DIVERSIDADE DE CAMINHOS PARA APLICAÇÕES EM TEMPO REAL

por

Bruno Felisberto Martins Ribeiro

UFRJ

Tese submetida para a obtenção do título de

Mestre em Ciências em Engenharia de Sistemas e Computação

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

da COPPE/UFRJ

por

Bruno Felisberto Martins Ribeiro

Julho 2003

DIVERSIDADE DE CAMINHOS PARA APLICAÇÕES EM TEMPO REAL

Bruno Felisberto Martins Ribeiro

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. Virgílio Augusto Fernandes Almeida, Ph.D.

Prof. Donald F. Towsley, Ph.D.

Prof. Valmir Carneiro Barbosa, Ph.D.

RIO DE JANEIRO, RJ - BRASIL

JULHO DE 2003

RIBEIRO, BRUNO FELISBERTO MARTINS

Diversidade de Caminhos para Aplicações em Tempo Real [Rio de Janeiro] 2003

XIV, 59 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia de Sistemas e Computação, 2003)

Tese - Universidade Federal do Rio de Janeiro, COPPE

1. Diversidade de Caminhos
2. Análise de Desempenho
3. Qualidade de Serviço
4. Análise Estatística de Perdas

I. COPPE/UFRJ   II. Título (Série)

Dedico este trabalho a todos os que me ensinaram

e me capacitaram a sentir-me um ser-humano completo,

em especial à minha maior professora, minha mãe.

Agradecimentos

Quero agradecer a todos, tentando não esquecer de ninguém, o que para mim significa um esforço sobre-humano; por isso mesmo, os agradecimentos estão em ordem cronológica, para facilitar.

Primeiramente à heroína que é minha mãe, exemplo de força, inteligência e coragem. A maior parte de TUDO o que sou e serei devo a ela, a minha melhor e mais exigente professora. Às minhas irmãs Aline e Flávia: que são pessoas quase tão maravilhosas quanto eu (quase :).

Aos meus famigliares: Tia Lúcia, Tio Fred, Renata e Luciana, que tiveram participação ativa na minha formação; amo todos vocês. Aos meus primos Alexandre, Rafael, à Tia Leda, ao Tio Luís e a todo o pessoal do Rocha.

A todos os meus amigos, especialmente aos meus amigos de colégio, que moram em meu coração, e acompanharam todo o meu desenvolvimento na "informática", desde o dia em que fui apresentado a um computador, onde ficamos um olhando para o outro esperando alguém saber ligar aquilo (um PC-XT), até os dias de hoje, onde vejo suas carreiras em franca ascensão; tenho muito orgulho de vocês todos, especialmente destes: Nine, Bruna, Carol, CD e Guto (Gordo). Não há como descrever cada um deles porque a dissertação é sobre diversidade de caminhos e cada um deles daria uma. Obrigado, adoro vocês! Um agradecimento também à professora Bernardete do CPII.

Aos meus amigos de graduação, cujas conversas tanto me enriqueceram e aos meus amigos da Intelig, empresa na qual fui tão bem recebido, mesmo sendo um peixe fora d'água.

A outro grande amigo meu, que não limitou-se a ser apenas um excelente orientador de Iniciação Científica, sendo um marco em minha vida. Uma pessoa fantástica, que me ensinou os primeiros passos em pesquisa, a qual tenho e sempre terei como exemplo a seguir; obrigado Collier.

Aos meus amigos do mestrado, que aos poucos aprenderam a conviver comigo, o que não é nada fácil. Ao Cesar, Guto, Bernardo, Kelvin, à Drikinha, ao Guilherme, Flávio, Alysson, Daniel e Fernando. Obrigado pelas discussões, esclarecimentos e pela amizade.

À Ana Paula, que fez dessa jornada um evento menos solitário e mais bonito; compartilhando de todas as minhas idéias, tomando ciência de cada passo da jornada, coitada. Uma pessoinha que se instalou em meu coração e que não sai mais (e não é porque ela é pequena demais para encontrar) e que muito me ajudou a escrever este trabalho.

Um agradecimento também à professora Rosa por ter tido extrema paciência comigo e por seus diversos ensinamentos durante este produtivo período. Agradecimentos também a todos os professores da COPPE: Valmir, Cláudio Esperança, entre outros, que tanto me ensinaram.

Também à pessoa que deveria estar trabalhando nas melhores universidade dos Estados Unidos, mas preferiu orientar alunos como eu, aqui no Brasil, e torná-los pesquisadores. Além de tudo, ainda tem um grande coração. Não tenho palavras para agradecê-lo; em verdade as tenho sim, o meu agradecimento, implícito, estará em cada *paper* publicado e em cada trabalho de sucesso que fizer pelo resto de minha vida. Muito obrigado Edmundo, espero que consiga corresponder à energia que foi gasta comigo e um dia, quem sabe, poder chegar aonde você chegou.

Por último e não menos importante, ao Povo brasileiro, que me deu a oportunidade de freqüentar, durante 14 anos, exceletes instituições educacionais públicas; espero ter sido um bom investimento.

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

DIVERSIDADE DE CAMINHOS PARA APLICAÇÕES EM TEMPO REAL

Bruno Felisberto Martins Ribeiro

Julho / 2003

Orientador: Edmundo de Souza e Silva

Programa: Engenharia de Sistemas e Computação

O envio de mídia continuamente por redes sem garantia de qualidade de serviço é um dos maiores problemas para a popularização de distribuição de conteúdo via Internet. Recentemente, muitos trabalhos foram publicados na literatura visando amenizar os efeitos da falta de qualidade da Internet, melhorando as estatísticas de perda fim-a-fim, alguns destes usando diversidade de caminhos, que é simplesmente enviar tráfego por dois ou mais caminhos distintos concomitantemente, sem replicação de pacotes. Uma das principais vantagens deste método reside no fato que não aumenta a quantidade de tráfego gerado ou não implica em modificações na estrutura da Internet.

A diversidade de caminhos foi inicialmente proposta por Maxemchuk para melhorar o roteamento de pacotes em redes comutadas. Recentemente foi usada para aplicações multimídia sobre redes ATM, Internet, redes sem fio e até ad-hoc. A proposta deste trabalho é estudar a técnica em detalhes, visando esclarecer alguns pontos não abordados na literatura, identificando problemas novos e alguns potenciais benefícios da abordagem de diversidade de caminhos. O primeiro ponto abordado será a performance da diversidade de caminhos quando usados mais de dois caminhos, nesta análise concluirá que, dependendo dos requisitos de qualidade da aplicação, um maior número de caminhos pode ou não ser vatajoso. O segundo ponto abordado será como o método se comporta quando uma quantidade substancial do tráfego da rede é tráfego que faz uso da diversidade de caminhos, onde se concluirá que o melhor cenário é quando apenas a fonte analisada usa a diversidade de caminhos.

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

PATH DIVERSITY FOR REAL TIME APLICATIONS

Bruno Felisberto Martins Ribeiro

July / 2003

Advisor: Edmundo de Souza e Silva

Departament: Engenharia de Sistemas e Computação

Streaming media over best-effort networks is one of the main issues for large scale deployment of content delivery over the Internet. A variety of techiniques have been proposed to cope with Internet's poor quality of service. One of these is to achieve better end-to-end loss behavior through path diversity. In this techinique traffic is sent to the receiver using more than one path, without replicating any packets. One of the main advantages of this method is that neither extra traffic is generated nor any modifications is required in the Internet's infrastructure.

Path diversity was first introduced by Maxemchuk as a routing mechanism. Recently the techinique has draw considerable attention as a way to improve the QoS as seen by the application and has been proposed for ATM, IP, wireless and ad-hoc networks.

The objective of this work is to investigate some of the path diversity issues that have not been addressed in the literature, ilustrating many of the existing tradeoffs. The first analysis is the performance of path diversity when more than two paths are used, where the conclusion is that increasing the number of paths may or may not be advantageous depending on the aplication. The second analysis is its scalability, i. e., when a large portion of the network traffic is path diversity traffic. This analysis will show that the best scenario for an application is when only the target source uses path diversity.

Sumário

Lista de Figuras

Lista de Tabelas

Palavras-chave

1. Diversidade de Caminhos.

2. Análise de Desempenho.

3. Transmissão de Mídia Contínua.

4. Qualidade de Serviço.

Capítulo 1

Introdução

NOS últimos anos, as aplicações de tempo real - transmissão de vídeo para entretenimento, telefonia IP, sites WWW multimídia, teleconferência, jogos iterativos, mundos virtuais, ensino a distância - têm crescido em importância e exigem uma qualidade de serviço que se diferencia bastante das exigências feitas pelos tráfegos de aplicações tradicionais como, por exemplo, WWW e aplicações de e-mail. Em particular, as aplicações de tempo real são altamente sensíveis aos atrasos fim-a-fim bem como a variação deste (*jitter*), e o descarte de pacotes que acontece principalmente devido ao congestionamento dos rotedores.

Para alcançar a qualidade de serviço exigida por este novo conjunto de aplicações transmitidas pela Internet, dois diferentes grupos de propostas podem ser encontradas na literatura: propostas que exigem mudança nos nós da rede (seja mudança de software ou hardware) ou propostas que são implementadas nas aplicações. No caso das propostas que exigem mudança na infra-estrutura de rede destacam-se as propostas em relação a arquitetura, como por exemplo redes ATM que durante estabelecimento da conexão asseguram a qualidade de serviço necessária à aplicação [?] e políticas de armazenamento, descarte e serviço nas filas dos roteadores, como por exemplo o DiffServ e IntServ[?] .

Como exemplos de propostas implementadas nas aplicações, pode-se destacar FEC (forward error correction)[?, ?], packet interleaving [?] e o mecanismo que será abordado neste trabalho: diversidade de caminhos [?, ?, ?, ?]. Todos estes métodos visam minimizar os problemas causados pela utilização da Internet, cuja natureza é best-effort, ou seja, nenhuma garantia é assegurada as aplicações que utilizam o seu meio físico.

O mecanismo de diversidade de caminhos foi proposto em 1975 por Maxemchuck em sua tese de doutorado [?]. Inicialmente, este método foi proposto para facilitar o roteamento dos pacotes, aumentar tolerância a falhas e diminuir o impacto do tráfego de uma fonte em particular sobre as demais fontes que compartilham canais. Somente a partir de 2001 este método passou a ser extensivamente estudado para fornecer qualidade de serviço ao tráfego multimídia enviado através da Internet. Diversos modelos analíticos e de simulação mostraram a eficiência da utilização do método de diversidade de caminhos para aplicações de voz e vídeo.

Apesar de o método ter recebido bastante atenção na literatura, muitas questões ainda continuam em aberto, propiciando uma análise mais profunda e cuidadosa deste método, principalmente questões em relação aos prós e contras da utilização da técnica de diversidade de caminhos. Este trabalho se direciona a responder algumas dessas questões.

1.1 Objetivos da Dissertação

Esta dissertação tem como objetivo avaliar o desempenho do método de diversidade de caminhos em vários cenários. O trabalho foca principalmente dois pontos ainda não abordados na literatura: melhora da performance do método quando se utilizam mais de dois caminhos e investiga a escalabilidade do método em relação ao número de aplicações usando a mesma técnica.

No capítulo 2 apresenta-se uma revisão bibliográfica que aborda os requisitos necessários às aplicações de tempo real, além de apresentar diversos métodos propostos na literatura que possuem como objetivo atender a estes requisitos. Neste capítulo destaca-se a definição do método de diversidade de caminhos, bem como as principais características inerentes a este.

No capítulo 3 apresentam-se os estudos realizados com o objetivo de responder a seguinte pergunta: Aumentar o número de caminhos traz algum benefício a aplicação que está utilizando o método de diversidade de caminhos? Para responder a esta pergunta, foram considerados dois diferentes cenários: a utilização de caminhos homogêneos, onde todos os caminhos possuem a mesma probabilidade de um pacote ser entregue à aplicação com erro, e a utilização de caminhos heterogêneos, onde os caminhos diferem no parâmetro citado anteriormente. Diversos resultados são obtidos, e uma discussão em relação a estes é apresentada.

No capítulo 4 apresentam-se os estudos realizados com o objetivo de analisar a escalabilidade do método, sendo que esta análise é fundamental para qualquer proposta de um novo método. Neste capítulo, diferentemente do capítulo 3, será mostrado que a diversidade de caminhos implica na dependência de perda entre os diversos caminhos. Uma outra questão abordada refere-se à suavização do tráfego quando considera-se o tráfego conjunto de todas as fontes que enviam dados pelos mesmos canais. Por final, será estudada a combinação destes dois fatores e os resultados decorrentes desta combinação.

No capítulo 5 são feitas considerações finais sobre esta dissertação, as contribuições do trabalho apresentado e propostas para trabalhos futuros.

Capítulo 2

Aplicações em Tempo Real na Internet

2.1 Introdução

OS últimos anos têm testemunhado o explosivo crescimento no desenvolvimento e implementação de aplicações de rede que transmitem e recebem conteúdo de áudio e vídeo pela Internet. Novas aplicações multimídia (também conhecidas como aplicações de mídia contínua) - vídeo para entretenimento, telefonia IP, teleconferência, mundos virtuais, monitoramento em tempo real e muitos outros - são anunciadas diariamente.

Para que as aplicações em tempo real sejam realmente difundidas, os usuários necessitam de qualidade em suas aplicações. Os requisitos de serviço das aplicações em tempo real diferem significativamente daqueles das aplicações orientadas a dados como Web - imagem e texto - e-mail e ftp. Em particular, aplicações multimídia são sensíveis ao atraso fim-a-fim e a sua variação (conhecida como *delay jitter*), e ao descarte de pacotes. Estes requisitos sugerem que uma arquitetura de rede que foi definida primariamente para comunicação de dados pode não suportar bem aplicações multimídia. Várias propostas para evitar estes problemas podem ser encontradas na literatura, desde propostas de arquiteturas, como ATM, até políticas de armazenamento, serviço e descarte nos roteadores, como, por exemplo, Diffserv (*Differentiated Service*) e Intserv (*Integrated Service*) [?]. O principal problema destas propostas é a premissa de que os nós da rede precisam ser alterados de alguma forma, seja por software ou hardware, já que a base maior da Internet é de roteadores IP, onde o *best-effort* predomina e não dá garantia de qualidade a nenhum tráfego. Até os dias de hoje estas mudanças são difíceis de serem feitas pela sua complexidade e custo.

Existem outras propostas para entrega de conteúdo de aplicações em tempo real que não garantem explicitamente requisitos de qualidade, porém não supõem alterações da infra-estrutura da Internet. Estas propostas não eliminam a perda de pacotes nem garantem suas restrições de tempo, apenas tentam minimizar os problemas causados pela falta de garantias da rede. Os mecanismos são geralmente implementados na aplicação final, muitas vezes tornando uma aplicação simples demasiadamente complexa. Para a variação do atraso fim-a-fim tem-se o mecanismo *playout delay* [**?**]- tempo entre o pedido do cliente pela mídia escolhida e a sua apresentação. Para a perda de pacotes alguns esquemas de recuperação são usados: FEC (*forward error correction*), *packet interleaving* e diversidade de caminhos (*path diversity)* [**?**].

Este capítulo tem como objetivo apresentar os principais conceitos inerentes às aplicações multimídia que estão se tornando comuns na Internet. Na seção 2.2 são apresentadas as principais características das aplicações multimídia atuais. A seção 2.3 apresenta os principais métodos encontrados na literatura utilizados para prover a qualidade de serviço necessária as aplicações de tempo real. A seção 2.4 mostra métodos utilizados para amenizar o problema de erro introduzido nas aplicações decorrente da perda de pacotes. A principal seção do capítulo é a 2.5 que apresenta as definições do método de diversidade de caminhos, os trabalhos relacionados, bem como os principais resultados encontrados na literatura. A seção 2.6 apresenta as conclusões do capítulo.

2.2 Aplicações Multimídia

Inúmeras aplicações multimídia, conforme já citado anteriormente, estão se tornando comuns nos dias atuais. Estas podem ser agrupadas em três classes: *streaming* de áudio e vídeo armazenados; *streaming* de áudio e vídeo em tempo real sem interatividade e aplicações de áudio e vídeo em tempo real com interatividade.

2.2.1 Fluxo de áudio e vídeo armazenados

Nesta classe de aplicações, clientes solicitam arquivos de áudio ou vídeo (compactados) que estão armazenados em um servidor. Entre estes arquivos, tem-se, por exemplo, músicas, aulas, filmes, programas de TV, eventos, desenhos, entre outros. Para esta classe, três pontos principais devem ser estabelecidos:

- Armazenamento da mídia - Previamente, a mídia deve ser gravada e armazenada no servidor. Assim, o cliente pode usar os mecanismos utilizados em videocassetes: *play, rewind, fast-forward;*
- *Streaming* - Neste mecanismo, o cliente começa a exibir a mídia poucos segundos após o seu pedido. Este mecanismo evita a necessidade de *download* de toda mídia antes do início da sua exibição, evitando uma espera desnecessária;
- Exibição contínua - Uma vez que se inicie a exibição da mídia, espera-se que o tempo de demonstração transcorra de acordo com o tempo original de gravação. Assim, restrições de atraso para a entrega dos dados são estabelecidas. Os dados devem ser recebidos a tempo de serem exibidos pelo cliente, caso contrário podem ser consideradas inúteis.

