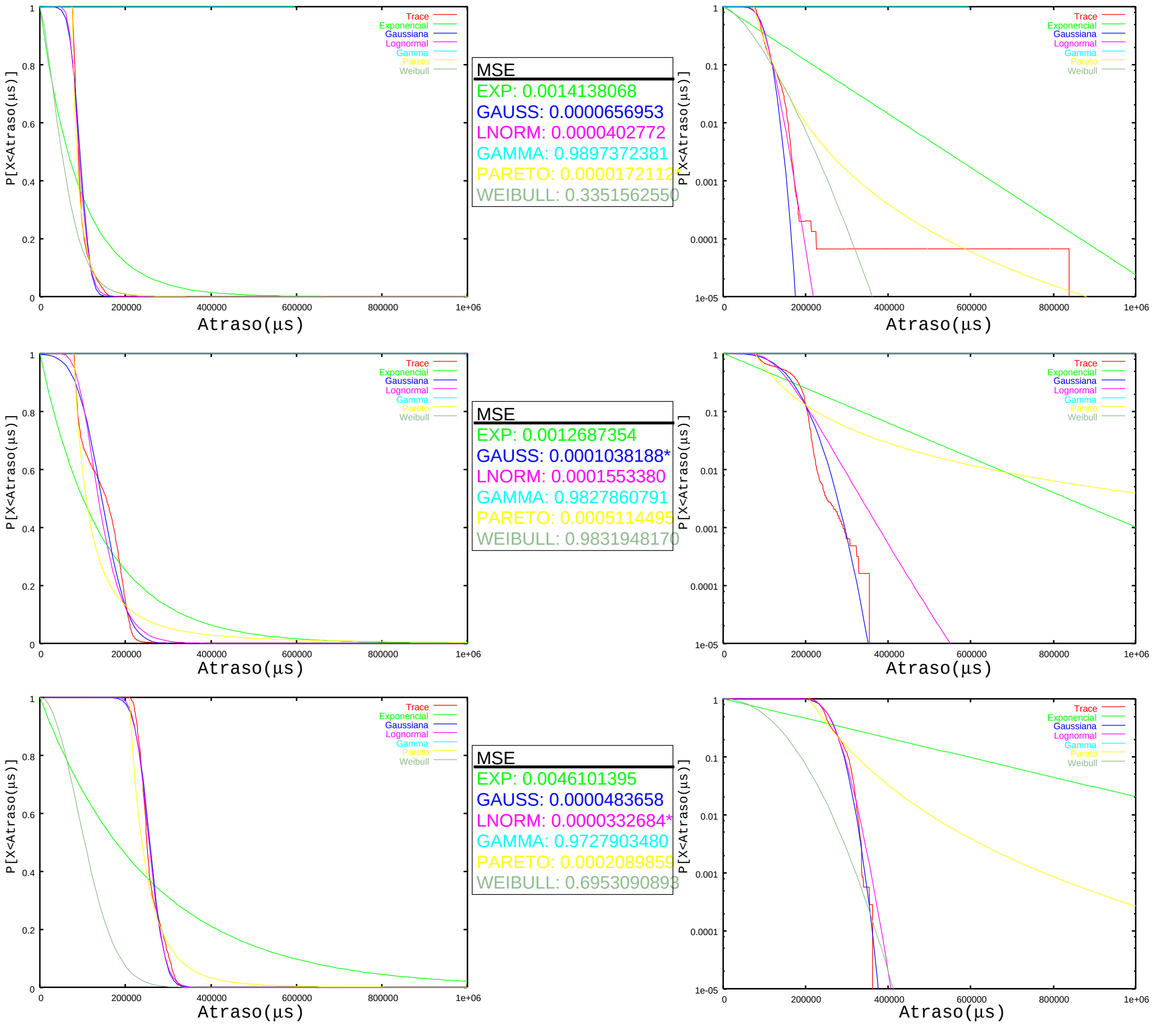

Figura 5.6: Resultados de experimentos entre a UFRJ e a UMASS.

foram submetidas à análise da distribuição do atraso. A Figura 5.7 ilustra o resultado obtido e as estimativas feitas para o *MSE*. O número de coletas com esse tipo de experimento não foi suficientemente grande para que conclusões possam ser tiradas. A Figura 5.7 mostra o resultado e o experimento que apresentou menor valor para o *MSE*. Neste caso a distribuição mais apropriada é a Lognormal.

Além dos experimentos feitos com a UMASS, um número reduzido de testes foram realizados entre a UFRJ e a UNIFACS (Universidade Salvador). Cerca de quatorze coletas foram feitas e, desses resultados, quatro diferentes distribuições foram indicadas pelo *MSE* para representar o atraso (Gaussiana, Lognormal, Exponencial e Pareto). Algumas dessas coletas estão ilustradas na Figura 5.8. Para a maioria das coletas, entre a UFRJ e a Unifacs, a distribuição Gaussiana mostrou-se

Distribuição	Freqüência
Pareto	80%
Gaussiana	10%
Lognormal	10%

Tabela 5.2: Freqüência obtida das distribuições nos experimentos entre a UFRJ e UMASS.

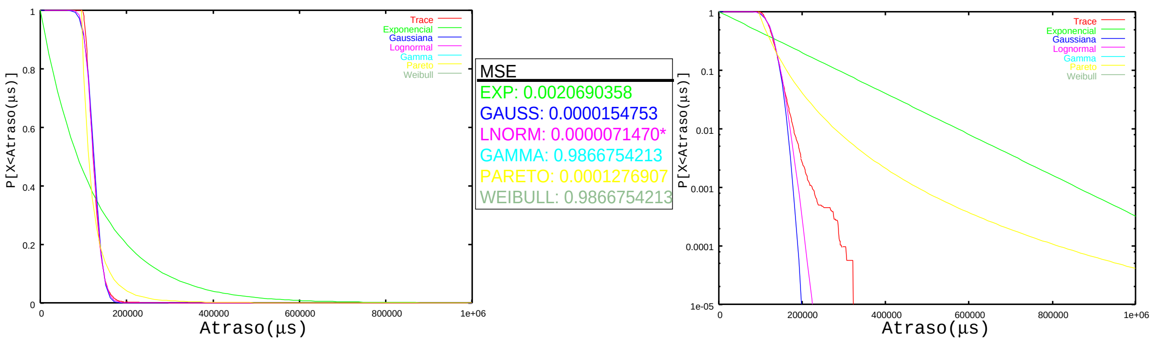

Figura 5.7: Resultado de uma estimativa de atraso para o tráfego de vídeo entre a UFRJ e a UMASS.

a mais adequada para caracterizar o atraso. Na Tabela 5.3 é apresentado um resumo do percentual de ocorrências dos resultados obtidos. Mas, novamente, é importante destacar a baixa precisão destas estimativas por conseqüência da aproximação imposta pelo GNUPLOT.

5.4 Uso dos Resultados na Criação de Modelos

Como foi visto na Seção 3.2.1, o *TANGRAM-II Modelenv* é um ambiente de modelagem que permite ao usuário a criação de modelos formado por "objetos" que

Distribuição	Freqüência
Exponencial	16%
Gaussiana	59%
Lognormal	9%
Pareto	16%

Tabela 5.3: Freqüência obtida das distribuições nos experimentos entre a UFRJ e UNIFACS.

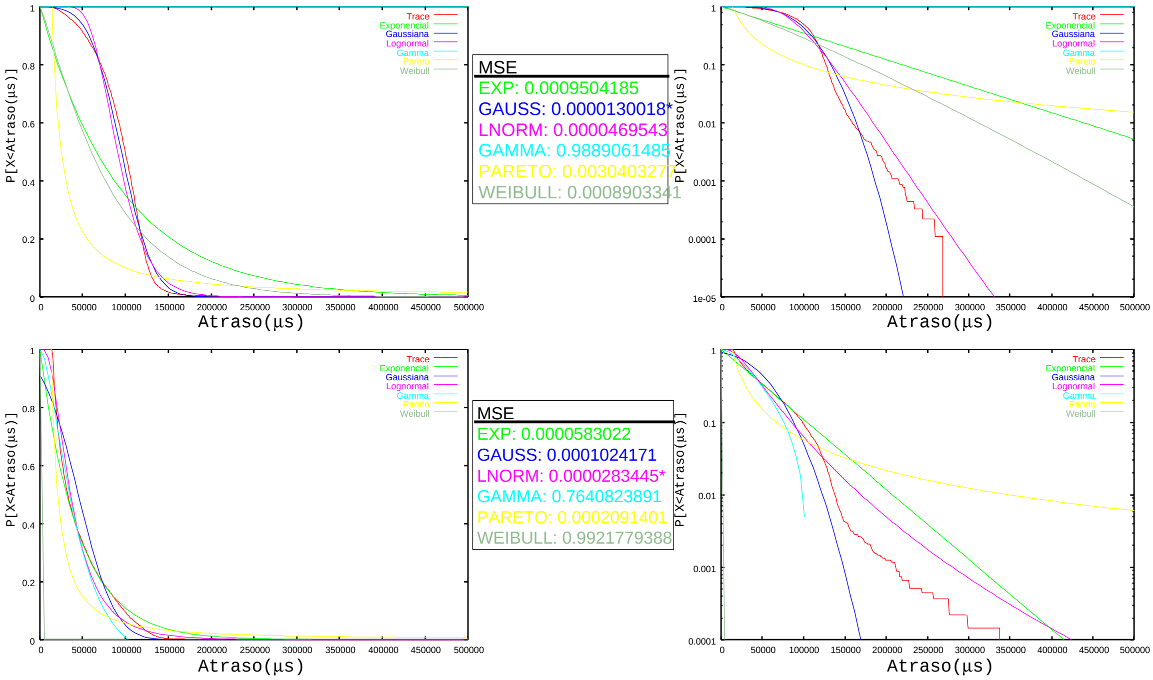

Figura 5.8: Resultados de experimentos entre a UFRJ e UNIFACS.