2.2.2 Fluxo de áudio e vídeo em tempo real sem interatividade

Esta classe de aplicações é similar ao sistema de rádio e televisão, exceto pelo fato das informações serem transmitidas pela Internet. Algumas características são particulares a este tipo de aplicações, como o fato da não possibilidade, por exemplo, da utilização do mecanismo *fast forward.* No entanto, como nas aplicações da classe da seção 2.2.1 a exibição contínua da mídia é requerida.

2.2.3 Áudio e vídeo em tempo real com interatividade

Esta classe de aplicações permite os clientes utilizarem áudio e vídeo de tal maneira que as informações podem ser trocadas entre os participantes de uma determinada seção, como por exemplo, videoconferência, telefonia IP. Para uma conversa com interação entre os diversos interlocutores, o atraso entre a fala e o movimento de um deles e a sua recepção pelos demais deve ser menor que poucos milisegundos. Para voz, atrasos menores que 150 milisegundos não são percebidos pelos usuários. No entanto, atrasos acima de 400 milisegundos afetam muito a qualidade da aplicação. Assim, restrições no atraso devem ser estabelecidas para que a aplicação tenha uma qualidade que satisfaça o usuário.

2.3 Métodos para melhorar a qualidade nas aplicações em tempo real

Como a Internet promove um serviço *best-effort*, onde não existe nenhuma garantia de entrega dos pacotes, do atraso fim-a-fim de um pacote individual ou da variação do atraso de pacotes dentro de um *stream* de pacotes, a melhoria da qualidade normalmente fica sob responsabilidade da aplicação. As aplicações de tempo real são as que possuem a maior demanda por qualidade, pois seus requisitos de atraso, variação do atraso e perda são sensivelmente maiores do que das aplicações como vídeo e áudio armazenados.

Além do *best*-effort, existem outras formas de transmissão de dados que, diferente do *best-effort*, possuem garantia de qualidade. Estes métodos supõem alterações no hardware ou software da rede. Nesta seção apresentam-se estes métodos, bem como o que deve ser feito quando opta-se por não se utilizar alterações explícitas, ficando com o serviço oferecido pelo *best-effort*.

2.3.1 Divisão dos recursos

A forma mais fácil de obter qualidade é garantir recursos de banda de forma exclusiva, mesmo que esta não esteja sendo utilizada. A telefonia convencional, por exemplo, é uma forma de reserva exclusiva. O atraso neste caso é determinístico e, se não considerarmos erros de transmissão, não há perda de pacotes.

Dois métodos de reserva de recursos são os mais utilizados: multiplexação por tempo (TDM - *time division multiplexing*) e por freqüência (FDM - *frequency division multiplexing*). Considere-se, por exemplo, um canal que suporte N diferentes nós e que sua taxa de transmissão é de R bps. TDM divide o tempo em frames de tempo, e cada frame de tempo é subdividido em N *slots* de tempo. Com a utilização do TDM, colisões são evitadas e a taxa de transmissão dedicada a cada nó é de R/N bps durante cada frame de tempo. A maior desvantagem deste método está no desperdício de banda: se somente um nó possui informações a serem transmitidas, ele deve esperar todos os slots de tempo destinados aos outros nós para transmitir novamente.

Enquanto o TDM divide o canal em termos de tempo, o FDM divide o canal em diferentes freqüências, sendo que cada nó terá a sua freqüência particular. As vantagens são as mesmas que as do TDM, sendo que também existe o problema de desperdício de banda.

Vale ressaltar que existem outros mecanismos que se destinam a divisão de recursos de maneira justa e que evite o desperdício dos recursos destinados às aplicações. No entanto, a idéia principal foi apresentar os métodos que garantem banda fixa e atraso determinístico.

2.3.2 Tráfego regulado

Na seção 2.3.1 apresentou-se mecanismos em que explicitamente os recursos são divididos, garantindo certamente a qualidade (banda, atraso). Os mecanismos apresentados nesta seção, no entanto, destinam-se ao policiamento dos tráfegos que passam por um caminho, garantindo a todos estes tráfegos a qualidade exigida quando estes foram admitidos na rede. No entanto, diferente do TDM e do FDM, esta garantia é estatística, ou seja, os roteadores da rede tentam atingir um certo nível de qualidade previamente acordado.

Entre estes mecanismos, pode-se citar o Intserv e o Diffserv [?]. O Intserv oferece garantias de qualidade por fluxo, diferente do Diffserv, que oferece garantias por conjunto de fluxos. As garantias do Intserv são melhores, mas exige muito mais processamento dos roteadores.

A grande desvantagem destes métodos é a alteração necessária na estrutura da rede e mesmo após as alterações, o processamento requerido por estas abordagem pode ser muito grande.

2.3.3 *Best-effort*

Nesta seção serão detalhados os problemas inerentes às redes *best-effort* e apresentados, sucintamente, os métodos que os minimizam e que não supõem nenhum tratamento especial da rede. A idéia é manter a complexidade da rede baixa e tentar diminuir os impactos da não garantia da qualidade de atraso e entrega.

2.3.3.1 Variação do Atraso

Um componente importante do atraso fim-a-fim na Internet é os atrasos aleatórios de permanência nas filas dos roteadores ao longo do caminho pelo qual são enviados os pacotes pertencentes a uma determinada aplicação. Devido a esta variação de atraso dentro da rede, o tempo entre o envio do pacote pela fonte de dados e o seu recebimento no receptor pode flutuar de pacote para pacote. Este fenômeno é chamado *jitter*.

Consideremos, como exemplo, dois pacotes que sejam enviados consecutivamente em uma aplicação de voz utilizando a Internet. Estes pacotes são gerados, no transmissor, a cada 20 ms. No entanto, no receptor a diferença entre os tempos de recebimento de cada um desses pacotes consecutivos pode ser maior ou menor que 20 ms, resultando, respectivamente, em um *jitter* positivo ou negativo. Se o receptor das informações ignora a presença do *jitter*, e exibe estas informações no momento que elas chegam, a qualidade do áudio pode se tornar extremamente ruim, afetando o entendimento da conversa entre os participantes.

Existem diferentes mecanismos na literatura destinados a minimizar os problemas gerados pela presença do *jitter* [?]. Entre estes, pode-se destacar o mecanismo de *playout delay,* onde o receptor armazena uma pequena porção do fluxo a ser utilizada antes que este seja escalonado para ser tocado. Para aplicações de tempo real com interatividade, o tempo de armazenamento não pode ser grande. Algumas propostas como *adaptative playout buffer* tentam encontrar uma boa relação entre o atraso total e o erro por demora na entrega, veja Towsley et al. [?] e Rosenberg et. al [?] para maiores detalhes.

2.3.3.2 Perda de Pacotes

Para aplicações multimídia como vídeo e áudio, a perda (tanto pelo enchimento das filas dos roteadores do caminho pelo qual a aplicação envia os dados, quanto por erros nos dados) de uma porção do fluxo é normalmente percebida pelo usuário como um ruído no fluxo de vídeo ou do áudio, afetando, sensivelmente, a qualidade da aplicação. Assim, diversos mecanismos para correção de perdas e, conseqüentemente, da melhoria da qualidade nas aplicações são apresentados na literatura [?].

Semelhante às aplicações que utilizam o TCP (*transport control protocol)* como protocolo da camada de transporte, as aplicações em tempo real podem utilizar o mecanismo de retransmissão para a recuperação de dados perdidos durante a transmissão destes. No entanto, as características da rede, bem como os requisitos de qualidade exigidos pela aplicação devem ser analisados. A retransmissão só é possível quando há tempo para se pedir novamente o pacote, e a retransmissão chegar (*round trip time* ou RTT) antes que o pacote seja necessário. Se o RTT é alto, isso significa que o tempo entre a chegada do pacote e a sua utilização é alto, fato inaceitável para muitas aplicações em tempo real, especialmente as com interatividade que, conforme visto anteriormente, possuem severos requisitos de atraso.

Como mecanismos mais utilizados para a correção de perda, podemos citar os mecanismos baseados em redundância, na utilização dos pacotes que já foram recebidos pela aplicação e mais recentemente, o mecanismo baseado na diversidade de caminhos [?, ?], que pode ser utilizado em conjunto com outros mecanismos, aumentando a sua eficiência.

2.4 Amenização do erro introduzido pela perda de pacotes nas aplicações em tempo real sobre redes *best-effort*

Nesta seção serão apresentadas as técnicas de correção de erro. O objetivo central deste trabalho é realizar um estudo profundo do mecanismo baseado na diversidade de caminhos, mostrando seus pontos fortes e fracos, alguns já apresentados na literatura e outros que ainda não foram abordados. Para completude do trabalho, na seção 2.4.1 serão apresentados os pontos fortes e fracos dos mecanismos baseados em redundância e na utilização de pacotes que já foram recebidos pela aplicação. Na seção 2.5, será abordado os principais conceitos e trabalhos relacionados na literatura a respeito do mecanismo de diversidade de caminhos, objetivo central do trabalho apresentado.

2.4.1 Recuperação de erros em aplicações em tempo real - redundância e utilização de pacotes recebidos

As principais formas de correção de erro presentes na literatura especializada são: inserção de silêncio, repetição do último pacote, interpolação de pacotes adjacentes, codificação do mesmo sinal com qualidade inferior inserindo esta redundância no fluxo e codificação com múltiplas descrições. O trabalho concentrar-se-á nas duas últimas e em suas aplicações na transmissão de vídeo e áudio.

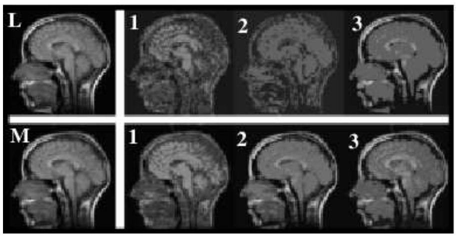

Figura 2.1: Exemplo de codificação por camadas (**L**) (acima) e por MDC (**M**) (abaixo).

2.4.1.1 *Multiple Description Coding* (MDC)

Codificação de múltiplas descrições (MDC), descrita em Kubin e Kleijn [?], é uma técnica onde o vídeo ou o áudio são reconstituídos através de várias descrições do original. Esta abordagem difere da codificação normal, onde as descrições são feitas umas a partir das outras, ou seja, existe uma descrição base, por exemplo, um quadro do tipo **I** num vídeo MPEG, e as outras são a diferença entre esse quadro e a imagem original. Na codificação MDC sempre os quadros capturam elementos diferentes da imagem ou da voz originais. A Figura 2.1 mostra a diferença entre as duas codificações.

Codificação em múltiplas descrições possui vantagens quando ocorrem perdas ou na ausência destas. Quando não há perda, os quadros "extras" podem ser usados para melhorar a imagem, diferente dos esquemas tradicionais de FEC, onde a redundância é jogada fora. Quando há perda de um quadro, os demais quadros que chegaram podem dar uma razoável reconstituição da imagem original. Na figura 2.1 fica claro que se forem perdidos os quadros 3 e 1 da codificação por camadas, a imagem ficaria quase irreconhecível. Por outro lado, se forem perdidos os quadros 1 e 3 da codificação MDC, ainda se poderia ter uma boa aproximação da imagem original.

2.4.1.2 *Forward Error Correction* (FEC)

A principal abordagem utilizada para recuperar a perda de pacotes é o envio de redundância, seja em pacotes separados, ou inserido em pacotes do próprio fluxo. Este método é chamado de *Forward Error Correction* (FEC). Alguns métodos pro-

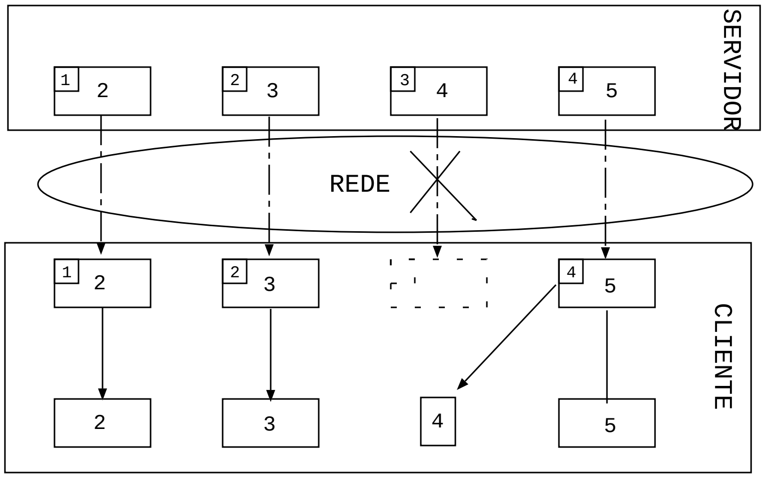

Figura 2.2: Exemplo de correção de erro com FEC, redundância codificada a uma taxa menor.

postos na literatura podem ser encontrados em D. R. Figueiredo e E. de Souza e Silva [?], Towsley et al. [?] e Dube et al. [?].

FEC pode ser usada de diversas formas; a mais comum é codificar cada amostra com uma taxa de compressão maior, resultando em amostras de pior qualidade, e inserir a n-ésima amostra comprimida no pacote da amostra $n+1$, conforme mostrado na Figura 2.2. Ao perder um pacote qualquer, a aplicação procura usar a sua redundância (do pacote), conforme exemplificado pelo pacote 3 na Figura 2.2.

FEC tem o problema de requerer maior taxa e ser pouco justo com as conexões TCP, pois no momento em que há maior perda, o fluxo de tempo real necessita de mais tráfego, aumentando assim a perda. Alguns métodos como Bolot, Fosse-Parisis e Towsley [?] diminuem este impacto fazendo controle da taxa máxima de envio. Outros métodos podem garantir a recuperação integral dos quadros, como em Figueiredo e de Souza e Silva [?].

2.4.1.3 Limitações da eficiência dos métodos de correção de erro

Apenas a taxa de perda não fornece informação suficiente para que se possa definir o melhor esquema de recuperação. É necessário, também, saber a distribuição do número de perdas consecutivas, bem como em alguns casos, a distribuição do número de pacotes recebidos entre duas perdas. Se as perdas de pacotes acontecem em rajadas, como verificado por experimentos na Internet (ver por exemplo [?, ?]), os esquemas de correção de erro descritos nas seções 2.4.1.2 e 2.4.1.1 podem não ser muito eficientes, especialmente se as rajadas são muito longas e a aplicação em tempo real possui requisitos de atraso muito próximos do tempo de viagem dos pacotes. Isso porque a aplicação se utiliza dos pacotes recebidos para a reconstituição (ou aproximação) do pacote perdido. Caso muitos pacotes sejam perdidos de uma só vez, não serão recebidos pacotes suficientes para "reconstituir" os pacotes perdidos. Uma forma de resolver o problema com FEC é fazer redundância do pacote $n + k$ no pacote n, com k maior que o tamanho da rajada de perda. Contudo, o emissor precisaria manter um atraso na recepção dos pacotes de, pelo menos, k pacotes, o que pode ser inviável para aplicações em tempo real com restrições no atraso, conforme dito anteriormente.

O problema das perdas em rajada, pode ser minimizado por dois métodos: *packet interleaving* e o uso de múltiplos caminhos (*path diversity*). *Packet interleaving* é feito enviando pacotes fora de seqüencia, conforme mostrado na Figura 2.3 Contudo, fica claro que, conforme acontece com o esquema cuja redundância do pacote n é enviada com o $n + k$, o receptor precisa armazenar k pacotes antes de reproduzi-los. Se o atraso é pequeno e as rajadas de erro são curtas, a técnica é eficiente, pois diminui a probabilidade de ocorrer uma rajada grande de perdas da seqüencia original. Caso contrário, o atraso da chegada de pacotes inserido pela rede, somado ao atraso do mecanismo (que é diretamente proporcional ao tamanho esperado das perdas), impossibilita a sua adoção se os requisitos de tempo são muito curtos, da mesma forma que o método de FEC quando o k é grande.

A utilização de diversidade de caminhos (múltiplos caminhos ou *path diversity*) para aplicações em tempo real também tem a finalidade de diminuir o tamanho da rajada de perda só que, diferente do método de *interleaving*, não há necessidade que o receptor armazene os pacotes recebidos antes de entregá-los ao decodificador. Na seção 2.5 será visto como a diversidade de caminhos funciona e quais são os trabalhos realizados na área.

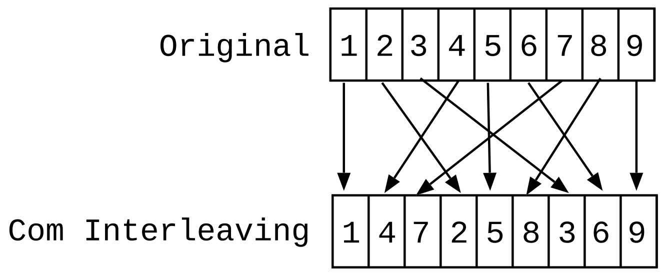

Figura 2.3: Exemplo de *packet interleaving*

2.5 Trabalhos Relacionados

2.5.1 Introdução

O objetivo principal da diversidade de caminhos é diminuir a probabilidade de ocorrer uma rajada longa de perda de pacotes, para que métodos de correção de erro como o FEC ou MDC possam ser mais eficazes. A diversidade de caminhos pode também atenuar as rajadas de fluxo de tráfego na rede, prover maior proteção contra *outages* de canais e ainda diminuir o *playout buffer* de algumas aplicações, como será visto adiante.

O modelo mais comum de diversidade de caminhos é o "Nth path, dual round robin", que é enviar os pacote em seqüência por cada um dos N caminhos.

Proposto inicialmente como um método de roteamento de pacotes, a diversidade de caminhos foi descrita com uma técnica para se obter uma transmissão de dados confiável para tráfegos comutados por circuitos (reais ou virtuais, como na tecnologia ATM (*Assyncronous Transfer Mode*)), ao mesmo tempo infringindo menor carga à rede que os métodos que utilizam apenas um caminho para a transmissão de dados. O mecanismo foi inicialmente proposto em 1975 por Maxemchuck em sua tese de doutorado [?], como um mecanismo de roteamento de pacotes. O objetivo de Maxemchuck foi diminuir as rajadas de tráfego na rede e evitar paralisações temporárias na transmissão de pacotes quando houvessem paralizações na rede. Como os algoritmos de roteamento não são rápidos o suficiente para selecionar outra rota alternativa em caso de uma paralização temporária, a diversidade de caminhos

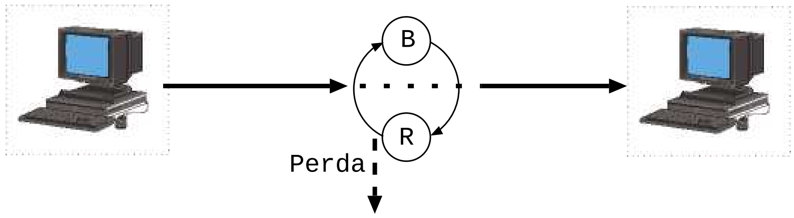

Figura 2.4: Modelo de perda de Gilbert.

é uma alternativa interessante para lidar com este problema.

Outros trabalhos foram escritos sobre as aplicações dessa abordagem, chamada até então de *dispersity routing*. Banejea [?] fez uso de simulações mostrando que o método permite, comparando as mesmas estruturas de rede, um número maior de fluxos do que a abordagem convencional, aumentando ao mesmo tempo, a confiabilidade da transmissão. Existe ainda um *survey* de 1997 [?] sobre *dispersity routing*. Cidon et al. [?] analisou *multi-path routing* (roteamento por múltiplos caminhos) em redes ATM, com vistas a diminuir o tempo de estabelecimento de uma conexão ATM.

Diversidade de caminhos para tráfegos multimídia, em redes comutadas por pacotes sem qualidade de serviço (*best-effort*), só seria extensivamente estudado em 2001. Abaixo serão abordados alguns pontos dos trabalhos já publicados, que serão analisados com maiores detalhes nas sessões 2.5.3 e 2.5.4. Os diversos trabalhos publicados têm muitos pontos em comum, como o uso de redes *best effort*.

Um ponto em comum em todos os trabalhos que fazem uso de modelos matemáticos para estudar o ganho da técnica de diversidade de caminhos é o modelo utilizado para a perda de pacotes, o modelo de Gilbert (Figura 2.4). Neste modelo o caminho perde pacotes se o seu estado (modelado por uma cadeia de Markov de dois estados) está em "Ruim" (**R** da Figura 2.4). A taxa de transição do estado "Bom" para o "Ruim" é μ_{Bom} e de "Ruim" para "Bom" é μ_{Ruim}.

Seja o intervalo entre envio de pacotes constante e igual a τ, a fórmula para a probabilidade de perda de um pacote dado que o pacote anterior foi ou não perdido é [?]

$$p_{i,j}(\tau) = \begin{cases} \frac{\mu_{\text{Ruim}}}{\mu_{\text{Bom}}+\mu_{\text{Ruim}}}\left(1 - e^{-[\mu_{\text{Bom}}+\mu_{\text{Ruim}}]\tau}\right) & i = \text{Ruim}, j = \text{Bom} \\ \frac{\mu_{\text{Bom}}}{\mu_{\text{Bom}}+\mu_{\text{Ruim}}}\left(1 - e^{-[\mu_{\text{Bom}}+\mu_{\text{Ruim}}]\tau}\right) & i = \text{Bom}, j = \text{Ruim} \\ 1 - p_{\text{Ruim,Bom}}(\tau) & i = \text{Ruim}, j = \text{Ruim} \\ 1 - p_{\text{Bom,Ruim}}(\tau) & i = \text{Bom}, j = \text{Bom} \end{cases}. \quad (2.1)$$

Pela equação 2.1, para o mesmo valores de μ_{Bom} e μ_{Ruim}, quanto maior a taxa de transmissão maior o número de perdas consecutivas.

O modelo de Gilbert pode ser considerado um bom modelo para a perda fim-a-fim. Segundo Yajnik et al. [?] e suas referências, o modelo de Gilbert é adequado para uma grande quantidade de experimentos feitos na Internet. Assim, este será o modelo de perdas adotado durante o trabalho, assim como foi o modelo adotado pelos outros trabalhos relacionados.

Vale ressaltar que as fontes de tráfego em tempo real podem ou não enviar pacotes em intervalos regulares de tempo. Durante este trabalho, e nos trabalhos listados acima, assume-se que os intervalos são regulares, o que de fato não compromete os principais resultados da análise, como será visto posteriormente.

2.5.2 Liang et al. [?]

Baseado na idéia de que a variação do atraso (*jitter delay*) está relacionado com a probabilidade de perda de pacotes durante sua transmissão, e que esta variação afeta a política de *playout buffer,* o trabalho de Liang et al. [?] visou estudar a determinação do *playout buffer* quando os pacotes da aplicação trafegam por diferentes caminhos, e, possivelmente diminuindo a probabilidade de perda .

Os autores realizaram experimentos na Internet entre dois diferentes pares origem / destino. Além de experimentos reais, também foram realizadas diferentes simulações, possibilitando a utilização de diferentes topologias e diferentes cargas de tráfego. Com estas simulações, foi possível mostrar que a utilização do mecanismo de diversidade de caminhos forneceu uma melhoria significativa na qualidade de voz e uma redução no atraso de entrega dos pacotes. Esta redução no atraso é essencial quando o esquema de playout buffer é usado, já que não existe a necessidade de um grande *buffer* para que se inicie a apresentação dos pacotes de voz.

Como demais resultados, foi mostrado que a eficiência da diversidade de caminhos depende da estatística de correlação entre estes caminhos e que o ganho obtido com a utilização deste mecanismo depende da diferença no atraso de propagação dos múltiplos caminhos e da variância do atraso da rede. Se um canal com baixa capacidade é compartilhado por múltiplos caminhos, foi mostrado que a correlação do atraso entre os fluxos é muito grande. Apesar disso, foi mostrado que a taxa de perda pode ser reduzida utilizando a diversidade de caminhos.

Outros trabalhos de Liang et al. [?, ?] também focam na diminuição do *playout buffer* usando diversidade de caminhos, mas ao mesmo tempo testando o comportamento de um fluxo vídeo (*streaming vídeo*) usando *Multiple Desciption Coding* (MDC). Os trabalhos mostram que usar múltiplos caminhos resulta em uma qualidade de vídeo melhor do que quando apenas um caminho é usado.

2.5.3 Apostolopoulos et al. [?, ?, ?, ?]

Nos trabalhos Apostolopoulos et al. [?, ?, ?, ?], a diversidade de caminhos foi utilizada em conjunto com aplicação de vídeo. O fluxo de vídeo é decodificado utilizando múltiplas resoluções (*multiple descriptions*), onde cada resolução é enviada por um caminho diferente, utilizando o mecanismo de diversidade de caminhos. Nestes trabalhos, foram modeladas e testadas as rajadas de perda e a distorção do vídeo recebido, assim como simulados os efeitos de caminhos que compartilham os mesmos pontos de congestionamento, com caminhos homogêneos e heterogêneos para no máximo dois caminhos. Nos trabalhos, três principais benefícios em relação à abordagem de apenas um caminho são listados: melhor comportamento da rede (uma vez que a aplicação vê um comportamento médio dos caminhos, i.e., média do atraso, média da perda, e média de outras estatísticas), menores rajadas de erro, menor probabilidade de *outage* e melhor qualidade de vídeo.

2.5.4 Golubchick et al. [?]

Golubchick et al. [?] tem como objetivo uma análise mais profunda das estatísticas de perda de pacotes usando múltiplos caminhos e é, entre todos os trabalhos analisados, o que mais se assemelha ao trabalho aqui proposto. O trabalho analisou a difusão de fluxos de vídeo pré-gravado, usando caminhos disjuntos (que não compartilham os mesmos pontos de congestionamento), com e sem FEC. O trabalho indica que a diversidade de caminhos, em geral, fornece melhores estatísticas de perda em rajada, autocorrelação de *lag* um, melhor relação sinal ruído e que a utilização de FEC juntamente com diversidade de caminho, fornece estatísticas ainda melhores. Contudo, distribuir vídeos pré-gravados não se enquadra exatamente na noção de aplicações em tempo real da forma como foram definidas durante este trabalho. Para aplicações de *Streaming* de vídeo e áudio pré-gravados existem técnicas, como *packet interleaving*, que podem ter quase a mesma eficiência da diversidade de caminhos, se achados os parâmetros certos, sem por isto herdar alguns de seus problemas.

2.5.4.1 Análise da probabilidade de perda em rajada

O trabalho mais completo encontrado na literatura sobre a probabilidade de perda em rajada é o de Golubchick et al. [?]. Neste é analisada a probabilidade de perda em rajada quando são usados apenas dois caminhos na diversidade de caminhos. Para tal, foi utilizado o modelo de Gilbert para caracterizar as potenciais correlações entre perdas consecutivas de um caminho e quando usados dois caminhos, os modelos de Gilbert dos caminhos são independentes.

Assumindo que a taxa de transmissão da aplicação não afeta as características de perda canal, então a probabilidade de perda de um pacote quando é usado apenas um caminho é igual à probabilidade do canal estar "Ruim", ou seja,

$$P_{sp}[\text{perda de pacote}] = \pi_{\text{Ruim}} = \frac{\mu_{\text{Bom}}}{\mu_{\text{Bom}} + \mu_{\text{Ruim}}}.$$

Para o caso de múltiplos caminhos, assuma que tem-se $N > 1$ caminhos e seja α_ia fração de pacotes enviados pelo caminho i, onde $\sum_{i=1}^{N} \alpha_i = 1$. Então a média da perda para o caso de múltiplos caminhos é

$$P_{mp}[\text{perda de pacote}] = \sum_{i=1}^{N} \alpha_i \frac{\mu_{\text{Bom}}}{\mu_{\text{Bom}} + \mu_{\text{Ruim}}}.$$

Se estes N caminhos são homogêneos, pode-se simplificar a expressão acima por

$$P_{mp}[\text{perda de pacote}] = \pi_{\text{Ruim}} = \frac{\mu_{\text{Bom}}}{\mu_{\text{Bom}} + \mu_{\text{Ruim}}}.$$

O trabalho de Golubchick et al. [?] também analisa o tamanho condicional da perda em rajadas. O caso de um caminho apenas foi descrito inicialmente por Bolot et al. [?], sendo a probabilidade do erro em rajada ser maior que $m \geq 1$ é:

$$P_{sp}[\text{rajada de perda de tamanho } m] = \tag{2.2}$$
$$\begin{cases} \pi_{\text{Bom}} p_{\text{Bom,Ruim}}(\tau) p_{\text{Ruim,Bom}}(\tau) & \text{para } m = 1 \\ \pi_{\text{Bom}} p_{\text{Bom,Ruim}}(\tau) \left[p_{\text{Ruim,Ruim}}(\tau)\right]^{m-1} p_{\text{Ruim,Bom}}(\tau) & \text{para } m \geq 2 \end{cases}.$$

A probabilidade de ter uma rajada de perda de qualquer tamanho é portanto

$$P_{sp}[\text{rajada de perda}] = \sum_{m=1}^{\infty} P_{sp}[\text{rajada de perda} = m] = \pi_{\text{Bom}} p_{\text{Bom,Ruim}}(\tau).$$

A probabilidade condicional de ter uma rajada de perda de pacotes de tamanho $m \geq 1$ é igual a

$$\begin{aligned} & P_{sp}[\text{rajada de perda de tamanho } m \mid \text{rajada de perda}] = \\ & \frac{P_{sp}[\text{rajada de perda de tamanho } m]}{P_{sp}[\text{rajada de perda}]} = \\ & \left[p_{\text{Ruim,Ruim}}(\tau)\right]^{m-1} p_{\text{Ruim,Bom}}(\tau), \qquad \text{para } m \geq 1 \end{aligned} \tag{2.3}$$

O trabalho de Golubchick et al. [?] faz também a mesma análise para dois caminhos e supõe que ambos os caminhos possuam as mesmas características:

$$P_{mp}[\text{rajada de perda de tamanho } m] = \tag{2.4}$$
$$\begin{cases} \pi_{\text{Bom}} \pi_{\text{Ruim}} p_{\text{Bom,Bom}}(2\tau) & \text{para } m = 1 \\ \pi_{\text{Bom}} \pi_{\text{Ruim}} \left[p_{\text{Ruim,Ruim}}(2\tau)\right]^{m-2} p_{\text{Bom,Ruim}}(2\tau) p_{\text{Ruim,Bom}}(2\tau) & \text{para } m \geq 2 \end{cases},$$

esta fórmula pode ser entendida da seguinte maneira:

Seja um instante aleatório do envio de um pacote qualquer, a probabilidade deste pacote não ser perdido é π_{Bom}. A probabilidade de perder o próximo pacote no outro camino é π_{Ruim} pois os caminhos são independentes. Para $m = 1$, o pacote seguinte ao perdido tem que ter sido recebido, este evento é igual à probabilidade do canal pelo qual se enviou o primeiro pacote ainda estar em "Bom": $p_{\text{Bom,Bom}}(2\tau)$. No caso $m \geq 2$, o canal pelo qual se enviou o primeiro pacote terá que perder o terceiro pacote, cujo evento tem probabilidade $p_{\text{Bom,Ruim}}(2\tau)$, a partir do terceiro pacote os pacotes são perdidos a medida que ambos os canais continuam ruins no momento do envio dos pacotes, estes eventos têm probabilidade $[p_{\text{Ruim,Ruim}}(2\tau)]^{m-2}$, finalmente, o último evento, é o recebimento do $(m+2)$-ésimo pacote, cujo evento tem probabilidade $p_{\text{Ruim,Bom}}(2\tau)$.

A probabilidade de ocorrer uma perda em rajada de qualquer tamanho é

$$P_{mp}[\text{rajada de perda}] = \sum_{m=1}^{\infty} P_{mp}[\text{rajada de perda de tamanho } m] = \pi_{\text{Bom}}\pi_{\text{Ruim}}.$$

A probabilidade condicional de ocorrer uma perda de pacotes em rajada de tamanho $m \geq 1$, condicionado em que houve uma rajada de erro é:

$$\begin{aligned} &P_{mp}[\text{rajada de perda de tamanho } m \,|\, \text{rajada de perda}] = \\ &\frac{P_{mp}[\text{rajada de perda de tamanho } m]}{P_{mp}[\text{rajada de perda}]} = \\ &\begin{cases} p_{\text{Bom,Bom}}(2\tau) & \text{para } m = 1 \\ [p_{\text{Ruim,Ruim}}(2\tau)]^{m-2}\, p_{\text{Bom,Ruim}}(2\tau) p_{\text{Ruim,Bom}}(2\tau) & \text{para } m \geq 2 \end{cases} . \end{aligned} \tag{2.5}$$

A partir da comparação da probabilidade condicional de ocorrer uma perda de pacotes em rajada de tamanho $m \geq 1$ de apenas um caminho com a de dois caminhos, pode-se mostrar que se $p_{\text{Bom,Ruim}}(2\tau)p_{\text{Ruim,Bom}}(2\tau) \leq p_{\text{Ruim,Ruim}}(\tau)p_{\text{Ruim,Bom}}(\tau)$, então múltiplos caminhos oferecem sempre uma probabilidade condicional de perda em rajada menor que quando apenas um caminho é utilizado.