interagem entre si. A Figura 5.9 representa um modelo de fila M/M/1/K. Maiores detalhes e exemplos mais complexos podem ser encontrados em [27].

Um dos componentes/objetos de um modelo pode representar o processo de retardo e/ou de perda em uma rede. Foi desenvolvido para o *TANGRAM-II Traffic Generator* um gerador de objetos cuja finalidade é utilizar resultados obtidos durante experimentos para auxiliar na criação de modelos de retardo e perda mais reais. Os objetos, denominado de "Canal", podem ser gerados a partir de parâmetros obtidos em coletas com o *TANGRAM-II Traffic Generator*. Após estimados o "Atraso em um sentido" e a distribuição que melhor caracteriza a coleta, um objeto pode ser gerado onde o atraso do "Canal" seguirá a mesma distribuição estimada pelo experimento. Além do atraso, a perda também pode ser semelhante àquela obtida no experimento. Um exemplo deste objeto gerado é ilustrado na Figura 5.10.

A Figura 5.11 ilustra um modelo criado no *TANGRAM-II Modelenv* onde a rede é representada por um objeto canal. Esse objeto foi gerado a partir de uma coleta de pacotes de uma transmissão de vídeo. Usando o "Canal" é possível desenvolver um modelo de aplicações, cliente e servidora, para transmissão de vídeo em tempo real. Uma série de simulações podem ser executadas, fazendo-se alterações nos parâmetros

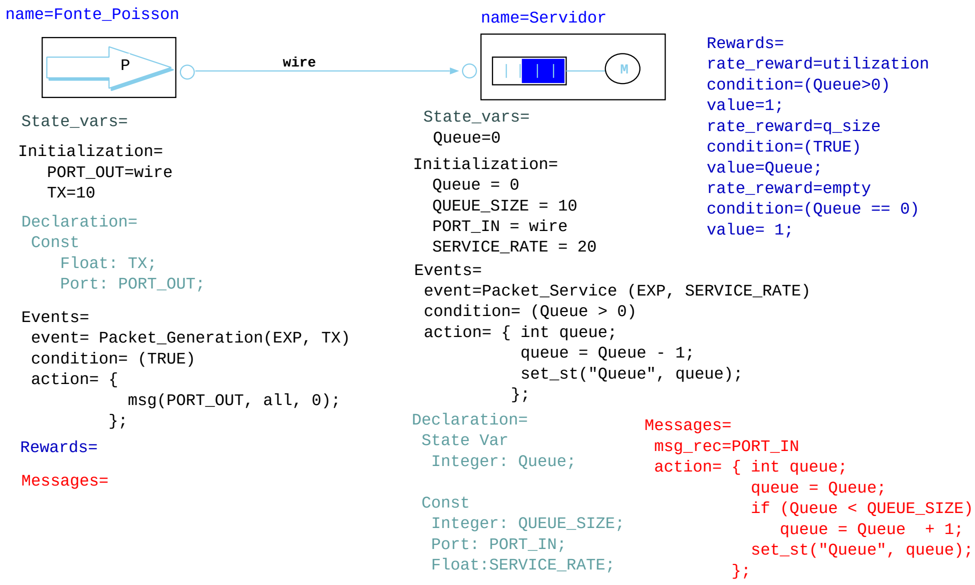

Figura 5.9: Exemplo de modelo do ambiente TANGRAM-II Modelenv.

dos objetos que representam as aplicações, até se chegar a um valor considerado ideal para determinadas características dessas aplicações, como por exemplo o tamanho do *buffer* na aplicação do cliente.

```
name=Channel
```

```
State_vars=
N_Pcts=0
Channel[0:10] =
Aux_Channel[0:10] =
```

```
Declaration=
Const
  Integer: CHANNEL_SIZE;
  Float  : P_LOSING_PKT, MEAN_DELAYER, VAR_DELAYER;
  Port   : SEND_PKT_PORT, RECV_PKT_PORT;

Var
  Integer: N_Pcts, Channel[CHANNEL_SIZE];
  Float  : Aux_Channel[CHANNEL_SIZE];

Rewards=

rate_reward = num_pkts_rew
condition = (FALSE)
value = 0;

rate_reward = send_time_rew
condition = (FALSE)
value = 0;
```

```
Initialization=
/* Inicializacao das Constantes */
P_LOSING_PKT = 0.02   /* probabilidade de perda de pacotes */
RECV_PKT_PORT = sender_channel_1
SEND_PKT_PORT = channel_1_receiver
CHANNEL_SIZE = 120    /* tamanho do canal */
MEAN_DELAYER = 10   /* media utilizada na distribuicao normal do atraso dos pacotes */
VAR_DELAYER = 54.975 /*variancia utilizada na distribuicao normal do atraso dos pacotes */
/* --- */

/* Inicializacao das Variaveis de Estado */
N_Pcts = 0           /* numero de pacotes no canal */
Channel[] = 0        /* vetor para armazenar os numeros de sequencia dos pacotes*/
Aux_Channel[] = 0.0    /* vetor para armazenar os tempos de envio dos pacotes */
/* --- */
```

```
Messages=
/* Ao receber uma mensagem do sender, este objeto simula, com probabilidade
   P_LOSING_PKT, a perda do pacote recebido. Se nao houver perda, este objeto
   atualiza o valor das variaveis de estado N_Pcts, que representa o numero de
   pacotes no canal, e Channel, que e um vetor utilizado para armazenar tais pacotes.
     */
msg_rec = RECV_PKT_PORT
action={
         int n_pcts;
         int channel[CHANNEL_SIZE], i;
         float aux_channel[CHANNEL_SIZE], message[2], temp;

         /* Inicializacao das Variaveis */
         get_msg_data(message);
         get_st(n_pcts,"N_Pcts");
         get_st(channel,"Channel[]");
         get_st(aux_channel,"Aux_Channel[]");
         /* --- */

         if (n_pcts < CHANNEL_SIZE)
         {
           channel[n_pcts] = (int) message[0];
           aux_channel[n_pcts] = message[1];
           n_pcts++;
         }
         clone_ev("Delayer");
         set_cr(num_pkts_rew,n_pcts);
         temp = message[1];
         set_cr(send_time_rew,temp);
         /* Atualizacao das Variaveis de Estado */
         set_st("N_Pcts",n_pcts);
         set_st("Channel[]",channel);
         set_st("Aux_Channel[]",aux_channel);
         /* --- */
       }:prob = 1 - P_LOSING_PKT;
       {
         ;
       }:prob = P_LOSING_PKT;
```

```
Events=
/* Este evento e o responsavel pelo atraso dos pacotes que chegam ao canal,
   simulando uma rede real */
event     = Delayer (GAUSS, MEAN_DELAYER, VAR_DELAYER)
condition = (N_Pcts > 0)
action = {
           /* Envia um pacote para o receptor e atualiza as variaveis
              de estado N_Pcts e Channel */
           int n_pcts, i, channel[CHANNEL_SIZE];
           float aux_channel[CHANNEL_SIZE], message[2];
           /* Inicializacao das Variaveis */
           get_st(n_pcts,"N_Pcts");
           /* --- */
           message[0] = Channel[0]; /* contem o SN */
           message[1] = Aux_Channel[0]; /* contem o tempo no qual o pacote foi enviado pelo emissor */
           msg(SEND_PKT_PORT,all,message);
           for (i = 1;i<=(CHANNEL_SIZE-1);i++)
           {
             channel[i-1] = Channel[i];
             aux_channel[i-1] = Aux_Channel[i];
           }
           channel[CHANNEL_SIZE-1]=0;
           aux_channel[CHANNEL_SIZE-1]=0;
           n_pcts--;
           set_cr(num_pkts_rew,n_pcts);
           /* Atualizando as variaveis de estado */
           set_st("N_Pcts",n_pcts);
           set_st("Channel[]",channel);
           set_st("Aux_Channel[]",aux_channel);
           /* --- */
         };
```

Figura 5.10: Objeto canal.