2.5.4.2 Autocorrelação da perda usando Gilbert como modelo de erro

Outra importante medida analisada em [?] foi a autocorrelação entre perdas. Seja $X(t) \in \{0,1\}$o processo estocástico que vale um quando há uma perda de um pacote no tempo t e zero caso contrário. A autocorrelação é definida como:

$$Cor(X(t), X(t+\tau)) = \frac{E[(X(t)X(t+\tau)] - E[X(t)]E[X(t+\tau)]}{E[X(t)^2] - E[X(t)]^2}$$

Resultados obtidos em [?] mostraram que a autocorrelação da perda (lag-1) para apenas um caminho pode ser expressa da seguinte maneira:

$$Cor(X(t), X(t+\tau)) = \frac{[\mu_0 + \mu_1]p_{1,1}(\tau) - \mu_0}{\mu_1},$$

donde se conclui que a autocorrelação quando $\tau \to \infty$ tende a zero e quando $\tau \to 0$ tende a um.

Lema 1. *Da mesma forma, a autocorrelação da perda lag-1 para dois caminhos é* $Cor(X(t), X(t+\tau)) = 0$.

Demonstração. $X(t)$ e $X(t+\tau)$ são variáveis aleatórias independentes, por isso $Cor(X(t), X(t+\tau))$ é zero. □

O lema 1 leva a concluir que não há autocorrelação de perda entre pacotes consecutivos, uma vez que foi assumida a independência entre as perdas dos dois caminhos.

A autocorrelação é um indicativo da interdependência entre o pacote enviado no tempo t e o pacote enviado no tempo $t+k\tau$, $k \in \mathbf{N}$. Quanto maior a autocorrelação, maior a probabilidade de se o pacote enviado no tempo t for perdido, o pacote enviado no tempo $t+k\tau$ também será perdido, dessa forma para diminuir as rajadas de perda a redução da autocorrelação entre perdas é desejada.

2.5.5 Outros Trabalhos

Os trabalhos de Lin et al. e Wang et al. [?, ?] têm como objetivo aplicar múltiplos caminhos a redes sem fio, a fim de minimizar os problemas de perda de pacotes e escassez de banda (capacidade de transmissão) inerentes a estas redes.

2.6 Conclusões do capítulo

Conforme foi apresentado neste capítulo, a utilização da diversidade de caminhos contribui para a melhoria da qualidade de serviço de diferentes aplicações na Internet, como voz e vídeo, principalmente se usadas em tempo real. Neste sentido muitos trabalhos tiveram importância vital para o desenvolvimento da técnica. No entanto, ainda há carência por um estudo detalhado que apresente, além dos benefícios do mecanismo, os cenários em que a utilização da diversidade de caminhos não apresenta ganho significativo de qualidade. Ou seja, deve-se apresentar um trabalho que não fique restrito somente no lado positivo do mecanismo, mas que também apresente os seus pontos negativos. Além dos pontos negativos, há necessidade também de um estudo sobre as conseqüências da larga utilização deste método.

Baseado neste princípio, o presente trabalho visa a apresentar um estudo detalhado de diversas características da diversidade de caminhos, juntamente com os seus prós e seus contras. Serão abordados alguns pontos considerados, até então triviais na literatura, onde a conclusão é exatamente a oposta à conclusão "trivial" encontrada na literatura. Estes serão os temas dos dois próximos capítulos.

Um dos pontos em que o presente trabalho se diferencia do trabalho de Golubchick et al. [?] é ao analisar a diversidade de caminhos para um número arbitrário de caminhos, análise esta que será feita no Capítulo 3.

Capítulo 3

Performance do Método com o Aumento do Número de Caminhos

3.1 Introdução

O objetivo deste capítulo é estudar o impacto do aumento do número de caminhos no desempenho do método. A pergunta a ser respondida é: Aumentar o número de caminhos traz algum benefício para a aplicação?

Na seção 3.2 será visto um modelo analítico para o cálculo da probabilidade de perda em rajada para um número arbitrário de caminhos. Este modelo analítico considera que todos os caminhos a serem analisados são homogêneos, ou seja, possuem a mesma probabilidade de erro/sucesso no envio de um pacote pela rede. Serão apresentados importantes resultados referentes as probabilidades de ocorrência de rajadas de erro/sucesso de pacotes, considerando tanto a utilização de um único caminho como também de múltiplos caminhos. Além dessas medidas, será apresentada correlação de perda utilizando diversidade de caminhos. Na seção 3.3 será mostrada análise semelhante quando os diversos caminhos são heterogêneos, ou seja, os caminhos possuem características diferentes, como por exemplo, a probabilidade de ocorrência de erros. A seção 3.4 apresentará as conclusões do capítulo, enfatizando as contribuições em relação aos outros trabalhos.

3.2 Modelo de caminhos homogêneos

Seja S uma variável aleatória que conta o número de pacotes consecutivos corretamente recebidos entre duas perdas e L uma variável aleatória que indica o número de pacotes consecutivamente perdidos entre dois pacotes recebidos. Ambas as variáveis serão associadas aos esquemas de um caminho apenas ou de múltiplos caminhos pelo subscrito "$_{sp}$" para um caminho e "$_{mp}$" para múltiplos caminhos, de modo que $P\{L_{sp}(\tau) > 0\}$ significa, a probabilidade de haver uma rajada de perda de qualquer tamanho, quando a fonte usa apenas um caminho e envia pacotes em intervalos regulares τ e $P\{L_{mp}(\tau, N) > 0\}$ tem um parâmetro a mais N que é o número de caminhos utilizados. Nesta seção serão analisadas quatro medidas de interesse sobre o comportamento do canal: $P\{L > m \,|\, L > 0\}$, $P\{L > 0\}$, $P\{S > 0\}$, $P\{S > m \,|\, S > 0\}$. Conforme descrito no capítulo 2 o intervalo de envio de pacotes é constante e de tamanho τ.

O modelo do comportamento do erro no canal é o mesmo usado durante o capítulo 2 (usado também nos trabalhos [**?**, **?**]). Segue-se aqui também as mesmas premissas descritas na seção 2.5, ou seja, todos os canais são homogêneos, os pacotes são enviados em seqüência por cada um dos N caminhos e o canal se encontra em estado estacionário.

3.2.1 Probabilidade de Rajadas de Erro/Sucesso

Seguindo o mesmo raciocínio utilizado na seção 2.5.4, podemos deduzir as probabilidades $P\{L_{sp}(\tau) > 0\}$, $P\{L_{mp}(\tau, N) > 0\}$, $P\{S_{sp}(\tau) > 0\}$ e $P\{S_{mp}(\tau, N) > 0\}$, onde N é o total de caminhos quando utiliza-se a diversidade de caminhos . Para no máximo dois caminhos estas equações estão descritas na seção 2.5.4. Resta deduzir $P\{L_{mp}(\tau, N) > 0\}$ para um número arbitrário de caminhos N:

Lema 2. $P\{L_{mp}(\tau, N) > 0\} = \pi_{Bom}\pi_{Ruim}$.

Demonstração. $P\{L_{mp}(\tau, N) > 0\} = P\{$O último pacote chegou em um tempo qualquer t (estado 0) e o próximo não chegou no tempo $t + \tau$ (estado 1)$\} = P\{$Caminho k estar no estado 0 em um tempo qualquer t e o caminho $k + 1 \,\mathbf{mod} N$ no estado

1 no tempo $t + \tau\} = \pi_{\text{Bom}}\pi_{\text{Ruim}}$. A utilização do termo $k + 1\,\mathbf{mod}N$ decorre da disciplina de Round Robin em que os canais são selecionados para o envio de um pacote. . $\square$

Usando a mesma abordagem, pode-se determinar $P\{S_{mp}(\tau, N) > 0\} = \pi_{\text{Ruim}}\pi_{\text{Bom}}$.

Vale notar que $P\{L_{sp}(\tau) > 0\} < P\{L_{mp}(\tau, N) > 0\}$ e $P\{S_{sp}(\tau) > 0\} < P\{S_{mp}(\tau.N) > 0\}$ (basta notar que $\pi_{\text{Ruim}} < p_{\text{Bom,Ruim}}$), o que significa que diversidade caminhos tem um número maior de rajadas de sucesso ou perda do que se fosse usado apenas um caminho, contudo, em contrapartida, essas rajadas são menores, pois a probabilidade de perda/sucesso é a mesma em ambos os casos.

$P\{L_{sp}(\tau) = m\}$ já foi definido na seção 2.5.4, assim como $P\{L_{mp}(\tau, 2) = m\}$.

O primeiro objetivo deste capítulo é definir $P\{L_{mp}(\tau, N) = m\}$ para um número arbitrário de caminhos:

Teorema 3.

$$P\{L_{mp}(\tau, N) = m\} = \begin{cases} \pi_b \pi_r^m \pi_b & se\ m < N-1, \\ \pi_b \pi_r^{N-1} p_{b,b}(N\tau) & se\ m = N-1, \\ \pi_b \pi_r^{N-1} p_{b,r}(N\tau)\left[p_{r,r}(N\tau)\right]^{m-N} p_{r,b}(N\tau) & se\ m \geq N \end{cases},$$

onde os sub-escritos significam "Bom" (b) e "Ruim" (r).

Demonstração. A prova será dividida em três partes:

Para $m < N - 1$:

Evento: Envio do primeiro pacote em um tempo arbitrário t e este pacote não ser perdido. **Probabilidade do evento**: π_{Bom}.

Evento: Envio de m pacotes e todos estes serem perdidos nos tempos $t + \tau, t + 2\tau, \ldots, t + m\tau$. **Probabilidade do evento**: Como todos os caminhos são independentes, a probabilidade deste evento é a multiplicação da probabilidade dos eventos isolados, ou seja, π_{Ruim}^m.

Evento: Envio do $m + 1$-ésimo pacote e este não ser perdido. **Probabilidade do evento**: Novamente, pela independência dos caminhos a probabildade deste evento não depende dos outros pacotes: π_{Bom}.

Para $m = N - 1$:

Evento: Envio do primeiro pacote em um tempo arbitrário t e este pacote não ser perdido. **Probabilidade do evento**: π_{Bom}.

Evento: Envio de m pacotes e todos estes serem perdidos nos tempos $t + \tau, t +$

$2\tau, \ldots, t+m\tau$. **Probabilidade do evento**: Como todos os caminhos são independentes, a probabilidade deste evento é a multiplicação da probabilidade dos eventos isolados, ou seja, π^m_{Ruim}.
Evento: Envio do $m+1$-ésimo pacote e este não ser perdido. **Probabilidade do evento**: Diferente do evento anterior, a probabilidade de sucesso deste pacote depende do fato do primeiro pacote ter sido bem sucedido, pois os dois irão pelo mesmo caminho: $p_{\text{b,b}}(N\tau)$.

Para $m > N$:
Evento: Envio do primeiro pacote em um tempo arbitrário t e este pacote não ser perdido. **Probabilidade do evento**: π_{Bom}.
Evento: Envio de $N-1$ pacotes e todos estes serem perdidos nos tempos $t+\tau, t+2\tau, \ldots, t+m\tau$. **Probabilidade do evento**: Como todos os caminhos são independentes, a probabilidade deste evento é a multiplicação da probabilidade dos eventos isolados, ou seja, π^m_{Ruim}.
Evento: Envio do N-ésimo pacote e este também ser perdido. **Probabilidade do evento**: Dado que o último pacote enviado por este canal foi bem sucedido: $p_{\text{r,b}}(N\tau)$.
Evento: Envio dos restantes dos pacotes perdidos. **Probabilidade do evento**: Dado que o último pacote enviado por este canal foi perdido: $p_{\text{r,r}}(N\tau)^{m-N}$.
Evento: Envio do $m+1$-ésimo pacote e este não ser perdido. **Probabilidade do evento**: O último pacote enviado por este caminho foi perdido, portanto: $p_{\text{b,r}}(N\tau)$.
□

Compare a fórmula do teorema 3 com a fórmula dada em [?] (seção 3.2.1, $N=2$). O termo $N\tau$ refere-se ao tempo total de duração de um ciclo da disciplina Round Robin de entrega de pacotes aos N canais.

A fórmula da perda condicional para um número arbitrário (N) de caminhos é pelo lema 2:

$$P\{L_{mp}(\tau,N) = m \mid L_{mp}(\tau,N) > 0\} = \begin{cases} \pi_{\text{r}}^{m-1}\pi_{\text{b}} & \text{se m} < N-1, \\ \pi_{\text{r}}^{N-2}p_{\text{b,b}}(N\tau) & \text{se m} = N-1, \\ \pi_{\text{r}}^{N-2}p_{\text{b,r}}(N\tau)\left[p_{\text{r,r}}(N\tau)\right]^{m-N} p_{\text{r,b}}(N\tau) & \text{se m} \geq N \end{cases} . \tag{3.1}$$

Será demonstrado a seguir que a probabilidade de haver rajadas menores que $N-1$ pode ser bem maior usando diversidade de caminhos. É interessante notar nos gráficos da perda condicional o "salto" entre $P\{L_{mp}(\tau, N) = N - 2 \mid L_{mp}(\tau, N) > 0\}$ e $P\{L_{mp}(\tau, N) = N - 1 \mid L_{mp}(\tau, N) > 0\}$. Apesar dessa característica depender dos parâmetros da fonte e do canal, este é um claro alerta sobre o método ser contraproducente se há um número muito grande de caminhos pois não há uma queda significativa na probabilidade de haver uma rajada muito longa, como pode-se observar no gráfico da Figura 3.1, a probabilidade de haver uma rajada maior ou igual a 4 usando-se 5 caminhos é muito maior do que com apenas dois caminhos.

3.2.2 Resultados

A resposta à pergunta "mais caminhos melhoram o método?" depende dos requisitos da aplicação, ou seja, uma quantidade maior de caminhos não necessariamente melhora o método. Para eventos muito raros, como uma total indisponibilidade de um caminho, o aumento do número de caminhos claramente diminui a probabilidade de que todos os caminhos tenham este mesmo problema ao mesmo tempo. Contudo, uma maior quantidade de caminhos pode não interessar quando se pensa em diminuir a probabilidade de ocorrer uma rajada pequena de perda de pacotes.

O gráfico da Figura 3.1 retrata a rajada de perda condicional de um modelo cuja probabilidade de perda de um pacote é (≈ 0.46) muito alta porque seu objetivo é apenas mostrar a relação entre a probabilidade de perda, o aumento do número de caminhos e a probabilidade condicional de uma rajada de perda $< N - 1$. Nota-se que o incremento do número de caminhos de 2 para 3 aumenta drasticamente a probabilidade de perda de pelo menos dois pacotes em rajada.

Comparando os gráficos das figuras 3.1, 3.2 e 3.3[1], percebe-se que com a diminuição da probabilidade de perda e o aumento do tempo médio em que o caminho fica "Bom", a probabilidade de haver rajadas menores que o número de caminhos diminui drasticamente. Os gráficos também mostram que o ganho em adicionar mais caminhos decai com o número de caminhos e, portanto, a aplicação deve avaliar se é benéfica a adição de mais um caminho de acordo com as características de perda dos caminhos. Considere-se, por exemplo, uma aplicação de transmissão de

[1] As probabilidades dos gráficos não são constantes, apenas decrescem muito vagarosamente.

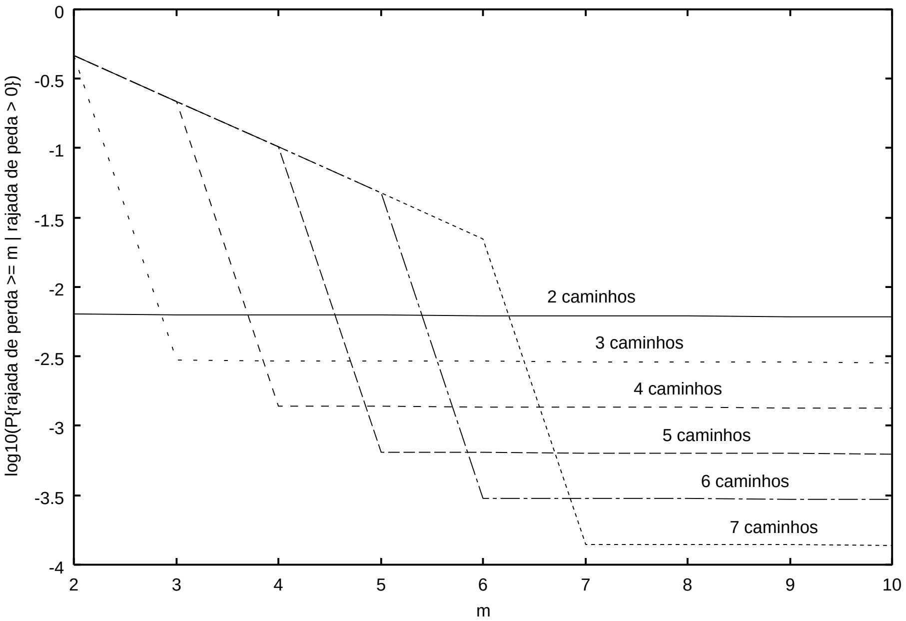

Figura 3.1: Gráfico de $\log_{10}(P\{L_{mp} \geq m \mid L_{mp} > 0\})$ com o número de caminhos N variando de 2 a 7, $\mu_{\text{Bom}} = 0.7$, $\mu_{\text{Ruim}} = 0.8$ e $\tau = 1/125$s ou 1.5Mbps se os pacotes têm tamanho 1500 bytes.

vídeo. Uma forma para se avaliar os ganhos de incrementar o número de caminhos é criando uma função de avaliação que leve em conta três fatores: a melhoria na qualidade do vídeo, utilizando-se por exemplo da função do ruído induzido por uma perda em rajada descrita em Liang et al. [?], a eficácia do esquema de correção de erros com a adição de mais um caminho e o custo de se adicionar mais um caminho. Como exemplo, o sistema de correção de erros descrito em [?] (onde a perda de um pacote só é corrigida se o próximo pacote chegar) atinge o seu melhor desempenho de correção de erros com dois caminhos para uma grande gama de parâmetros do modelo de Gilbert do canal. O gráfico 3.4 mostra $f(1/125, 2) - f(1/125, 5)$, onde $f(\tau, N) = \sum_{m=1}^{\infty} P[L_{mp}(\tau, N) = m \mid L_{mp}(\tau, N) > 0]/m$ é a fração de pacotes corrigido por este método, para $\mu_{\text{Bom}} \in (0, 3]$ e $\mu_{\text{Ruim}} \in (0, 8]$.

Esquemas de recuperação de erro (FEC) que necessitem $> N$ pacotes bem sucedidos para recuperar um perdido serão prejudicados se for usada a diversidade de caminhos, a não ser que os $\geq N$ pacotes sejam esperados pelo mesmo caminho, ou seja, que o algoritmo de redundância seja aplicado em cada caminho de forma independente dos outros. Este efeito tem relação direta com o fato de que

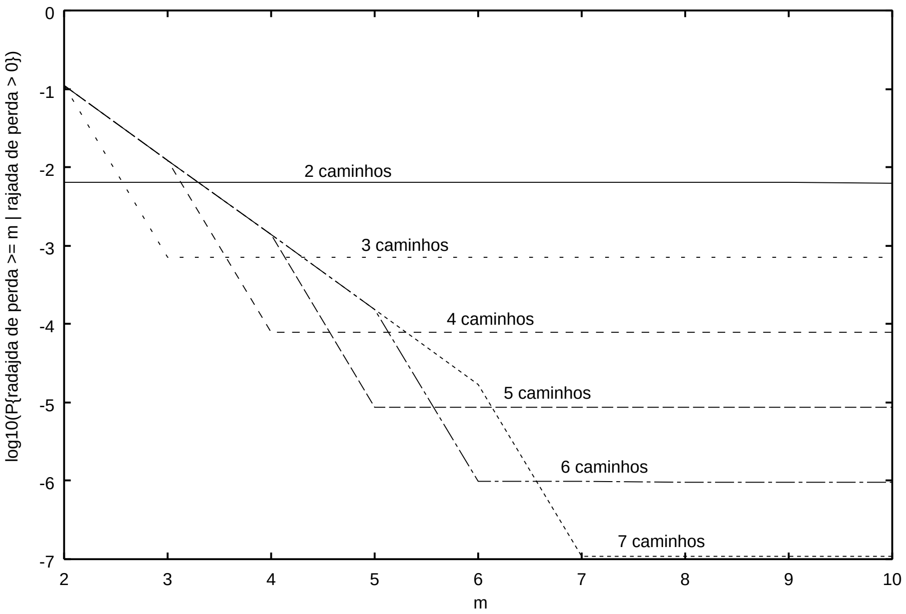

Figura 3.2: Gráfico de $\log_{10}(P\{L_{mp}(\tau, N) \geq m \,|\, L_{mp}(\tau, N) > 0\})$ com o número de caminhos N variando de 2 a 7, $\mu_{\text{Bom}} = 0.1$, $\mu_{\text{Ruim}} = 0.8$ e $\tau = 1/125$s ou 1.5Mbps se os pacotes têm tamanho igual a 1500 bytes.

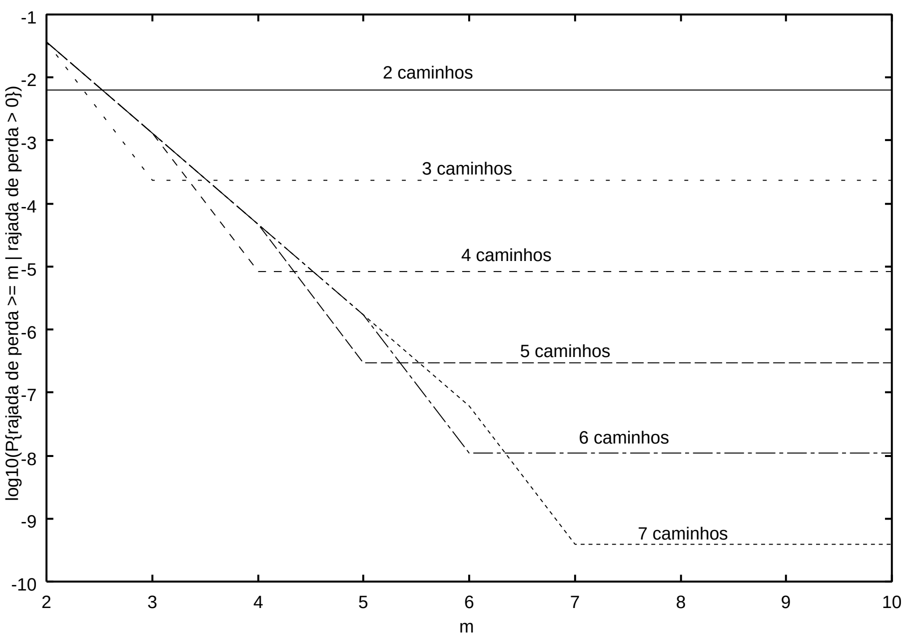

Figura 3.3: Gráfico de $\log_{10}(P\{L_{mp}(\tau, N) \geq m \,|\, L_{mp}(\tau, N) > 0\})$ com o número de caminhos N variando de 2 a 7, $\mu_{\text{Bom}} = 0.03$, $\mu_{\text{Ruim}} = 0.8$ e $\tau = 1/125$s ou 1.5Mbps se os pacotes têm tamanho igual a 1500 bytes.

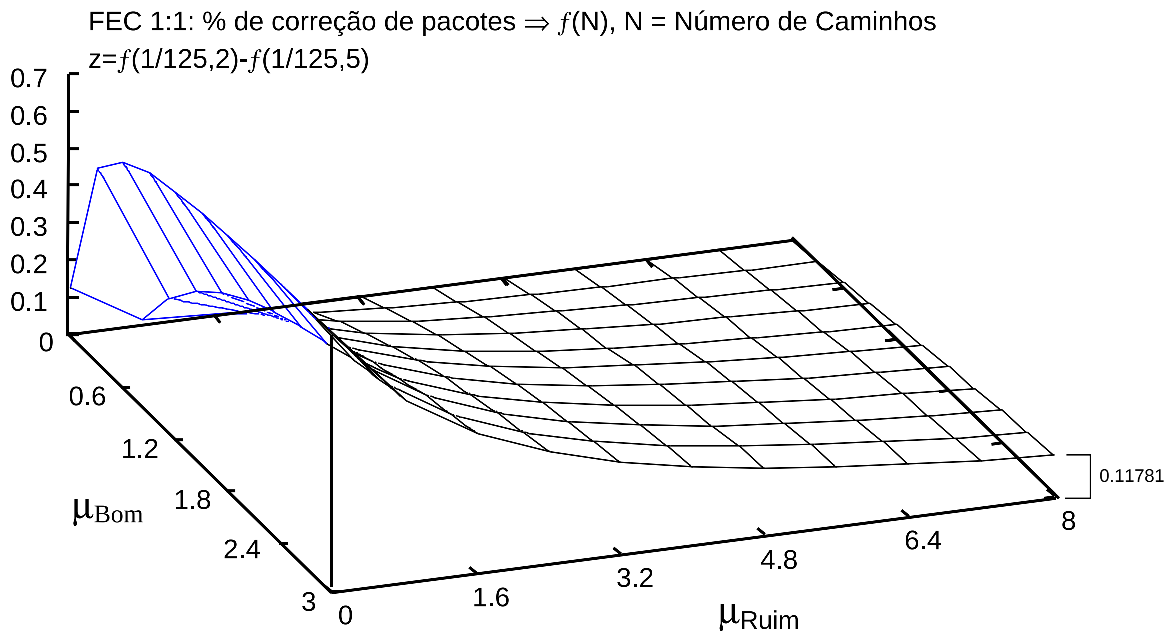

Figura 3.4: Diferença entre a fração de pacotes recuperados por FEC entre dois e cinco caminhos.

a probabilidade de haver uma rajada de erro ou sucesso, $P\{L > 0\}$e $P\{S > 0\}$ respectivamente, é maior usando diversidade de caminhos, o que leva a concluir que rajadas de erro ou sucesso são muito mais freqüentes, porém menores.

3.2.3 Correlação de perda com diversidade de caminhos

O objetivo desta seção é verificar se a autocorrelação da perda,

$$Cor(L_{mp}(t), L_{mp}(t+\delta)),$$

diminui com o aumento do número de caminhos utilizados. Aqui não será necessária a premissa de que os caminhos são homogêneos.

Dois casos devem ser considerados, baseados nos possíveis valores de δ:

- Para $\delta = kN,\ k \in \mathbb{N}$

$Cor(L_{mp}(t), L_{mp}(t+kN))$ é a autocorrelação de um dos N caminhos, então $Cor(L_{mp}(t), L_{mp}(t+$

$kN)) = Cor(X(t), X(t + kN))$ e usando [?]:

$$Cor(X(t), X(t + kN)) = \frac{(\mu_0 + \mu_1)p_{1,1}(kN) - \mu_0}{\mu_1}.$$

- Para $\delta \neq kN,\ k \in \mathbb{N}$.

Como se está enviando pacotes na disciplina *round robin* pelos diferentes caminhos e os caminhos são independentes, não há autocorrelação para $k \neq iN$, $i \in \mathbb{N}$. Por exemplo, se $N = 3$, obviamente a autocorrelação é zero para $k = 1, 2, 4, 5, \ldots$ mas não é zero para $k = 3, 6, 9, \ldots$.

A autocorrelação, apesar de ter sido utilizada em outros trabalhos na literatura ([?] entre outros), não é uma boa medida de desempenho do método. Observando $Cor(L_{mp}(t), L_{mp}(t+\delta))$ pode-se notar que esta é uma função decrescente em N, mas isso não implica diretamente que com utilização de mais caminhos as estatísticas de perda melhorem, como visto na seção 3.2.2. Conclusão: Reduzir a correlação de perdas à zero ($N \gg 1$) não significa necessariamente uma melhora nas estatísticas de perda.

3.3 Modelo com caminhos heterogêneos

O modelo proposto para estudar o método com caminhos heterogêneos é uma cadeia de Markov de tempo contínuo onde a perda de pacotes no caminho é modelada como um modelo de Gilbert e há um canal com características de perda diferente dos demais. Para analisar caminhos com diferentes modelos de erro foi utilizada a ferramenta Tangram-II [?]. Neste modelo cada caminho é representado por dois estados ("Bom" e "Ruim"), como na seção 3.2, a diferença fundamental é que os caminhos não são homogêneos e a distribuição do tempo entre envio de pacotes é exponencial.

Os parâmetros do modelo estão dispostos na tabela 3.1. A fonte é Poisson com taxa $\lambda = 20$ pacotes por unidade de tempo e, conforme anteriormente, a fonte escolhe os canais segundo a política Round-Robin.

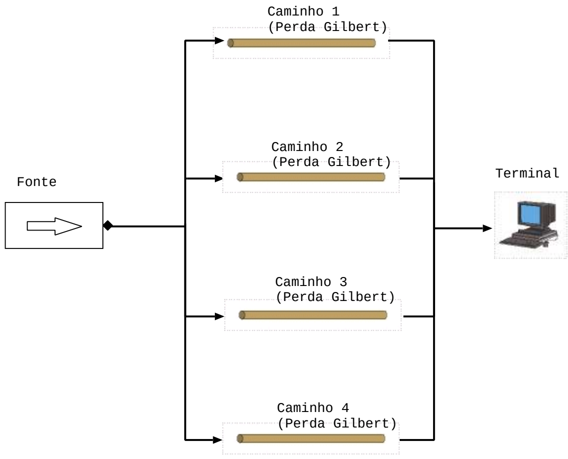

Figura 3.5: Modelo de diversidade de caminhos

Parâmetro	Exemplo Fig. 3.6	Exemplo Fig. 3.7
λ	20	20
$\mu^{(2,3,4)}_{Ruim,Bom}$	0.1	0.1
$\mu^{(2,3,4)}_{Bom,Ruim}$	0.9	0.9
$\mu^{(1)}_{Ruim,Bom}$	0.1	1
$\mu^{(1)}_{Bom,Ruim}$	0.1	10^{-5}

Tabela 3.1: Parâmetros do modelo

3.3.1 Resultados Analíticos

Neste seção será analisada a distribuição condicional da rajada de perda para identificar se, tendo um canal melhor, seria melhor enviar todo o tráfego por este canal. Serão analisados dois casos, o primeiro onde a diferença entre o melhor canal e os demais é muito grande e o segundo onde o canal melhor não é tão melhor que os demais canais. Para finalizar os estudos da utilização de caminhos heterogêneos, um outro objetivo é avaliar se a probabilidade condicional do tamanho das rajadas de perda ainda é a melhor estatística a ser comparada.

Para uma primeira análise, a diferença entre o canal "bom" e os "ruins" é muito grande exatamente para realçar as diferenças entre eles. Quanto menor a diferença entre o canal bom e os ruins, mas próximo dos resultados da seção 3.2.2 estarão os resultados onde foram considerados caminhos homogêneos.

Probabilidade	Exemplo Fig. 3.6	Exemplo Fig. 3.7
$P\{L_{sp\,1} > 0\}$	0.5	10^{-6}
$P\{L_{mp\,1,2} > 0\}$	≈ 0.7	≈ 0.45
$P\{L_{mp\,1,2,3} > 0\}$	≈ 0.766	≈ 0.6
$P\{L_{mp\,1,2,3,4} > 0\}$	≈ 0.8	≈ 0.675

Tabela 3.2: Probabilidade de uma rajada de perda de qualquer tamanho

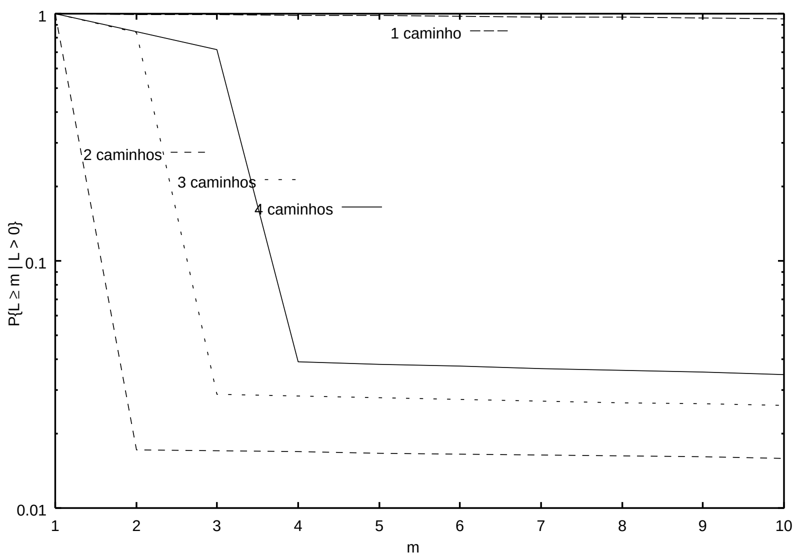

Figura 3.6: Distribuição condicional complementar de 1 à 4 caminhos, onde um deles tem sempre uma probabilidade de perda menor, mas com o mesmo tamanho médio do tamanho da rajada de perda.

Na figura 3.6 pode-se notar que a diversidade de caminhos realmente diminui muito a probabilidade de ter rajadas maior que o número de caminhos em relação a um só caminho, mesmo se os canais usados para dividir o tráfego têm probabilidades de perda maior que o melhor caminho. O trabalho de Golubchick et al. [?] também analisa o comportamento de dois caminhos heterogêneos, contudo a análise feita somente usa as probabilidades de perda condicional, o que não é uma medida apropriada (conforme será mostrado mais a frente) já que as probabilidades de perda de apenas um caminho e dos dois caminhos juntos não são as mesmas como eram em [?].