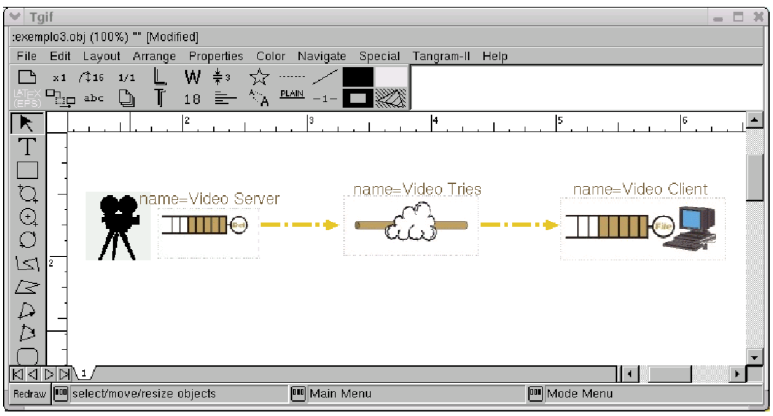

Figura 5.11: Exemplo de um modelo criado no ambiente *TANGRAM-II Modelenv* onde o objeto "Canal" foi gerado com parâmetros estimados de uma coleta.

Capítulo 6

Nova Proposta para Estimar Métricas Baseadas em Pares de Pacotes

A métrica "capacidade de transmissão do enlace no gargalo", definida anteriormente na Seção 3.1.4, pode ser estimada através de técnicas de medição ativa. O método proposto para a obtenção desta métrica, os problemas encontrados neste método, e uma nova proposta que tenta solucionar alguns desses problemas, serão apresentados a seguir. Ainda neste capítulo, serão descritos o método *Loss Pair* e as medidas por ele obtidas, além das implementações, desenvolvidas na ferramenta *TANGRAM-II Traffic Generator*, para estimar a capacidade do gargalo e as métricas do *Loss Pair* (por exemplo, Tamanho do *Buffer*) [30].

6.1 Método "Pares de Pacotes"

O método "Pares de Pacotes" consiste na emissão de dois pacotes, separados por um intervalo de tempo bem próximo de zero, que atravessam o mesmo caminho na rede até chegarem a um único destino. A coleta desses pacotes vem sendo usada para identificar diversos parâmetros da rede. Keshav, em [23], foi o primeiro a usar o método para medir a capacidade de transmissão do enlace no gargalo. Depois disso, diversas propostas e ferramentas utilizaram o método ou uma variação dele, para estimar a mesma métrica [9, 24, 22, 6, 25] ou outras medidas baseadas no envio de pares de pacotes [30].

6.1.1 Usando o Método “Pares de Pacotes” para Estimar Capacidade de Transmissão do Enlace no Gargalo

A proposta da técnica “Pares de Pacotes” para estimar a capacidade de transmissão do enlace no gargalo é baseada no intervalo entre as chegadas dos pacotes. Esse comportamento foi originalmente ilustrado em [20]. A suposição principal da técnica é que a dispersão entre os pacotes, encontrada nos pares coletados, é causada pela menor capacidade de transmissão ao longo do caminho. Os pacotes que são gerados de uma mesma origem e separados por intervalos de tempo bem próximos de zero, possuem o espaçamento entre eles mantido até que passem por um enlace com capacidade de transmissão inferior à do emissor. Essa separação, causada pelo tempo de transmissão deste enlace que é superior aos experimentados nos enlaces anteriores, é mantida até o destino das mensagens. A menos que seja encontrado ao longo do resto do caminho um outro enlace com uma capacidade ainda menor. A Figura 6.1 ilustra o efeito de dispersão entre os pacotes causada pela diferença entre a capacidade de transmissão dos enlaces ao longo do caminho.

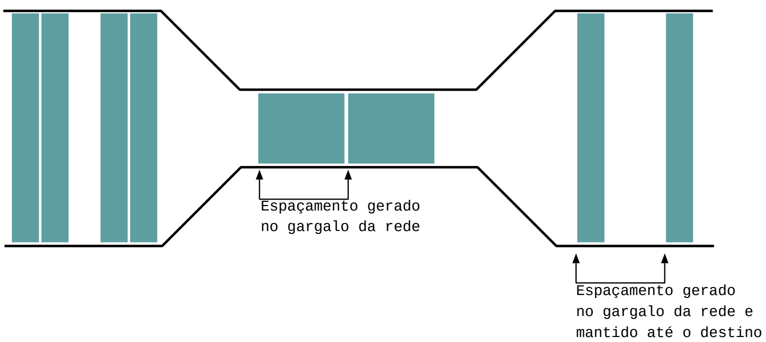

Figura 6.1: Ilustração do funcionamento do método Pares de Pacotes com a dispersão imposta pelo enlace de menor capacidade.

Com a variação do intervalo de tempo entre as chegadas e o tamanho dos pacotes é possível estimar a capacidade de transmissão do gargalo. A divisão do tamanho do pacote pelo intervalo de tempo entre as chegadas equivale à métrica desejada. Supondo a inexistência de tráfego concorrente, apenas um par de sondas já é o suficiente para a obtenção desta métrica. Porém, nem sempre esta suposição é verdadeira. Para melhorar a previsão da estimativa, uma série de pares devem ser usados. Um histograma com os resultados obtidos é gerado, e o intervalo de tempo com maior probabilidade no histograma é usado para calcular a métrica. O tamanho dos pacotes em cada par, deve ser idêntico, garantindo que a dispersão entre os pacotes não aumente nem diminua por conseqüência da diferença dos tamanhos.

O algoritmo para estimar a capacidade de transmissão do enlace no gargalo, através de "Pares de Pacotes" foi implementado no *TANGRAM-II Traffic Generator*. Uma série de outras ferramentas possuem esta técnica implementada e [8, 25] são apenas duas delas. Porém, a implementação do método na ferramenta foi a opção encontrada para possibilitar uma avaliação mais detalhada da técnica com todas as variações desejadas. O objetivo é possibilitar que comparações sejam feitas entre esta técnica e uma nova proposta, que será apresentada na Seção 6.2 deste trabalho.

Uma série de experimentos foram realizados usando a ferramenta. A Figura 6.2 ilustra o resultado obtido em uma rede local. A rede usada para um desses testes é a do laboratório LAND da UFRJ. Todas as interfaces disponíveis nesta rede têm capacidade de transmissão igual a 100 Mb/s, equivalente ao resultado estimado no teste mostrado na Figura 6.2. Nesta figura é possível identificar que o pico do gráfico (valor mais freqüente no histograma) indica a capacidade de transmissão estimada para aquele experimento. Percebe-se ainda, que outros valores, além dos estimados para 100 Mb/s, são encontrados no gráfico. Isso ocorre devido à interferência do tráfego concorrente na dispersão dos pacotes.

Durante os experimentos executados entre a UFRJ e a UMASS para estimar o "Atraso em um sentido", descrito na Seção 5.4, foram realizados também experimentos usando a técnica de pares de pacotes. Alguns resultados das estimativas feitas da capacidade de transmissão do gargalo serão apresentados ainda nesta seção.

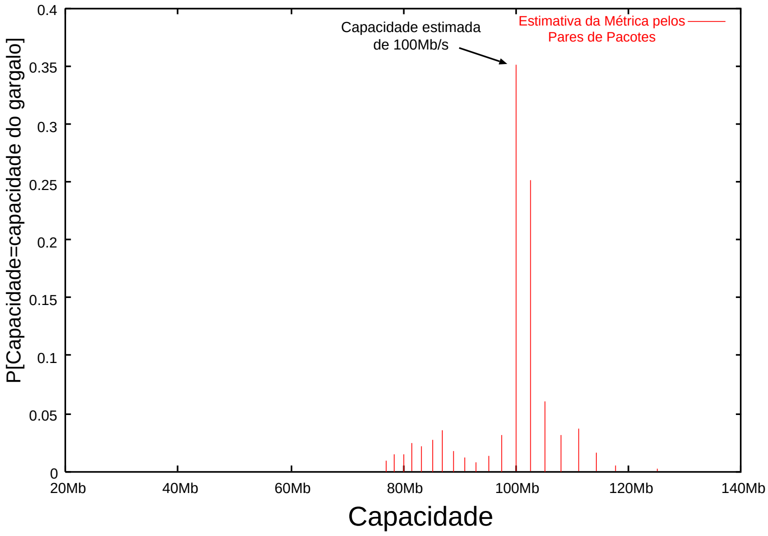

Figura 6.2: Estimativa da capacidade de transmissão do gargalo em uma rede local.