A Figura 3.7 mostra o mesmo comportamento da Figura 3.6, mas neste caso o caminho principal (o melhor) tem um probabilidade de perda muito pequena, com

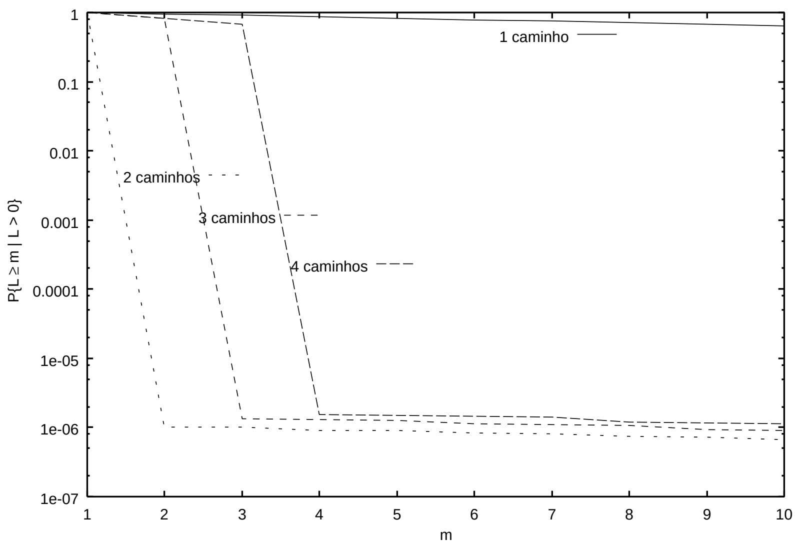

Figura 3.7: Distribuição condicional complementar de 1 à 4 caminhos, onde um deles tem sempre uma probabilidade e rajadas de perda menores (com um caminho usa-se o melhor caminho).

rajadas pequenas de perda. Os gráficos das Figuras 3.6 e 3.7 não representam a freqüência em que estas rajadas de perda acontecem. A Figura 3.8 mostra o gráfico 3.7 descondicionado em $P\{L > 0\}$, o que leva a concluir que, neste exemplo, no melhor caminho a freqüencia com que acontecem as rajadas grandes de erro é igual ao caso de múltiplos caminhos e a freqüencia das rajadas menores (menor que o número de caminhos) é muito menor. Selecionar entre um, dois ou mais caminhos depende exclusivamente do tipo de aplicação, do tipo de esquema de correção de erro adotado e das estatísticas do canal. Apesar de ao selecionar quatro caminhos no segundo caso (Figura 3.8) não serem obtidas estatísticas melhores que quando usado apenas um canal, dois canais ainda pode ser interessante para algum tipo de aplicação que não seja prejudicada com rajadas de perda pequenas. Balanceamento de tráfego pode também ser usado para atingir melhores estatísticas quando usado mais de um caminho.

Uma observação que vale nota nos gráficos 3.6, 3.7 e 3.8 é a vantagem de se usar menos caminhos nestes casos. Causa uma certa estranheza o fato de na distribuição de probabilidade condicional a curva “4 canais” estar sempre acima da curva “2 canais”, mas a verdade é que este fenômeno só se observa para valores de rajada

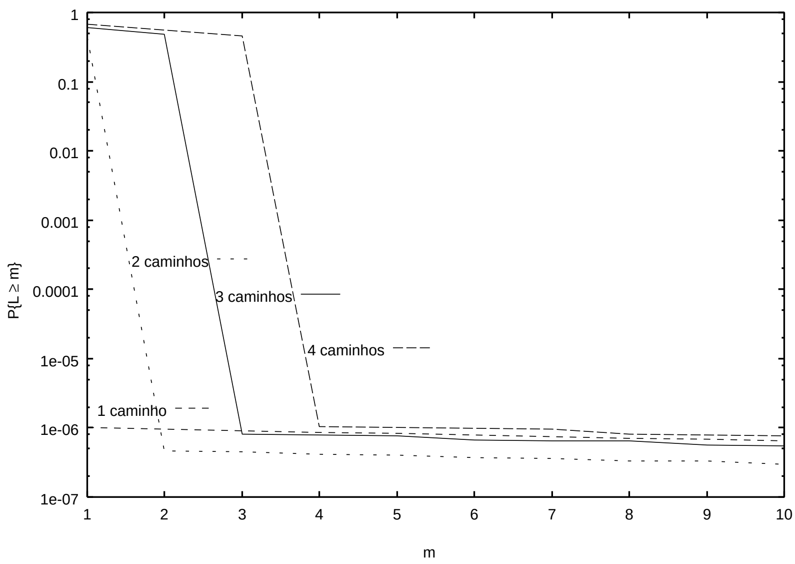

Figura 3.8: Distribuição complementar de 1 à 4 caminhos, onde um deles tem sempre uma probabilidade e rajadas de perda menores (com um caminho usa-se o melhor caminho).

menores que ≈ 200 no melhor caso, valores bastante altos. Como a probabilidade de ocorrer uma rajada maior que 200 é muito pequena, então melhorar o método para rajadas maiores que 200 não é uma abordagem que cause impacto na qualidade da transmissão.

3.4 Conclusões do capítulo

O aumento do número de caminhos melhora o método? Não há uma resposta única a esta questão. A seção 3.2.2 analisa caminhos homogêneos com pacotes enviados a intervalos regulares e conclui que o total de caminhos que melhora o método de diversidades de caminhos depende dos requisitos de qualidade da aplicação, assim como a seção 3.3.1. A seção 3.3.1 analisa caminhos desbalanceados e mostra que a distribuição condicional da rajada de perda pode não ser a melhor estatística para a comparação entre os caminhos, concluindo também que enviar todo o tráfego pelo melhor caminho é uma boa alternativa se o melhor canal não possui rajadas de erro muito longas.

Capítulo 4

Estudo da Escalabilidade do Método

4.1 Introdução

NESTE capítulo será analisada a escalabilidade do método de diversidade de caminhos, ou seja, como este se comportaria se fosse largamente utilizado por usuários na Internet ou numa rede privada. A análise da escalabilidade é fundamental para qualquer proposta, porém ainda não foi pesquisada pela literatura e será analisada no decorrer deste capítulo.

Consideraremos para análise os seguintes casos:

1. Todas as aplicações multimídia utilizam diversidade de caminhos.
2. Todas as aplicações multimídia, exceto uma, utilizam diversidade de caminhos . A única que não utiliza diversidade de caminhos usa apenas um caminho.
3. Nenhuma das aplicações multimídia utilizam diversidade de caminhos.
4. Apenas uma das aplicações multimídia utiliza diversidade de caminhos.

O estudo da diversidade de caminhos apresentado no capítulo 3, supôs independência nas perdas entre os caminhos, o que implicitamente é afirmar que existe independência entre os tráfegos que percorrem os diversos caminhos. Entretanto, como a diversidade de caminhos divide o tráfego pelos diferentes caminhos está independência não é alcançada. Na literatura, caminhos disjuntos não possuem dependência de perda porque assume-se que o tráfego enviado por somente uma fonte não é o suficiente para criar uma grande dependência de perda entre os diversos caminhos.

Entretanto, o cenário descrito anteriormente muda quando o tráfego que utiliza diversidade de caminhos começa a ser significativo em relação aos demais tráfegos que compartilham os diversos canais da rede. O objetivo principal deste capítulo é explorar diversos cenários e mostrar a existência de dependência entre perdas, bem como a suavização do tráfego quando diversas fontes são consideradas.

Este capítulo tem a seguinte estrutura. A seção 4.2 apresenta o modelo que será utilizado para responder as questões de dependência de perda e da suavização do tráfego que compartilha diversos caminhos. Na sub-seção 4.3.1 será estudado o impacto de uma porção significativa do tráfego de entrada nos roteadores, vindo das mesmas fontes, na correlação das perdas nas filas destes roteadores. Ao mesmo tempo que a diversidade de caminhos cria uma correlação de perdas, o tráfego conjunto de todas as fontes é suavizado, diminuindo o tamanho das rajadas de perda. Esta afirmativa será mostrada na sub-seção 4.3.2. A combinação dos dois fatores: dependência e suavização será abordada na seção 3.2.2, onde serão apresentadas também as conclusões.

4.2 Modelo

O modelo para estudar a escalabilidade do método pode ser visualizado na figura 4.1. O tráfego multimídia e representado por fontes On-Off, que somadas ao tráfego de fundo ("Tráfego de fundo") e o tráfego da fonte multimídia que se quer analisar ("Font"→dest") formam o tráfego de entrada nos roteadores "Roteador 1" e "Roteador 2". A medida de interesse deste sistema é a distribuição das rajadas de perda entre a fonte multimídia "Font" e o seu destino "dest". Os roteadores neste modelo representam roteadores da "borda da rede" (*border routers*). Para as fontes multimídia (exceto a fonte "src"), a distribuição tempo entre pacotes é exponencial, juntamente com a distribuição dos períodos ativo ("On") e inativo ("Off"). A distribuição do tempo de serviço dos pacotes também é exponencial. A única dis-

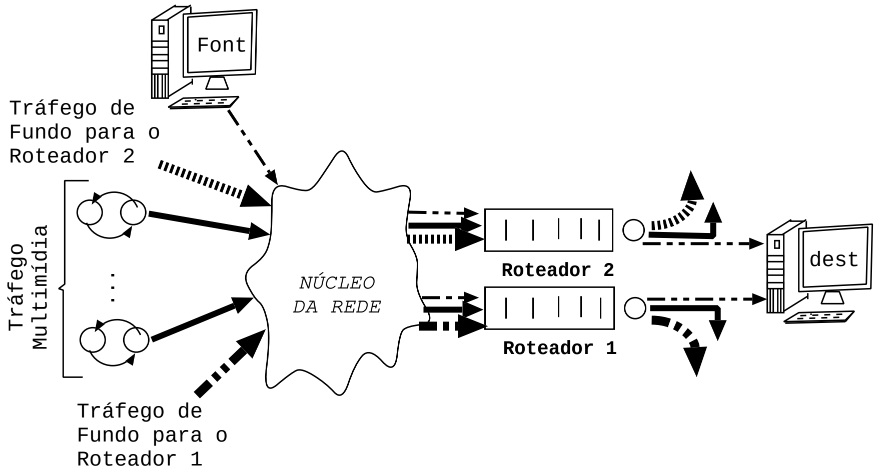

Figura 4.1: Modelo para análise da escalabilidade do método.

tribuição não exponencial é a distribuição do tempo entre pacotes de "Font", que é determinística e constante.

4.3 Propriedades da diversidade de caminhos

O objetivo desta seção é avaliar quais são as propriedades do tráfego de fontes multimídia que utilizam múltiplos caminhos (diversidade de caminhos) e suas diferenças em relação ao tráfego de fontes que fazem uso de apenas um caminho. Para avaliar estas propriedades serão feitas algumas simplificações no modelo da Figura 4.1, indicadas na descrição do modelo de cada seção.

4.3.1 Diversidade de caminhos e a dependência da perda em caminhos completamente disjuntos

A primeira avaliação remonta à assertiva dos capítulos 2 e 3 sobre a independência entre perdas de caminhos que não compartilhem os mesmos canais. Esta assertiva somente é verdadeira se o tráfego que passa pelos dois caminhos possui correlação nula. No entanto, será mostrado que a diversidade de caminhos cria uma dependência de perdas entre caminhos completamente disjuntos (caminhos que não compartilham nenhum roteador). Para tal, a medida de interesse neste caso é a correlação entre perdas nos roteadores "Roteador 1" e "Roteador 2". Esta medida foi escolhida porque indica o grau de dependência entre as duas filas.

Para as análises das seções 4.3.1.1 e 4.3.1.2 será utilizada a ferramenta Tangram-II [?] com modelos de simulação e analíticos respectivamente.

4.3.1.1 Autocorrelação nas filas sem atraso na rede

O modelo da Figura 4.1 considerado neste item contém apenas duas fontes multimídia, sem tráfego de fundo, onde o tráfego enviado por "Font" (20 em 20 milisegundos (constante)) é muito menor que o enviado pelas outras duas fontes On/Off multimídia (1 em 1 milisegundo (em média)) . Outra simplificação é a suposição de que a diferença de tempo entre o envio de um pacote e a sua recepção, causado pela rede ("Núcleo da Rede"), é sempre constante e igual para ambos os caminhos. Na seção 4.3.1.2 será vista a sensibilidade relativa à variação deste atraso.

A medida obtida foi a autocorrelação da perda entre os pacotes enviados por "Font". Se analisada a autocorrelação dos pacotes com intervalos múltiplos de um número par por τ (intervalo entre envio de pacotes), esta medida é também a média das autocorrelações de fila cheia dos roteadores "Ro nteador 1" e "Roteador 2". Considere 4 pacotes, numerados de 1 a 4, como são 2 roteadores, os pacotes 1 e 3 são enviados ao mesmo roteador, assim como os pacotes 2 e 4, portanto a autocorrelação de perda destes pacotes é a média da autocorrelação de perda dos roteadores. No caso dos pacotes 1 e 2 (que são enviados para roteadores diferentes), esta medida é a correlação de perdas entre os dois roteadores. Pode-se mostrar esta propriedade pelo seguinte lema:

Lema 4. *Seja $X(t)$ o processo estocástico que vale zero caso o pacote enviado pela fonte "Font" no tempo t tenha sido bem sucedido e um caso contrário e $Y_i(t)$ o mesmo processo estocástico só que neste caso os pacotes só serão enviados ao roteador i. A autocorrelação de perda $Cor(X(t), X(t + 2k\tau))$, $k \in \mathbf{N}$ é a média da correlação das filas dos roteadores, ou seja, $Cor(Y_1(0), Y_1(2k\tau))/2 + Cor(Y_2(0), Y_2(2k\tau))/2$ se o processo de perda for estacionário.*

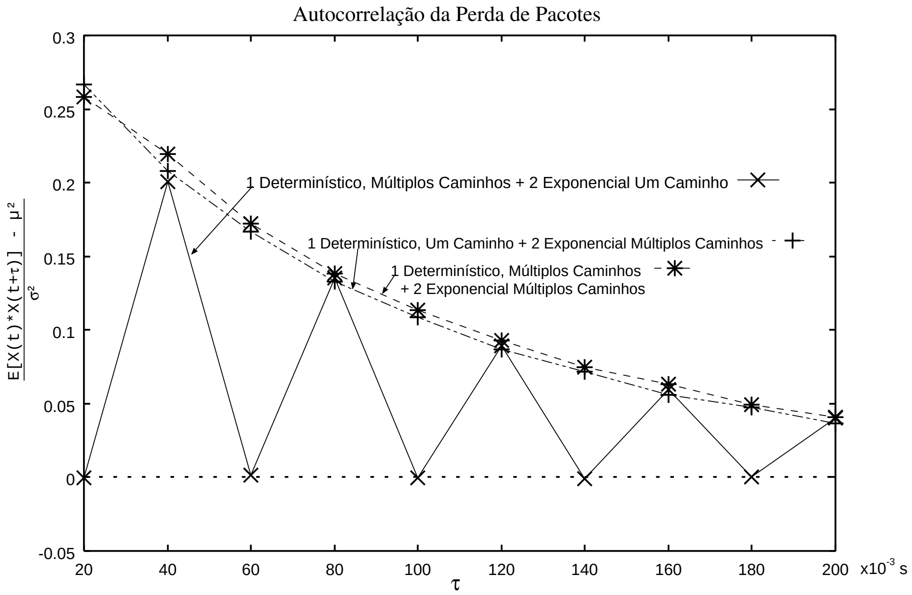

Figura 4.2: Autocorrelação da perda de pacotes da fonte multimídia determinística.

Demonstração. Pela definição de processo estacionário

$$Cor(Y(t), Y(t+2k\tau)) = Cor(Y(0), Y(2k\tau)).$$

Pode-se pensar ao invés de tempo, em pacotes, ou seja, no tempo $t + 2k\tau$ é enviado o $(t/\tau + 2k)$-ésimo pacote e nos casos onde a divisão não é inteira, a autocorrelação é zero. Se t/τ é ímpar, então todos os pacotes $(t/\tau + 2k)$ serão enviados ao primeiro roteador, $Cor(X(t), X(t+2k\tau) \mid t/\tau \text{ ímpar})$ que é $Cor(Y_1(0), Y_1(2k\tau))$, uma vez que a perda é estacionária e o mesmo vale para t/τ par: $Cor(Y_2(0), Y_2(2k\tau))$. Então

$$Cor(X(t), X(t+2k\tau)) = Cor(Y_1(0), Y_1(2k\tau))P[t/\tau \text{ ímpar}] + Cor(Y_2(0), Y_2(2k\tau))P[t/\tau \text{ par}],$$

que é a média da autocorrelação nas filas dos roteadores. □

A figura 4.2 mostra o gráfico da autocorrelação de perda em três casos a saber (para nestes modelos foram usados os parâmetros da tabela 4.1):

- Cada uma das duas On/Offs de tráfego multimídia envia pacotes por um dos caminhos "2 Exponencial Um Caminho" e a fonte multimídia analisada envia tráfego pelos dois caminhos.
- As duas On/Offs de tráfego multimídia enviam pacotes por ambos os caminhos e a fonte multimídia analisada envia tráfego por apenas um caminho e

Tipo	TX Serviço/Transmissão	On⇒Off	Off⇒On	Máximo da Fila
Fonte Multimídia On/Off 1	12	0.1	0.1	-
Fonte Multimídia Determinística	1	-	-	-
Fonte Multimídia On/Off 2	12	0.1	0.1	
Router 1	7	-	-	10
Router 2	7	-	-	10

Tabela 4.1: Parâmetros do modelo

- Todas as fontes de tráfego enviam pacotes por ambos os caminhos.

Como podemos observar pelo gráfico da Figura4.2, quando a maior parte do tráfego usa diversidade de caminhos há correlação entre as filas dos roteadores, o que é ruim pois quanto maior a correlação entre as filas, menor o ganho do uso da diversidade de caminhos. No caso extremo, onde a razão entre a correlação entre as filas e a autocorrelação das filas é um, os dois caminhos distintos se comportam como se fossem um só.

4.3.1.2 Sensibilidade em relação a atrasos aleatórios na rede

No caso estudado da sub-seção 4.3.1.1 obtém-se uma autocorrelação maior que se poderia obter em uma rede real pois quase todo o tráfego de entrada nos roteadores é fortemente correlacionado e não há variação do atraso nos dois caminhos. Esta correlação pode diminuir à medida que sejam considerados os atrasos aleatórios na rede, exatamente o foco de estudo desta sub-seção. O modelo e os parâmetros são os mesmo da seção 4.3.1.1 com uma única diferença: no caminho entre a "Fonte Multimídia On/Off 2" e o "Routeador 2" existe um atraso aleatório de distribuição Normal ou Exponencial. Na distribuição Normal, a variância é a metade da média. Os parâmetros foram escolhidos com o intuito de avaliar o impacto de grandes diferenças de atrasos entre os caminhos "Fonte Multimídia On/Off 1 → Routeador 1" e "Fonte Multimídia On/Off 2 → Routeador 2". A fonte de tráfego que está sendo medida não experimenta nenhum tipo de atraso, diminuindo ainda mais a correlação de perdas.

A Figura 4.3 mostra o mesmo modelo da seção 4.3.1.1 só que desta vez com

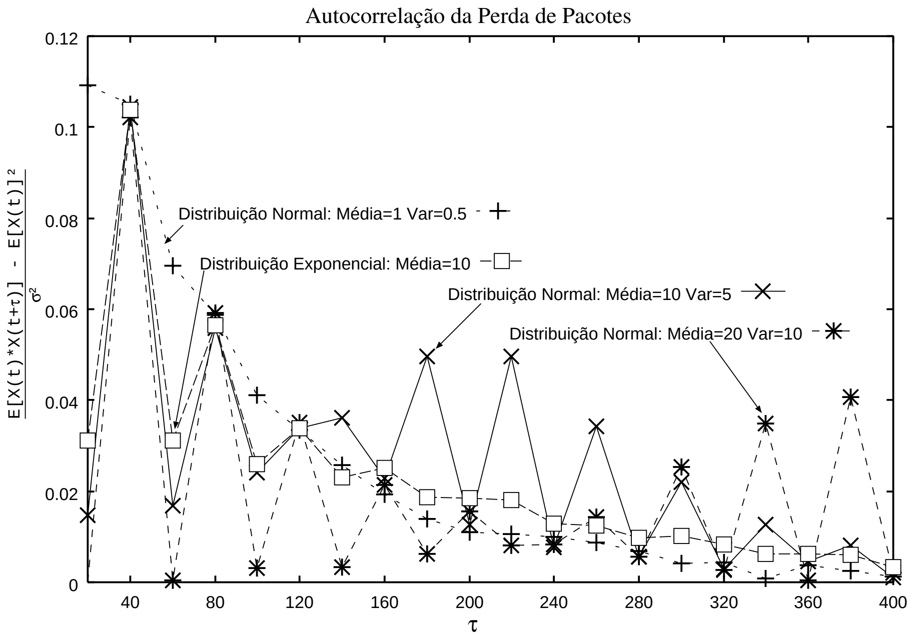

Figura 4.3: Autocorrelação entre da perda de pacotes da fonte multimídia determinística com atrasos aleatórios.

atrasos aleatórios. Por esta figura percebe-se que quanto maior o atraso inserido em relação ao tempo que a fonte fica ativa (período em "On") menor é a correlação entre as filas dos roteadores 1 e 2.

4.3.2 Suavização do tráfego

Ao mesmo tempo que a diversidade de caminhos cria uma dependência entre as perdas nas filas dos roteadores, ela também suaviza o tráfego das fontes o que implica diretamente em rajadas de perda menores nas filas dos roteadores.

Assumindo que as fontes On/Offs de tráfego multimídia transmitam tráfego à mesma taxa e que esta seja muito maior que a taxa de transição de "On" para "Off", pode-se usar a técnica de fluidos para resolver o modelo. Neste modelo existem S fontes On/Off de tráfego multimídia e a correlação nas duas filas é um, mas o nosso interesse é só na suavização causada pelo método de diversidade de caminhos.

Trabalhos anteriores sobre dispersão de tráfego (*traffic dispersion*), veja [?] para uma coletânea dos trabalhos, mostraram que dividindo o tráfego por múltiplos ca-

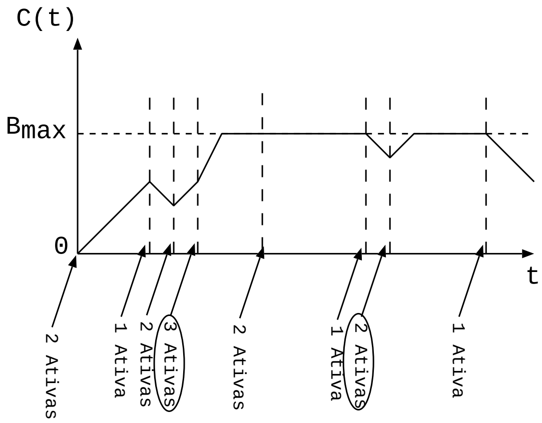

Figura 4.4: Exemplo do gráfico de $C(t)$ versus t. "i Ativas" significa que i fontes estão ativas.

minhos diminui a razão entre a taxa de transmissão de pico e a taxa média de transmissão, o que afeta diretamente o tamanho das rajadas de perda. Estes trabalhos entretanto não calculam em quanto é diminuída a rajada de perda.

Seja $X(t)$ o processo estocástico que representa o número de fontes On/Off ativas e $\mathbf{P}$ a matriz de probabilidade de transição associada ao processo discreto equivalente a $X(t)$, onde o tempo entre transições é exponencialmente distribuído com taxa Λ (ver [?] para maiores detalhes). Seja também r_i a recompensa associada ao estado i, que representa a taxa com que o reservatório do fluido (fila do roteador) está enchendo/esvaziando e $\mathbf{e}$ um vetor coluna onde todos os elementos são iguais a 1. A quantidade de fluido no reservatório no tempo t é dado por $C(t)$, que varia no intervalo $[0, B_{max}]$ (mais detalhes em [?]), portanto $C(t) \geq B_{max}$ significa que o reservatório está cheio no tempo t. O objetivo é achar a distribuição de T_{full}, o tempo que o reservatório permanece cheio. Para tal, primeiramente, será calculada a probabilidade de ter enchido o reservatório no momento em que i fontes estão ativas, com esta probabilidade é possível determinar a distribuição do tempo em que o reservatório ficará cheio. Os estados circulados por elipses ("3 Ativas" e "2 Ativas") na figura 4.4 são os estados de entrada no reservatório cheio.

Como o modelo é Markoviano (sem memória), uma vez encontrado o estado de entrada no reservatório cheio, o tempo que o reservatório continua cheio é igual ao tempo que, a partir do estado de entrada, o sistema fica com i fontes ativas dado que $r_i \geq 0$, ou seja, o reservatório está sempre sendo enchido ou mantido constante. Seja S_2 o conjunto de estados i onde $r_i \geq 0$ e S_1 o conjunto de estados i onde $r_i < 0$. $\mathbf{w}^*$ é um vetor linha em que o seu j-ésimo componente é igual à probabilidade de o estado de entrada ser j, $j \in S_2$. Assim, $\mathbf{w}^*$ é o vetor com as probabilidades dos estados no momento em que o reservatório enche.

Se $\mathbf{P}_{22}$ é a matriz sub-estocástica de probabilidades de transição entre os estados de S_2 e $\mathbf{P}_{11}$ a matriz sub-estocástica de probabilidades de transição dos estados de S_1, pode-se dividir a matriz de transição de probabilidade $\mathbf{P}$ da fonte conforme abaixo:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix},$$

e $\mathbf{P}_{12}$ a a matriz de probabilidade de transição entre os estados que esvaziam o reservatório e os que enchem ou mantém constante o reservatório.

O componente m, n da matriz $\mathbf{I}+\mathbf{P}_{22}+\mathbf{P}_{22}^2+\mathbf{P}_{22}^3+\ldots = (\mathbf{I}-\mathbf{P}_{22})^{-1}$ é o número médio de visitas aos estados que enchem o reservatório dado que se começou esta jornada pelo estado $m \in S_2$. No caso específico das fontes On/Off, só se pode sair para os estados $\mathbf{P}_{11}$ pelo estado que possui o menor número de fontes ativa entre os estados de S_2.

A probabilidade condicional de haver i fontes ativas dado que o reservatório está cheio ($\lim_{t\to\infty} P[X(t) = i \,|\, C(t) \geq B_{max}]$) poderia ser encontrada usando $\mathbf{w}^*$ e $\mathbf{P}_{22}$ da seguinte forma:

$$\mathbf{w}^* (\mathbf{I} - \mathbf{P}_{22})^{-1} = \mathbf{v}_{S_2}, \tag{4.1}$$

$\mathbf{v}_{S_2}^* = \mathbf{v}_{S_2}/(\mathbf{v}_{S_2}\mathbf{e})$ é um vetor cujo k-ésimo componente é igual a $\lim_{t\to\infty} P[X(t) = j \,|\, C(t) \geq B_{max}]$, $j \in S_2$.

O objetivo agora é achar $\mathbf{w}^*$ sabendo-se $\mathbf{v}_{S_2}$. A complexidade para calcular $\mathbf{v}_{S_2}$ é a complexidade de resolver um sistema linear de ordem $N+1$ se todas as fontes On-Off são homogêneas, veja Pallares-Ségarra e García-Haro [**?**] e suas referências para achar métodos aproximados ou mais precisos de calcular $\mathbf{v}_{S_2}$ para fontes On/Off homogêneas ou heterogêneas. A partir de $\mathbf{v}_{S_2}$, pela equação 4.1, pode-se achar $\mathbf{w}^*$ com $\mathbf{w} = \mathbf{v}_{S_2} - \mathbf{v}_{S_2}\mathbf{P}_{22}$ e $\mathbf{w}^* = \mathbf{w}/(\mathbf{w}\,\mathbf{e})$. Com $\mathbf{w}^*$ pode-se achar a média do tamanho da rajada de perda através do número médio de visitas, já que $[\mathbf{w}^*(\mathbf{I}-\mathbf{P}_{22})^{-1}]\,\mathbf{e}$ é o número médio de visitas aos os estados de S_2 após o reservatório ter enchido. Cada transição do modelo dura em média $1/\Lambda$, portanto

$$E[T_{full}] = \frac{[\mathbf{w}^*(\mathbf{I}-\mathbf{P}_{22})^{-1}]\,\mathbf{e}}{\Lambda}.$$

A equação $\mathbf{w}\,(\mathbf{I}-\mathbf{P}_{22})^{-1} = \mathbf{v}_{S_2}$ tem um vetor solução $\mathbf{w}$ não nulo se $\mathbf{P}_{22}$ for sub-estocástica. De fato, uma vez obtido o vetor $\mathbf{w}^*$ da probabilidade condicional dos estados em que primeiro enchem o reservatório, a distribuição deste tempo pode ser encontrada usando diversos métodos da literatura, em [?] pode-se encontrar alguns destes métodos para matrizes com estruturas genéricas.

O cenário de exemplo tem as fontes On/Off com a média do tempo de ir de "Off" para "On" igual a 10^4 segundos, mesma média do tempo de "On" para "Off". A capacidade do canal é 16.5Mbps e cada fonte transmite tráfego à 2Mbps, existem 30 fontes no total. Nota-se que a carga no sistema é muito alta, $\approx 91\%$, exatamente para medir o comportamento do sistema em um caso extremo.

Através do método descrito acima, é possível obter $E[T_{full}] = 1544.8$ caso as fontes façam **diversidade de caminhos** e $E[T_{full}] = 2209.0$ se as fontes usam **apenas um caminho**. A figura 4.5 mostra a distribuição do tempo em que o reservatório permanece cheio, o que será também, dividindo o eixo do tempo do gráfico pelo tempo entre envio de pacotes, o gráfico da distribuição do tamanho da rajada de perda (rajadas que seriam muito grandes pois o sistema está com carga altíssima).

4.4 Simulações com o TCP como tráfego de fundo

Caminhos disjuntos são caminhos que não compartilham os mesmos canais, ou se os compartilham, estes não são os pontos de gargalo na rede, de forma que a perda de pacote nestes caminho sejam independentes entre si.

Durante esta seção serão analisados cenários onde um tráfego multimídia toma metade da capacidade do canal e o método de diversidade de caminhos é largamente

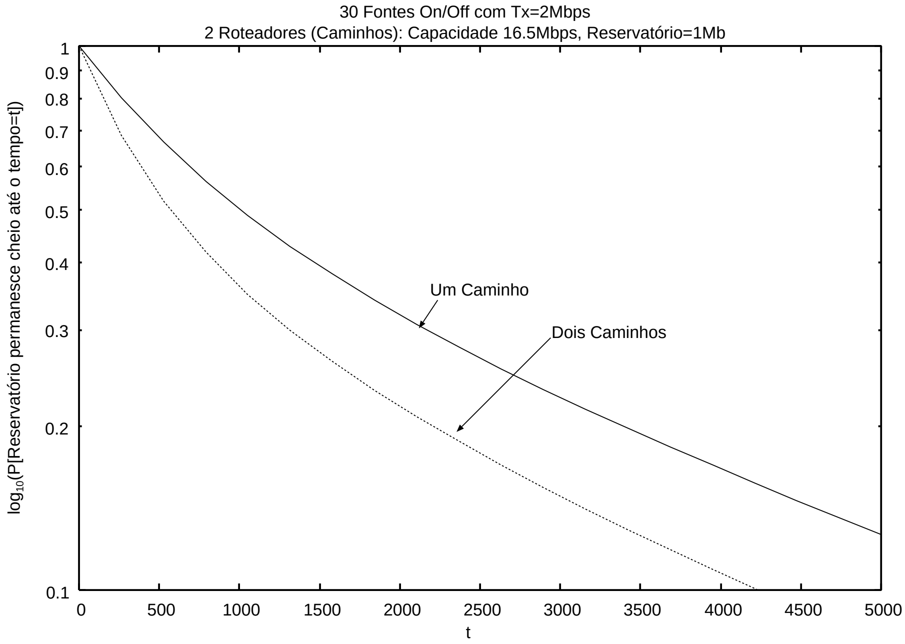

Figura 4.5: Distribuição do tempo em que o reservatório permanece cheio.

utilizado pelas aplicações multimídia. O objetivo principal é analisar como os comportamentos descritos nas sub-seções 4.3.1.1 e 4.3.2 interagem entre si e como é o comportamento das rajadas de perda e a sua relação com o que já foi visto neste capítulo.

Ferramentas utilizadas nesta seção: No primeiro exemplo (modelo de fluido) foi utilizada a ferramenta Tangram-II [?] onde foi feito um modelo do TCP simplificado e no segundo exemplo (modelo de pacotes) foi utilizada a ferramenta *ns* [?] pois esta última possui um modelo de TCP mais completo.

Simulação usando fluidos com o TCP sem controle de fluxo

A primeira análise da relação entre o aumento da correlação de perda e a suavização do tráfego nas rajadas de perda será feita usando-se o modelo completo da figura 4.1, onde o "Background Traffic" são fontes que simplificam o comportamento de fontes TCP (Reno, detalhes dos tipos de TCP podem ser encontrados em [?]), pois não simulam o controle de fluxo do TCP. Somente é simulado o controle de congestionamento (*fast start* e *fast recovery*) conforme indicado abaixo. Sem o controle de fluxo, o comportamento das fontes TCP é independente do tamanho do reservatório (o que não é verdade no caso real) até que haja uma perda na mesma. Esta independência do tamanho do reservatório por parte da inexistência do controle de congestionamento ameniza a dependência entre os reservatórios ocasionada pelas fontes de tráfego multimídia. Entretanto, conforme será visto a seguir, ainda assim este grau de independência do tráfego de fundo não é suficiente para eliminar a forte dependência entre as perdas nos dois reservatórios. Todas as distribuições do modelo são exponenciais, exceto quando explicitamente dito o contrário.