6.1.2 Problemas com o Método de Keshav

Diversos trabalhos se propuseram a avaliar a qualidade e a precisão do método de pares de pacotes [24, 6, 22]. Alguns desses trabalhos relatam a imprecisão da técnica para pacotes muito pequenos. Em [6] é indicado um tamanho específico para os pacotes de aproximadamente 800 bytes. Outros trabalhos propõem variações desse método, como a utilização de mais de dois pacotes (Trens de Pacotes).

Dentre as avaliações feitas relativas à técnica "Pares de Pacotes", pelo menos um ponto é acordado por todos os estudos: o tráfego concorrente atrapalha consideravelmente a estimativa da métrica. A Figura 6.3 ilustra uma estimativa feita da capacidade de transmissão do enlace no gargalo usando o método de pares de pacotes que supomos ser errada. Apesar de não passar de uma suposição, por não conhecermos a real capacidade de todo o caminho percorrido pelas sondas, uma série de experimentos estimam um outro valor para a "menor capacidade de transmissão ao longo do caminho" de medições aplicadas entre esses mesmos pontos (esses resultados serão apresentados na Seção 6.2). Esse problema ocorre devido ao alto tráfego concorrente presente no gargalo da rede durante os experimentos.

O problema do tráfego concorrente para o cálculo da métrica é que a "capacidade de transmissão do enlace no gargalo" é inversamente proporcional ao intervalo entre

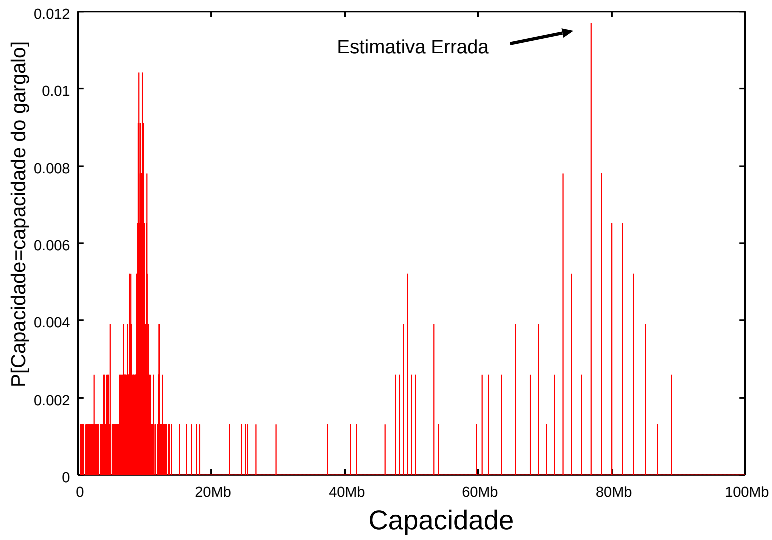

Figura 6.3: Estimativa errada da capacidade de transmissão do gargalo no caminho entre a UFRJ e a UMASS, causada por um alto tráfego concorrente.

as chegadas dos pacotes. A influência causada por este tráfego pode ser caracterizada de duas formas: primeiro, a presença de pacotes em frente aos pares na fila dos roteadores, após já terem passado pelo nó gargalo da rede, pode ocasionar uma redução na dispersão existente entre as duas sondas. A conseqüência disso é a super-estimativa da capacidade de transmissão do gargalo; segundo, a "injeção" de tráfego entre os pacotes pode resultar em um acréscimo da dispersão das sondas e causar uma estimativa inferior à capacidade real de transmissão do gargalo.

6.2 Uma Nova Proposta para Estimar a Capacidade do Gargalo

Devido a todas as dificuldades impostas para estimar a "capacidade de transmissão do enlace no gargalo" apresentadas anteriormente, usando a técnica "Pares de Pacotes", o desenvolvimento de um novo método que possibilite a obtenção dessa métrica, mesmo em situação de um alto tráfego concorrente, é mais uma motivação deste trabalho. A proposta desenvolvida baseia-se na dispersão entre as chegadas dos pares de pacotes para estimar a capacidade do gargalo. No entanto, apenas pares selecionados são usados para fazer o cálculo da métrica. Como o tráfego concorrente atrapalha de forma considerável as estimativas dos pares de pacotes, o novo método se propõe a selecionar apenas alguns pares para uso na estimativa da métrica.

O método proposto neste trabalho prevê a escolha das sondas que serão, ou não, usadas durante o cálculo da medida. O parâmetro base desta seleção é o "atraso em um sentido" sofrido pelo primeiro pacote do par. O objetivo com esta seleção é usar apenas pares que, supostamente, sofreram pouca ou nenhuma influência do tráfego concorrente durante a travessia do caminho na rede.

A seleção dos pares é feita a partir do atraso sofrido pelo primeiro deles. A suposição dessa seleção é de que os pares que não tiveram um retardo grande, têm uma menor probabilidade de terem sofrido influências consideráveis do tráfego concorrente. Para isso, é calculado o menor "atraso em um sentido" entre todas as sondas usadas na medição. Isso possibilita supor que pares, cujo primeiro pacote obteve um "atraso em um sentido" próximo a esse valor, possuem uma probabilidade menor de terem encontrado situações de alto tráfego concorrente durante suas transmissões. Assim, os pares de pacotes, usados para o cálculo da métrica, devem ter o atraso de suas primeiras sondas "próximo" ao menor retardo encontrado em todo o experimento.

O algoritmo implementado, para fazer a seleção dos pares a serem usados, considera um valor de retardo "próximo", quando este tem no máximo dez por cento a mais do que o menor atraso encontrado entre todas as amostras. Apesar de qualquer outro fator de variação para os atrasos poder ser usado na seleção dos pares, esse valor se mostrou satisfatório com os resultados obtidos a partir dos experimentos.

Durante os experimentos executados entre a UFRJ e a UMASS, algumas estimativas erradas foram obtidas do método tradicional de pares de pacotes, como foi ilustrado na Figura 6.3. As mesmas coletas foram aplicadas ao novo algoritmo proposto. O resultado obtido por este algoritmo e a comparação com o método tradicional que não faz seleção dos pares, estão ilustrados na Figura 6.4.

É importante perceber que a estimativa feita pelo novo algoritmo praticamente reduz o erro da medida devido ao tráfego concorrente. Obviamente, no pior caso,

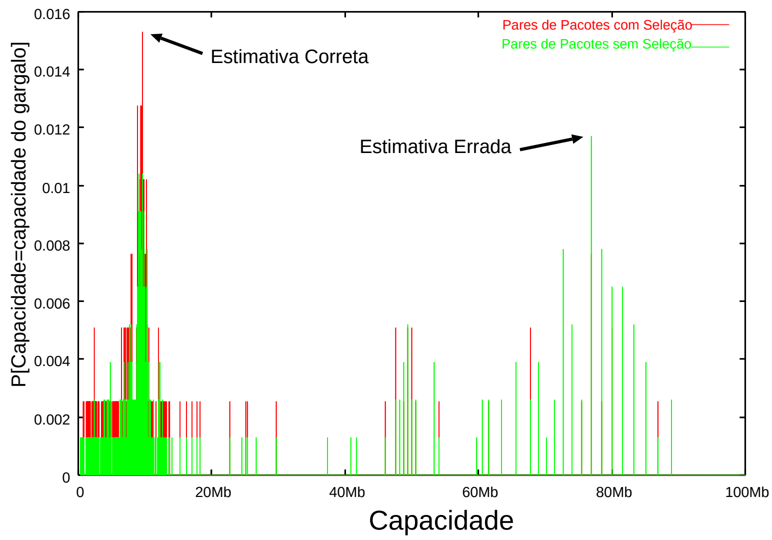

Figura 6.4: Estimativa correta da capacidade de transmissão do gargalo no caminho entre a UFRJ e a UMASS, mesmo com um alto tráfego concorrente.

a estimativa desta proposta é igual ao uso do método tradicional para a estimativa da "capacidade de transmissão no enlace do gargalo".

O algoritmo, descrito acima, foi implementado na ferramenta *TANGRAM-II Traffic Generator*. Como já foi comentado anteriormente, uma série de experimentos foram aplicados entre a UFRJ e a UMASS. Alguns dos resultados obtidos, a partir dessa coleta, estão ilustrados na Figura 6.5.

Durante o período de experimentos nem sempre foi utilizada a mesma topologia de rede. Além da possibilidade de rotas diferentes entre os dois pontos de medição, o roteador de saída de uma das redes foi alterado ao longo dos meses em que as coletas foram feitas. Tanto a possibilidade de diferentes rotas, quanto as alterações na estrutura da rede, explicam a diferença nas estimativas entre os diferentes dias. Essa alteração está ilustrada nas Figuras 6.5 inferiores, no sentido UFRJ-UMASS.

6.3 Trabalho Futuro com Método *Loss Pair*

Como foi mencionado acima, existem várias propostas baseadas na técnica "Pares de Pacotes" para estimar a capacidade do gargalo. Porém, algumas propostas vêm sendo definidas, a partir de variações do "Pares de Pacotes", para o cálculo de outras métricas. Em [30] foi definido o método *Loss Pair*, onde o autor propõe uma variação do método de "Pares de Pacotes" para a obtenção de medidas relacionadas ao gargalo da rede.

A proposta do método *Loss Pair*, de Liu e Crovela [30], é uma tentativa de estimar outros parâmetros, além da capacidade de transmissão do gargalo em um caminho de rede. O envio dos pacotes neste método funciona de forma semelhante ao do "Pares de Pacotes". Dois pacotes são gerados, de uma mesma origem, em direção a um mesmo destino. Essas sondas atravessam todo o caminho até chegaram ao gargalo da rede. Como o nó gargalo possivelmente tem uma alta taxa de utilização e, em conseqüência, o *buffer* encontra-se praticamente cheio, um dos dois pacotes é descartado por falta de espaço no *buffer*. O outro pacote, que entrou na fila bem próximo de uma situação de perda, é atendido e segue o restante do caminho até alcançar o destino.

Os pacotes que chegam ao destino, sem o duplo sucesso dos seus pares, possibilitam extrações de informações importantes a respeito do gargalo. Os pacotes que não foram descartados enfrentaram situações bem próximas ao período de perda desta fila. Assim, o autor propõe algoritmos para estimar métricas como o tamanho do *Buffer* na "máquina gargalo" e alguns parâmetros de possíveis políticas de descarte implementadas por esta máquina. Detalhes desses métodos são descritos no final desta seção.

Para fazer uma estimativa correta da métrica, algumas características são assumidas como válidas pelo método durante os experimentos. Caso alguma das suposições não seja verdadeira, a estimativa feita pode conter erros grosseiros. Além das suposições do método de "Pares de Pacotes", onde os pacotes são injetados a uma distância bem próxima de zero e que os pacotes supostamente seguem a mesma rota ao longo do caminho, outras particularidades são necessárias para a técnica de *Loss Pair*: o nó do caminho, com menor capacidade de transmissão deve ser também aquele onde ocorrem a maioria das perdas. Além disso, os gargalos, de perda e transmissão, não podem mudar durante toda a medição.

Para estimar as métricas, o autor aplica experimentos de "ida e volta". A partir das coletas, são identificados os pares que perderam um dos seus pacotes. Um histograma do "retardo de ida e volta", experimentado pelas sondas sem pares, é gerado. O valor mais comum do histograma (pico da curva) é considerado pelo autor a situação onde os pacotes só enfrentaram a fila do gargalo. O menor valor de retardo, entre todas as sondas do experimento, é considerado o tempo de propagação somado ao tempo de transmissão, ou seja, esse pacote não enfrentou fila durante todo o caminho de "ida e volta" na rede. A subtração do "retardo mais comum" pelo "menor atraso" resulta em uma variável que o autor chama de "tempo total para escoamento da fila cheia" ("*Full Queue Drain Time*"). A partir da estimativa dessa variável, são propostos métodos para estimar métricas.

A primeira métrica a ser calculada é a capacidade de transmissão no gargalo. Os pares, cujos dois pacotes obtiveram sucesso na transmissão, podem ser usados para fazer a estimativa, mas ferramentas específicas para esta finalidade podem também ser aplicadas. A obtenção desse parâmetro é fundamental para as estimativas de outras métricas usando o *Loss Pair*.

Estimando "Tamanho do *Buffer*"

Com as métricas "tempo total para escoamento da fila cheia" e "capacidade de transmissão", calcular o tamanho do *buffer* é simples. A divisão do tempo total que a máquina leva para esvaziar sua fila, pela sua capacidade de transmissão, equivale ao tamanho total da fila. Em [30] o autor apresenta, ainda, uma série de validações e avalia a precisão das estimativas de tamanho do *Buffer* feitas usando a ferramenta *Network Simulator* [19].

Estimando "Taxa de Descarte" e Parâmetros de Políticas de Descarte

A estimativa dos parâmetros de políticas de descarte proposto por Liu e Crovella, além de todas as suposições já definidas pelo método *Loss Pair*, dependem da existência de mecanismos para esta finalidade implementados no gargalo da rede. A proposta tenta identificar parâmetros dos mecanismos de descarte de fila, RED[16, 13] e BLUE [46, 47], em funcionamento no gargalo da rede. Para isso, uma variável, definida como "Taxa de Descarte", precisa ser estimada. Essa variável, calculada através da equação 6.1, representa a taxa de *Loss Pair* com atraso igual a "r", em relação ao número de pares que não perdeu nenhum dos dois pacotes também com atraso igual a "r".

$$Dx = \frac{[nmero\, de\ Loss\ Pair\ , RTT =' r']}{[nmero\, de\ Pares\ Ok\ , RTT =' r']} \tag{6.1}$$

Cada política de fila define os parâmetros para o descarte dos pacotes. O objetivo, usando o *Loss Pair*, é estimar os valores definidos para os parâmetros como "*Min Threshold*", "*Max Threshold*", e "*Taxa de Descarte*". Para isso, gráficos são gerados das "Taxas de Descarte" obtidas para as determinadas ocupações de *Buffer*. Alguns exemplos, apresentados em [30], demonstram que as estimativas feitas desses parâmetros podem ser consideradas bem próximas das definidas no modelo usado no trabalho.

Toda a descrição feita da técnica *Loss Pair* é importante para que fiquem claras as vantagens das implementações feitas na ferramenta. Todo o método descrito, para as estimativas com a técnica *Loss Pair*, foi implementado no *TANGRAM-II Traffic Generator*. No entanto, diferente do método usado por Liu e Crovella, nesta implementação, as medidas podem ser obtidas em um único sentido, e usando uma técnica que acreditamos ser mais precisa para estimativa da "Capacidade de transmissão do gargalo". Dessa forma, acredita-se que os resultados obtidos, com esse novo modelo de aplicação da técnica *Loss Pair*, podem oferecer uma maior precisão e as suposições do método satisfeitas mais facilmente.

Algumas vantagens justificam o uso do modelo proposto neste trabalho para a aplicação da técnica *Loss Pair*. As estimativas feitas em um sentido reduzem a probabilidade de ter mais do que um "gargalo de perda" ou que esse ponto não seja o mesmo do "gargalo de transmissão". Desta forma, as chances do experimento ser invalidado diminuem, por não terem sido cumpridas as exigências definidas pelo método.

Outra característica, a favor do uso da ferramenta *TANGRAM-II Traffic Generator* para a estimativa das métricas de *Loss Pair*, é a precisão no cálculo da capacidade de transmissão do gargalo. O método usado para o cálculo da capacidade em [30] não é indicado. Porém, ele é feito a partir da "ida e volta" dos pares de pacotes. Acreditamos que o uso do método proposto, no Capítulo 5, para o cálculo da capacidade de transmissão, aumentará a precisão da estimativa. Pois, como o método *Loss Pair* prevê situação de alto tráfego concorrente, uma vez que a máquina "gargalo" deve estar descartando pacotes, o uso da seleção dos pares que estimarão a capacidade de transmissão pode ser considerada, no mínimo, bastante razoável. Porém, apesar de terem sido implementadas na ferramenta, não foram aplicados um número suficientemente grande de teste com este método para que resultados possam ser mostrados. Assim, propomos que estas avaliações sejam feitas como trabalhos futuros.