Duas simulações foram feitas, uma com 100 e outra com 250 fontes TCP para cada roteador. Ambos os modelos possuem 100 fontes On/Off representando o tráfego multimídia para ambos os roteadores. O tempo médio no estado ativo das fontes On/Off é de 2000 segundos e o tempo médio inativo é de 200 segundos e a sua taxa de transmissão quando ativa é de 1.2Mbps. Cada conexão TCP é simulada com quatro estágios: não enviando tráfego e enviando tráfego à taxas 0.002, 0.01 e 0.05 Mbps. O controle de congestionamento entra em ação quando o reservatório passa de dois limites estabelecidos (perto do reservatório cheio). O primeiro limite (o mais baixo) a probabilidade de perda é menor que no segundo. A probabilidade de perda é na verdade dividida entre a probabilidade de decrementar a sua janela pela metade e a probabilidade de ir para o estado "não enviando". Uma destas ações somente acontece depois de um intervalo de tempo exponencialmente distribuído com média 1/RTT. Para resolver o modelo foi usada a aproximação da taxa de transmissão de pacotes por um fluido contínuo, como já realizado na seção 4.3.2.

A figura 4.6 mostra a probabilidade condicional de perder m pacotes consecutivos ($P\{L_{sp}(\tau) = m \,|\, L_{sp}(\tau) > 0\}$ e $P\{L_{mp}(\tau, N) = m \,|\, L_{mp}(\tau, N) > 0\}$) se as fontes multimídia (On/Offs) utilizam de apenas um caminho. A correlação de perda entre os reservatórios dos dois roteadores é representada pela diferença entre todas as fonte fazendo diversidade de caminhos ("Dois caminhos sem bottlenecks independentes" do gráfico da Figura 4.6) e somente uma fonte fazendo diversidade de caminhos ("Dois caminhos com bottlenecks independentes" do gráfico da Figura 4.6). Apesar de haver esta correlação entre perdas nos reservatórios dos roteadores, conforme já apontado pela sub-seção 4.3.1.1, o gráfico da figura 4.6 mostra que ainda há

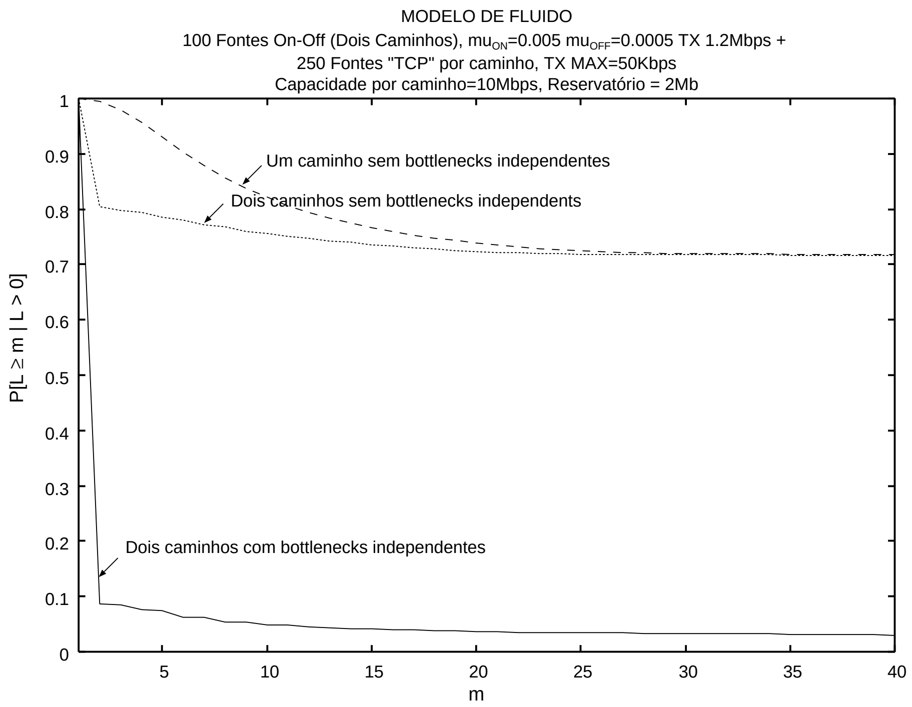

Figura 4.6: 250 fontes TCP para cada roteador, com 100 fontes On/Off usando diversidade de caminhos

vantagens em se usar diversidade de caminhos, principalmente pela suavização do tráfego causado pela diversidade de caminhos, conforme mostrado nos exemplos da sub-seção 4.3.2.

O gráfico da figura 4.7 mostra o que acontece quando se diminui o tráfego de fundo. Percebe-se que há uma sensível diminuição na probabilidade de haver uma grande rajada de perda quando as fontes todas utilizam diversidade de caminhos em relação ao caso de apenas um caminho. Isso se deve à suavização de tráfego que a diversidade de caminhos causa: Com menos tráfego de fundo o tráfego que chega ao roteador é mais "suave". Se no TCP modelado houvesse controle de fluxo, é possível que o tamanho das rajadas de perda diminuísse sensivelmente, conforme acontece no próximo exemplo, onde o TCP foi simulado por pacote e com todos os seus mecanismos implementados em cada detalhe.

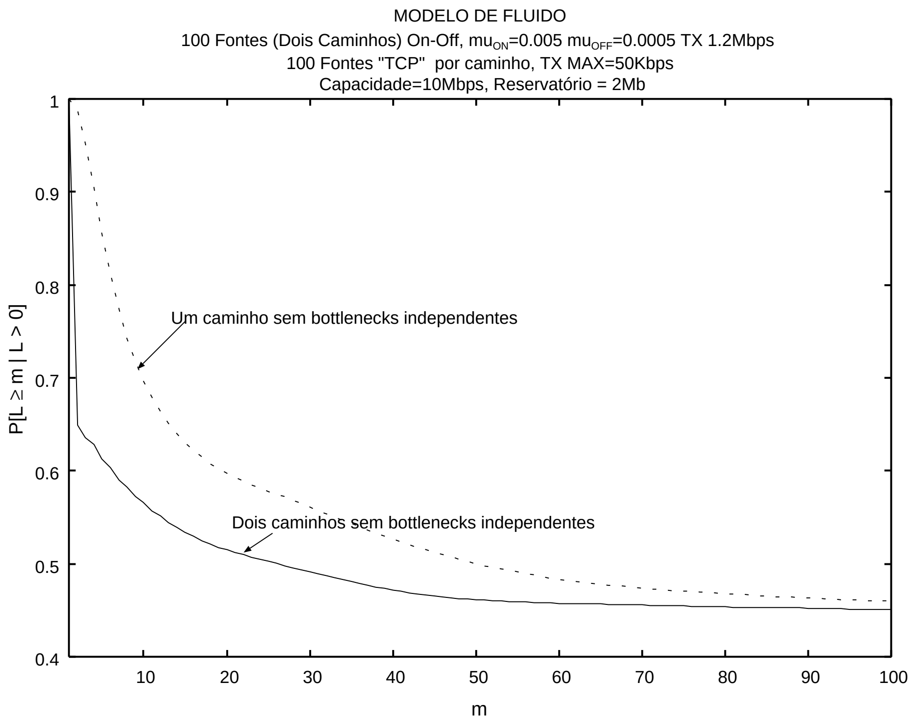

Figura 4.7: 100 fontes TCP para cada roteador, com 100 fontes On/Off usando diversidade de caminhos

Simulação por pacotes com o TCP completo

Na figura 4.8 está descrito o modelo simulado onde todos os canais são *full-duplex*. A distribuição do tempo em que as fontes multimídia ficam ativas é exponencialmente distribuído com média 10^3 segundos, o mesmo tempo que passam inativas e suas taxas de transmissão são constantes e igual a 1Mbps. A capacidade máxima da fila de todos os roteadores é de 50 pacotes e a política de descarte é *Drop-Tail* [?]. Neste modelo foi usado um total de 80 fontes multimídia, gerando um tráfego médio de 40Mbps e 168 fontes TCP com um tráfego máximo de 40.32Mbps, este tráfego é o tráfego total que irá aos destinos "DEST1" somado a "DEST2".

O tempo total simulado no modelo foi de 83 horas (leva aproximadamente 10 dias em uma máquina *Pentium 4* 1.8GHz) e foram simuladas de 4 a 8 rodadas. Pelo tempo de simulação ser muito grande, o modelo pode ter entrado em estado estacionário, o que explica as rodadas terem quase as mesmas estatísticas (em todas as rodadas foram obtidos os mesmos resultados do gráfico da figura 4.9).

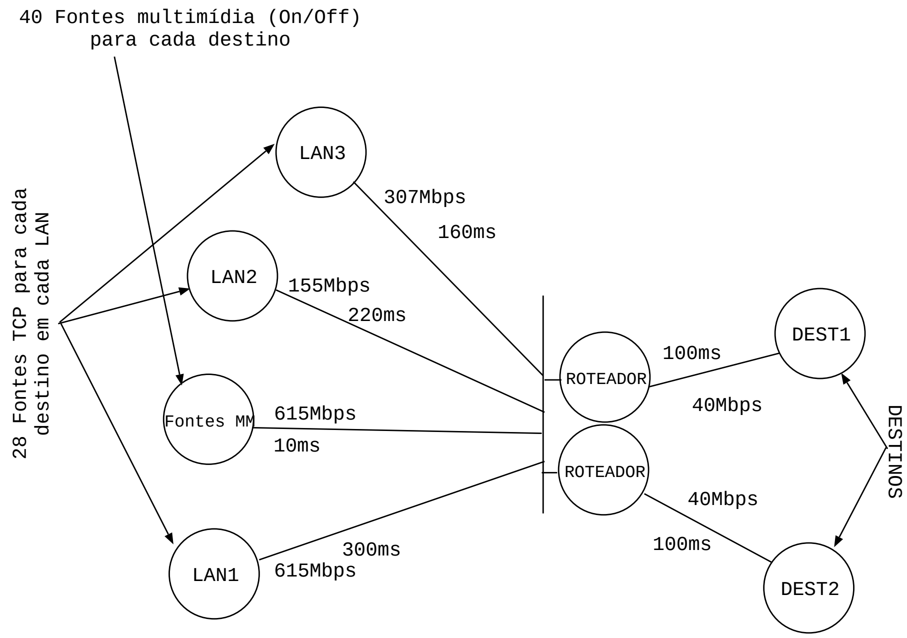

Figura 4.8: Modelo do NS

No gráfico da figura 4.9, pode-se notar o mesmo tipo de comportamento do que o obtido na figura 4.6, ou seja, somente uma fonte usando diversidade de caminhos é o cenário onde se obtém as melhores estatísticas para a fonte que faz diversidade de caminhos ("Apenas uma fazendo diversidade de caminhos"). O segundo melhor cenário é quando todas utilizam a diversidade de caminhos ("Todas as fontes usando diversidade de caminhos"). O terceiro melhor cenário para uma fonte usando diversidade de caminhos (e o melhor cenário para uma fonte que utiliza apena um caminho) é onde todas exceto uma fazem diversidade de caminhos ("Todas exceto uma fonte usando diversidade de caminhos"). O pior de todos os cenários é quando nenhuma fonte usa diversidade de caminhos ("Nenhuma fonte usando diversidade de caminhos").

Novamente observa-se a correlação de perda entre as duas filas quando há várias fontes fazendo diversidade de caminhos. Esta correlação é representada pela diferença nas curvas entre "Apenas uma fazendo diversidade de caminhos" e "Nenhuma fonte usando diversidade de caminhos" ser muito maior que a diferença entre "Todas as fontes usando diversidade de caminhos" e "Todas exceto uma fonte usando diversidade de caminhos". Ao mesmo tempo observa-se pela comparação das curvas

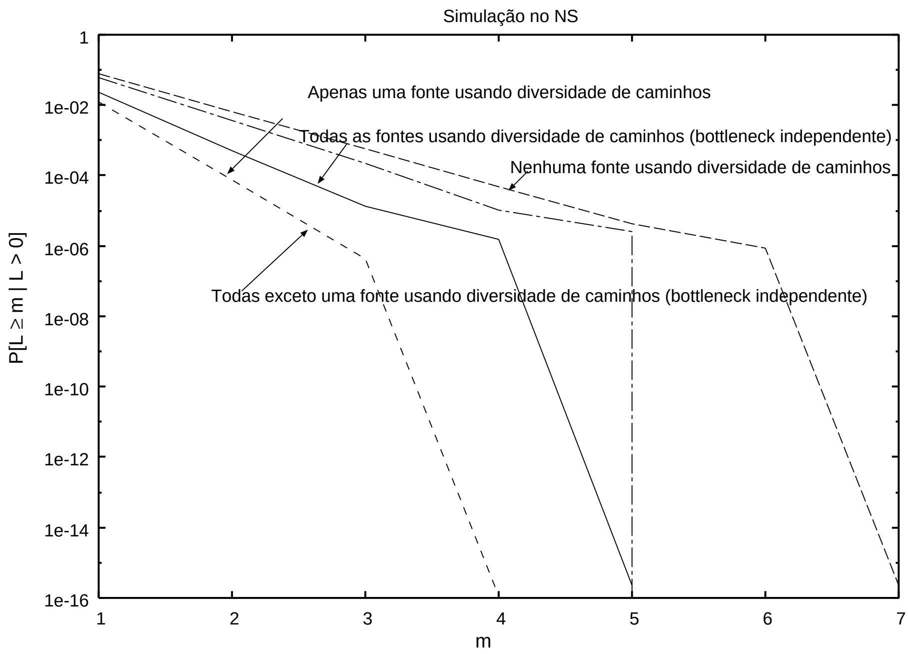

Figura 4.9: Distribuição condicional da rajada de perda para o modelo da figura 4.8.

de "Todas exceto uma fazendo diversidade de caminhos" e "Nenhuma fonte usando diversidade de caminhos" que a diversidade de caminhos diminui as rajadas de perda.

Capítulo 5

Conclusões

Devido à grande importância da qualidade de serviço nas aplicações em tempo real, diversos mecanismos de melhoria desta qualidade foram propostos na literatura. Entre estes mecanismos está a diversidade de caminhos que recentemente tem recebido grande atenção e muitos esforços de pesquisa.

O objetivo deste trabalho foi levantar dois pontos não cobertos pela literatura. O primeiro era verificar quais as vantagens e desvantagens ao se aplicar o método a mais de dois caminhos. O segundo era averigüar a performance do método e quais são os fatores que a influenciam caso a diversidade de caminhos fosse largamente utilizado por usuários na Internet ou em uma rede privada.

A primeira análise, capítulo 3, concluiu que a utilização de mais caminhos não necessariamente melhora a qualidade de serviço da aplicação. O impacto do aumento do número de caminhos sobre a qualidade de serviço depende dos requisitos da aplicação, mais precisamente, como ela se comporta com rajadas de perda de tamanho menor que o número de caminhos utilizados. Esta análise leva a concluir que se uma aplicação tem a opção de aumentar o número de caminhos utilizados, deve-se primeiro avaliar o possível impacto desta escolha sobre sua qualidade de serviço, calculando pela equação 3.1 se a inclusão do novo canal é vantajosa ou não para os seus requisitos de qualidade. O trabalho de Golubchick et al. [?] é o trabalho na literatura que mais se aproxima da análise feita neste capítulo, em especial a equação 3.1 que é a extensão para $N > 2$ caminhos de uma equação de Golubchick et al. [?] para o caso de dois caminhos.

A segunda análise, capítulo 4, estudou a escalabilidade da diversidade de caminhos com um grande número de clientes de tempo real. Este estudo visou identificar e isolar os diversos fatores que afetam a diversidade de caminhos quando há um grande número de clientes de tempo real, além de também quantizar a melhoria na qualidade de serviço em alguns cenários. Concluiu-se que para melhorar a qualidade de serviço para todas as aplicações, o melhor é utilizar a diversidade de caminhos em todos os clientes de tempo real. O melhor cenário individual, onde se está procurando a melhoria de apenas um cliente em tempo real, é que os demais clientes não utilizem a diversidade de caminhos e apenas este cliente de interesse a utilize.

Para estes estudos foi introduzida uma metodologia para análise de cálculo da distribuição do tempo que um reservatório enchido por uma fonte permanece cheio. Além desta análise, uma equação mais geral para a distribuição de rajada de perda condicional foi encontrada, facilitando a verificação se o acréscimo de mais um caminho é benéfico ou não para a aplicação. Também foram criados diversos modelos que auxiliam a análise de características de aplicações de rede de computadores.

5.1 Trabalhos Futuros

Apesar de já existirem muitos pontos analisados pela literatura e por este presente trabalho, muitos outros pontos novos e extensões deste trabalho podem ser exploradas, algumas são citadas abaixo:

- Incluir as equações de caminhos heterogêneos que, dentre todas as tarefas, parece ser a mais imediata.
- Fazer um estudo comparativo entre a diversidade de caminhos e a técnica de *packet interleaving* para aplicações que não sejam de tempo real.
- Melhoria do método estacionário para calcular a probabilidade do reservatório estar cheio e a fonte estar em um determinado estado. Os métodos utilizados

na seção 4.3.2 e descritos nas referências de [?] são numericamente instáveis para um grande número de estados.

- Quantificar, para alguns tipos de aplicações, qual o número ideal de caminhos de acordo com as estatísticas dos caminhos disponíveis.

Referências Bibliográficas