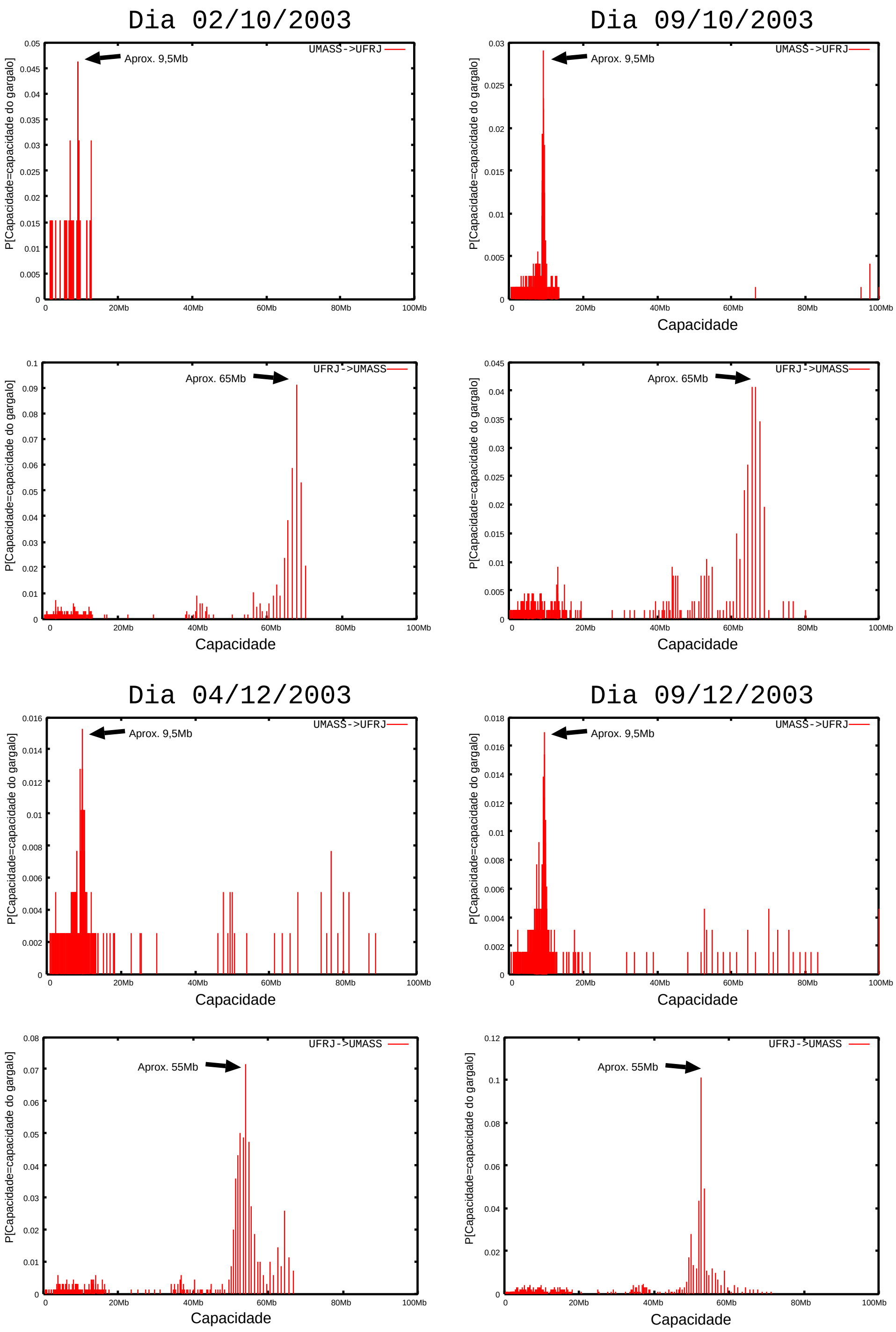

Figura 6.5: Estimativa da capacidade de transmissão do gargalo no caminho entre a UFRJ e a UMASS.

Capítulo 7

Conclusão

A necessidade de conhecer as características das redes, principalmente a Internet, é cada vez maior. A forma como a Internet foi concebida, baseada na técnica de comutação de pacotes com o uso de datagramas e sem nenhuma reserva de banda nem priorização, impede que garantias da qualidade do serviço sejam oferecidas. Conseqüentemente, dificulta que aplicações com requisitos mínimos para um bom funcionamento sejam implantadas. Por isso, conhecer o estado e as características da rede é fundamental para a implementação desse tipo de aplicações.

O objetivo deste trabalho foi fazer uma avaliação de métricas obtidas através de medições ativas, em especial as relacionadas ao atraso fim-a-fim, sem uso de infra-estruturas previamente montadas para esta finalidade. Para isso, algoritmos foram estudados, avaliados e implementados em uma ferramenta, outros foram propostos e também implementados, para possibilitar que sejam estimados os atrasos das sondas em um sentido. Todas as implementações foram validadas através de uma série de experimentos que foram realizados na Internet, o que possibilitou uma análise do atraso dos pacotes entre dois pontos da rede. Com a possibilidade de se estimar o atraso fim-a-fim, uma série de outras métricas podem também ser estimadas, o que resultou na proposta de novos métodos para obtenção das mesmas usando o atraso em um sentido.

7.1 Contribuições

Um dos resultados deste trabalho é uma ferramenta de medição ativa, ímpar no que diz respeito às medidas de desempenho estimadas por ela. As características disponíveis na ferramenta *TANGRAM-II Traffic Generator* estão descritas no Apêndice A deste trabalho. Usando essa ferramenta, vários experimentos foram executados na Internet, permitindo que algumas conclusões fossem tiradas deste trabalho.

Foram implementados e validados algoritmos da literatura para remoção do *Skew* e do *Offset* no cálculo do Atraso em um sentido.

Com os resultados dos experimentos e as validações executadas, é possível concluir que as estimativas feitas para o atraso em um sentido, sem uso de qualquer dispositivo para sincronização dos relógios, possuem alta precisão. Isso permite que experimentos sejam executados e avaliações do comportamento do atraso sofrido pelas sondas sejam estudadas. Os algoritmos para a estimativa do atraso em um sentido possuem um padrão de geração de sondas. Foi proposto um método que consiste na geração de sondas (segundo os padrões exigidos) em paralelo aos dados transmitidos por uma aplicação qualquer (exemplo: aplicação para transmissão de vídeo). Desta forma é possível estimar o atraso dos pacotes de uma aplicação, independente da forma como os dados são enviados por ela.

A série de experimentos executados entre a UFRJ e a UMASS, e entre a UFRJ e a UNIFACS, permitiu identificar outros problemas, causados por atualização nos relógios e congelamento do sistema aperacional. Foi constatado que esses problemas atrapalham, de forma considerável, as estimativas feitas do atraso das sondas. Porém, os algoritmos, que foram propostos neste trabalho, se mostraram eficientes na remoção do problema, possibilitando que os resultados do atraso em um sentido sejam estimados nos experimentos, mesmo com a ocorrência desses eventos.

Com os resultados dos experimentos, foram estimadas distribuições para caracterização do atraso das sondas baseadas no método de momentos. A identificação da distribuição que melhor caracteriza o atraso pode ser feita graficamente e através do

MSE. Os resultados coletados da Unifacs apontam a distribuição Gaussiana como melhor para caracterizar os atrasos obtidos nesse caminho.

O uso do atraso fim-a-fim para auxiliar a estimativa de outras métricas foi mais uma contribuição deste trabalho. Durante a série de experimentos executados entre a UFRJ e a UMASS, a geração de pares de pacotes foi também realizada. A partir dessa coleta, foram identificados problemas para a estimativa da métrica "capacidade de transmissão no enlace do gargalo". Assim, um método que seleciona os pares a serem usados no cálculo foi proposto e implementado. Os resultados obtidos com esse método mostraram-se mais precisos do que os obtidos com o método tradicional. No pior caso, o resultado do algoritmo proposto é igual à estimativa feita pelo método proposto na literatura que não faz a seleção dos pacotes.

A nova proposta para estimar a "capacidade de transmissão do gargalo" e a possibilidade de se estimar o atraso em um sentido, incrementa também a precisão na estimativa de outras métricas, como as obtidas pelo *Loss Pair*. Em muitos casos, como na proposta do método *Loss Pair*, as estimativas das métricas eram feitas baseadas no "atraso de ida e volta" dos pacotes, isso por não existir uma ferramenta disponível que estimasse o atraso fim-a-fim. Porém, as implementações atuais do *TANGRAM-II Traffic Generator* criam um novo paradigma para essas estimativas, com ele é possível calcular métricas relativas a um único sentido da rede, o que certamente aumenta a precisão e a qualidade da estimativa.

7.2 Trabalhos Futuros

Apesar das implementações na ferramenta *TANGRAM-II Traffic Generator* estarem concluídas, algumas avaliações podem ser feitas das propostas implementadas. Uma delas é a aplicação de outros testes na Internet, variando os pontos de coleta, e verificar se comportamentos semelhantes aos descritos neste trabalho são encontrados em outros testes.

Outra avaliação a ser realizada é a das estimativas feitas usando o *Loss Pair*. Para isso, seria adequado o uso de um ambiente controlado para a aplicação dos testes, onde são conhecidos parâmetros como o tamanho de *Buffer* dos equipamentos, e onde seja possível criar algumas situações, como o saturamento do equipamento gargalo e implementações de políticas de descarte. A validade dessa proposta já foi inclusive discutida durante as sessões de seminários entre a UFRJ e a UMASS, mas ainda há pontos bastante relevantes que devem ser estudados.

Mais um trabalho que deve ser executado é tentar quantificar a diferença de precisão entre as propostas para estimar a "capacidade de transmissão do gargalo". Durante este trabalho foi identificado que, para situações de alto tráfego, a proposta que seleciona os pares a serem usados no cálculo incrementa a precisão da estimativa. Porém, seria interessante avaliar o quanto a nova proposta é superior à proposta que não prevê a seleção dos pares baseados no atraso fim-a-fim da primeira sonda.

Referências Bibliográficas

[1] IPFIX WG - IETF Web Page. http://www.ietf.org/html.charters/ipfix-charter.html.

[2] Ipma project web page. http://www.merit.edu/ipma/.

[3] Nimi project web page. http://www.ncne.nlanr.net/nimi/.

[4] Ripe project web page. http://www.ripe.net/ttm/.

[5] Surveyor project web page. http://www.advanced.org/surveyor/.

[6] C. Dovrolis, P. R., e Moore, D. What do packet dispersion techniques measures? In *IEEE Infocom* (2001).

[7] Carmo, R., de Carvalho, L., de Souza e Silva, E., Diniz, M., e Muntz, R. Performance/Availability Modeling with the TANGRAM-II Modeling Environment. *Performance Evaluation 33* (1998), 45–65.

[8] Carter, R. L., e Crovella, M. E. Bprobe, 1996. http://www.cs.bu.edu/faculty/crovella/src/bprobe/Tools.html.

[9] Carter, R. L., e Crovella, M. E. Measuring bottleneck link speed in packet-switched networks. In *Performance Evaluation* (1996).

[10] Cheng, W. C. http://bourbon.cs.umd.edu:8001/tgif/. TANGRAM Graphic Interface Facility.

[11] Chuah, C.-N., e Katz, R. Characterizing packet audio streams from internet multimedia applications. In *IEEE ICC* (2002).

[12] CISCO. Netflow. ftp://www.cisco.com/netflow.

[13] D. Clark, J. Crowcroft, B. Davie, et al. Recommendations on queue management and congestion avoidance in the internet, April 1999. RFC 2309.

[14] Daniel Menasché, Edmundo de Souza e Silva, e Rosa Leão. Ferramenta de vídeo em tempo real e aplicações de ensino a distância. In *WRNP* (2003).

[15] de Souza e Silva, E. Leão, R. Duarte, F., et al. Modelagem e análise de redes com o conjunto de ferramentas tangram-ii. In *SBRC* (2003).

[16] Floyd, S., e Jacobson, V. Random early detection gateways for congestion avoidance. In *IEEE/ACM Transactions on Networking* (1993), pp. 397–413.

[17] G. Almes, S. K., e Zekauskas, M. A One-way Delay Metric for IPPM, September 1999. RFC 2679.

[18] Grossglauser, M., e Rexford, J. Passive traffic measurement for IP operations. to appear as a chapter in The Internet as a Large-Scale Complex System, Oxford University Press.

[19] Information Sciences Institute. Ns-2 simulator. http://www.isi.edu/nsnam/ns/.

[20] Jacobson, V. Congestion avoidance and control. In *ACM SIGCOMM* (1988).

[21] Jain, R. K. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling.* John Wiley and Sons, 1991.

[22] K. Harfoush, A. B., e Byers, J. Measuring bottleneck bandwidth of targeted path segments. In *IEEE Infocom* (2003).

[23] Keshav, S. A control-theoretic approach to flow control. In *ACM SIGCOMM* (1991).

[24] Lai, K., e Baker, M. Measuring bandwidth. In *IEEE Infocom* (1999).

[25] Lai, K., e Baker, M. Nettimer: A tool for measuring bottleneck link bandwidth. In *3rd USENIX Symposium on Internet Technologies and Systems* (2001).

[26] LAND. Tangram-II Traffic Generator. http://www.land.ufrj.br/tools/tangram2/traffic/traffic.h

[27] LAND. Tangram-IIv2.0. http://www.land.ufrj.br/tools/tangram2/tangram2.html.

[28] Lawrence Berkeley National Laboratory. Traceroute, 1989. ftp://ftp.ee.lbl.gov/traceroute.tar.gz.

[29] Li Zhang, Z. L., e Xia, C. H. Clock synchronization algorithms for network measurements. In *IEEE/Infocom* (2002).

[30] Liu, J., e Crovella, M. Using loss pairs to discover network. In *ACM SIGCOMM* (2001).

[31] Loung, D. D., e Biro, J. Needed services for network performance evaluation. In *IFIP Workshop on Performance Modeling and Evaluation of ATM Networks* (2000).

[32] Masato Tsuru, T. T., e Oie, Y. Estimation of clock offset from one-way delay measurement on asymmetric paths. In *SAINT International Symposium on Applications and the Internet* (2002).

[33] Mendenhall, W., e Sincich, T. *Statistics for Engineering and the Sciences*, 3rd ed. Macmillan, 1991.

[34] Mills, D. Network time protocol (version 3) - specification implementation and analysis, March 1992. RFC 1305.

[35] Muuss, M. Ping, 1989. ftp://ftp.arl.mil/pub/ping.shar.

[36] Pasztor, A., e Veitch, D. PC Based Precision Timing Without GPS. In *ACM Sigmetrics*. 2002.

[37] Paxson, V. On calibrating measurements of packet transit times. In *ACM/Sigmetrics* (1998).

[38] Rosa M.M. Leão, E. d. S. e. S., e de Lucena, S. C. A Set of tools for Traffic Modelling, Analysis and Experimentation. In *Proceedings of TOOLS'2000* (2000).

[39] Rytgaard, M. Estimation in Pareto distribution. *Austin Bulletin 20*, 2 (November 1990), 48–54.

[40] Sue B. Moon, P. S., e Towsley, D. Estimation and removal of clock skew for network delay measurements. In *IEEE/Infocom*. 1999.

[41] U. Hofmann, T. Pfeiffenberger, B. H. One-way-delay measurements with cm toolset. In *IPCCC* (2000).

[42] V. Paxson, A. A., e Mathis, M. Experiences with NIMI. *Proceedings of Passive and Active Measurement* (2000).

[43] V. Paxson, J. Mahdavi, A. A., e Mathis, M. An architecture for large-scale internet measurement. *IEEE Communications* (1998).

[44] V. Paxson, J. Mahdavi, A. A., e Mathis, M. Creating a scalable architecture for internet measurement. *INET* (1998).

[45] Van Jacobson, C. L., e McCanne, S. Tcpdump. ftp://ftp.ee.lbl.gov/tcpdump.tar.Z.

[46] W. Feng, D. Kandlur, D. S., e Shin, K. Blue: An alternative approach to active queue management. In *NOSSDAV* (2001).

[47] W. Feng, D. Kandlur, D. S., e Shin, K. Stochastic fair blue: A queue management algorithm for enforcing fairness. In *IEEE/INFOCOM* (2001).

[48] Williams, T., e Kelley, C. http://www.gnuplot.info/. GNUPLOT v3.7.

Apêndice A

TANGRAM-II Traffic Generator Nova Versão

AO longo de toda essa dissertação foram descritas implementações desenvolvidas para a obtenção de algumas medidas de interesse, validação dos algoritmos e avaliação dos experimentos coletados. O resultado de todas essas implementações é a segunda versão da ferramenta de medição ativa (*Traffic Generator*) do ambiente *TANGRAM-II*. Como foi visto na Seção 3.2, o gerador de tráfego foi concebido para dar ao usuário a possibilidade de geração de tráfego *IP* ou *ATM* nativo e a coleta de estatísticas deste tráfego gerado. Até sua segunda versão, a ferramenta possibilitava estimar apenas algumas métricas de interesse como: *jitter*, perda, sucesso e vazão. Em seguida, será apresentada a ferramenta em sua nova versão e a estrutura atual em que se encontra, assim como a interface gráfica desenvolvida para ela, onde é mantido um padrão gráfico semelhante ao construído no ambiente *TANGRAM-II* e na versão anterior do *TANGRAM-II Traffic Generator*.

A.1 Usando o *TANGRAM-II Traffic Generator*

Devido aos requisitos definidos para algumas das métricas de interesse, o formato da geração das sondas foi utilizado como parâmetro na criação da interface gráfica.

Basicamente, nove diferentes tipos de medições podem ser aplicados pela ferramenta. A variação destes modos de medição e as métricas estimadas por eles estão ilustradas na Figura A.1.

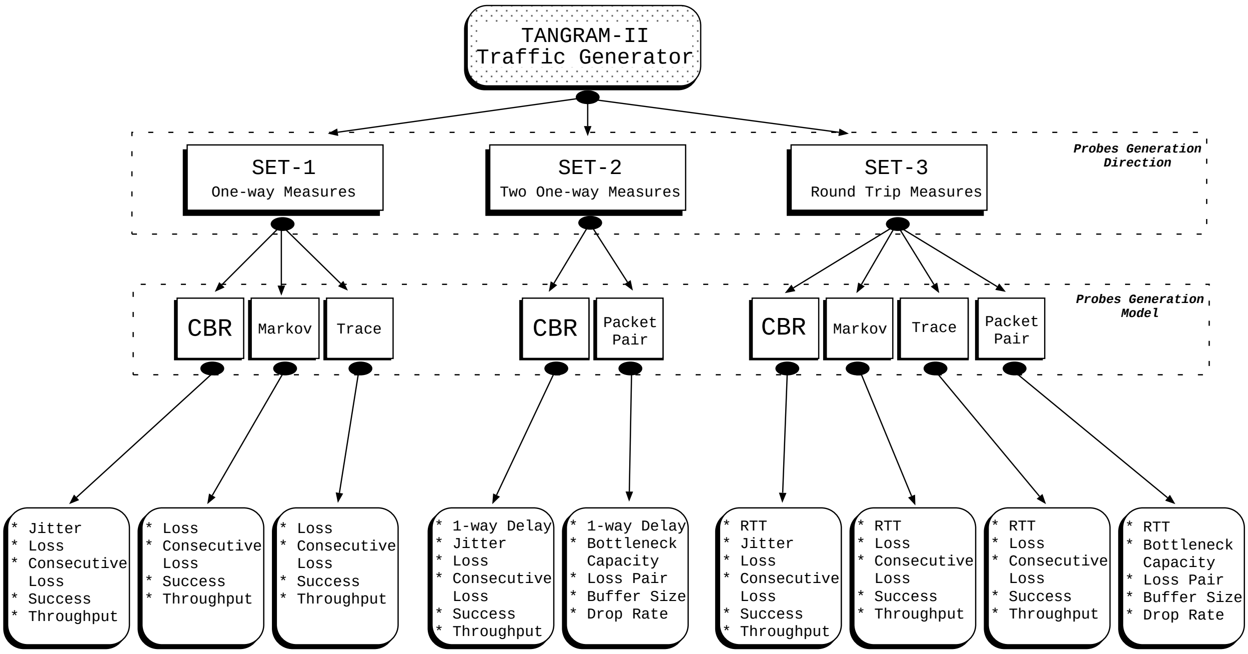

Figura A.1: Estrutura da nova versão da ferramenta.

A técnica a ser usada durante um experimento depende das métricas desejadas. Dessa forma, antes de executar uma medição, o usuário deve definir quais são as medidas de interesse e, assim, ativar a técnica mais apropriada para a execução do seu experimento. A Figura A.2 ilustra algumas telas que são encontradas pelos usuários do *Tangram-II Traffic Generator*.

Para ativar a técnica desejada na geração das sondas, dois níveis de escolha são requisitados ao usuário: primeiro, o usuário deve optar pelo sentido em que as sondas serão enviadas. Como já foi descrito anteriormente neste trabalho, o sentido utilizada durante a medição é, em muitos casos, pré-requisitos para a extração das medidas. Para cada sentido escolhido, um conjunto de modelos de geração de sondas é disponibilizado. Por isso, Set1, Set2 e Set3 representam as possíveis direções de geração das sondas disponíveis na ferramenta.

A segunda escolha a ser feita determina o modelo de geração das sondas que o usuário pretende usar, durante o experimento, com a ferramenta. Novamente, a

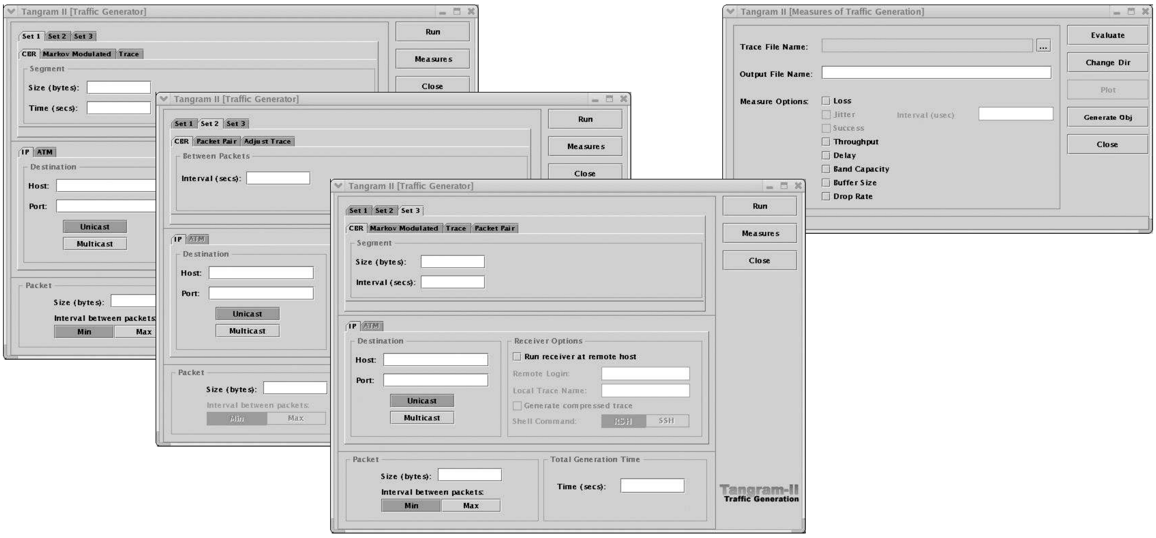

Figura A.2: Interfaces da nova versão da ferramenta.

seleção desse parâmetro deve ser em função das métricas de interesse desejadas pelo usuário.

Direção da Geração das Sondas

- *Set1:* Neste formato, sondas são geradas em um único sentido. Uma única fonte gera o tráfego que é coletado pelo destinatário, após atravessar um caminho da rede.

- *Set2:* Com esta opção, emissões de sondas são ativadas nos dois sentidos e de forma paralela. É como se dois geradores de tráfego, de um único sentido (Set1), fossem iniciados de forma simultânea, mas em sentidos opostos.

- *Set3:* Para trabalhar de forma semelhante à ferramenta *Ping*, um dos modelos do Set3 podem ser utilizados. Porém, diferente do *Ping* que trabalha com o protocolo *ICMP*, pacotes *UDP* são gerados do *Tangram-II Traffic Generator*. Outra diferença é a precisão de tempo e os vários modelos de geração das sondas, existentes nesta ferramenta, e que não se encontram disponíveis se usado o gerador de pacotes *ICMP*.

Modelo de Geração das Sondas

- *CBR*: No modelo de geração de tráfego CBR (*Constant Bit Rate*), as sondas são geradas seguindo um modelo determinístico, onde os intervalos entre os envios são constantes.

- *Modulado por Markov* No Modulado por Markov (ou *Markov Modulated*), o tráfego gerado segue um modelo de Markov, de tempo contínuo, com número finito de estados. Uma taxa de envio do tráfego associada a cada estado da cadeia deve ser definida.

- *Trace*: O modelo de geração "*Trace*" prevê que o envio das sondas de medição são determinadas por um arquivo de entrada. Este modelo possibilita a representação de situações reais de uso das aplicações na rede, como por exemplo a transmissão de um vídeo em tempo real, onde no *trace* constam o tamanho dos quadros (*frames*) e os instantes do tempo de geração dos pacotes.

- *Pares de Pacotes*: Neste caso são geradas rajadas de sondas consecutivas, em intervalos constantes de tempo.

Após a geração e a coleta de um determinado formato de medição, uma série de medidas podem ser estimadas. Em função da técnica utilizada na geração das sondas, as possíveis métricas ficam habilitadas na ferramenta para serem estimadas. O usuário pode selecionar as medidas de desempenho desejadas e a ferramenta, automaticamente, faz os cálculos necessários para cada uma das métricas gerando os gráficos dos resultados. Além da visualização gráfica básica dos resultados, uma série de variações podem ser aplicadas à ferramenta. Maiores detalhes da ferramenta poderão ser encontrados no manual do usuário, que é disponibilizado em conjunto com o manual do resto da ferramenta *Tangram-II* [27, 26].