Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica
Departamento de Comunicações

MODELAMENTO DE UM SISTEMA DE REDE LOCAL MULTICANAL BASEADO EM TEORIA DE FILAS COM PRIORIDADES

Autora: Luciana Santos Pupo
Orientador: Reginaldo Palazzo Jr.

Este exemplar corresponde à redação final da tese defendida por Luciana Santos Pupo e aprovada pela Comissão Julgadora em 1º/04/91

Orientador

Tese apresentada à Faculdade de Engenharia Elétrica da UNICAMP para obtenção do grau de Mestre em Engenharia Elétrica.

Campinas-SP - Abril, 1991

UNICAMP
BIBLIOTECA CENTRAL

Agradecimentos

Gostaria de agradecer à CAPES pela Bolsa de Estudos oferecida durante toda a elaboração deste trabalho e à Telebrás, pela complementação e facilidades proporcionadas. Ao meu orientador, Prof. Dr. Reginaldo Palazzo Jr., agradeço pela oportunidade, pela orientação séria e enriquecedora, assim como, pela atenção e compreensão dispensadas. Agradeço, pelas contribuições e pelo tempo dedicado à leitura deste trabalho, aos membros da Banca Examinadora, Prof. Dr. Edson Moschim e, especialmente, ao Prof. Dr. Pushpa N. Rathie que me encaminhou à Engenharia Elétrica. E ainda, a Eunice Pissolato e Edson Ursini pelo tempo e atenção dedicados à orientação para obtenção de parâmetros de tráfego reais. Gostaria também de agradecer ao Paulo, meu marido, pelo apoio, compreensão e pela constante troca de idéias, ao meu filho, Caio, pela enorme alegria que nos trouxe e aos meus pais, pelas oportunidades que sempre proporcionaram à minha formação. Agradeço, ainda, aos meus parentes, amigos, colegas e funcionários que, de forma direta ou indireta, contribuiram para a estabilidade emocional, indispensável à conclusão deste trabalho.

Resumo

O desempenho de um sistema de rede local multicanal é analisado através da relação entre vazão e atraso médio do pacote para os protocolos M-CSMA. Filas são consideradas nos receptores para armazenar o tráfego bloqueado na recepção das mensagens. Em situações de tráfego misto, prioridades são associadas aos pacotes que requerem um atraso de processamento muito pequeno.

Abstract

The performance of a Multichannel Local Area Network System is analysed by the relationship between throughput and mean packet delay with the M-CSMA protocols. Queues are considered in the receptors to store the blocked traffic at the reception of the messages. For mixed traffic situations, priorities are associated with packets that require very small processing delay.

Conteúdo

1 Introdução 1

2 Redes Locais: Protocolos e Topologias 4

2.1 Introdução 4

2.2 Redes Locais 4

2.3 Protocolos de Múltiplo Acesso 8

2.3.1 Técnicas de Acesso com Alocação Fixa 9

2.3.2 Técnicas de Acesso Aleatório com Contenção 10

2.3.3 Técnicas de Acesso com Alocação por Demanda . . 15

2.3.4 Estratégias Adaptativas e Modos Mistos 16

2.4 Topologias 16

2.4.1 Redes em Estrela (*Star*) 17

2.4.2 Redes em Anel (*Ring*) 18

2.4.3 Redes em Barramento (*Bus*) 20

2.4.4 Redes em Árvore (*Tree*) 21

2.5 Protocolos e Topologias Associados 21

2.5.1 Redes em Anel 22

2.5.2 Redes em Barramento 24

2.6 Conclusões 25

3 M-LAN – Análise de Desempenho 26

3.1 Introdução 26

3.2 Rede local Multicanal 28

3.3 Análise da vazão 30

3.3.1 Protocolo CSMA não-persistente 30

3.3.2 Protocolos M-CSMA não-persistente 36

3.3.3 Protocolo M-CSMA-IC com colisões no receptor . . 46

3.3.4 Protocolos M-CSMA-CD não-persistentes 50
3.4 Análise do atraso do pacote 50
3.4.1 CSMA 50
3.4.2 M-CSMA-RC 52
3.4.3 M-CSMA-IC 54
3.4.4 M-CSMA-IC com colisões no receptor 54
3.5 Conclusões 56

4 Modelamento de uma M-LAN com Filas **58**
4.1 Introdução 58
4.2 Teoria de Filas 59
4.3 Especificação de Sistemas de Filas 60
4.3.1 Identificação do Processo Estocástico 62
4.3.2 Identificação da Estrutura e Disciplina 64
4.4 Modelamento de uma M-LAN com sobrecarga 65
4.4.1 Fila $M/M/m/K/L$ 66
4.4.2 Fila $M/M/1/K/L$ 70
4.5 Parâmetros de Desempenho 72
4.5.1 Atraso com o protocolo M-CSMA-IC com filas . . . 73
4.6 Filas com Prioridades 74
4.6.1 M-CSMA com prioridades 79
4.7 Conclusões 80

5 Resultados **82**
5.1 Introdução 82
5.2 Atraso Médio do Pacote: M-CSMA-IC-RC 83
5.3 M-CSMA-IC com Fila $M/M/1/K/L$ no Receptor 85
5.4 M-CSMA-IC com Fila $M/M/1$ com Prioridades 89

6 Conclusões Finais **99**

A M-CSMA-IC-RC: Divergência na Vazão **101**

B M-CSMA-IC-RC: Expressões para o Atraso **106**

Lista de Figuras

2.1 Topologia em Estrela. 17
2.2 Topologia em Anel. 19
2.3 Concentrador de Repetidores. 20
2.4 Topologia em Barramento. 20
2.5 Topologia em Árvore. 22
2.6 Registrador com Inserção de Retardo. 23

3.1 Arquitetura de uma M-LAN. 28
3.2 CSMA não-persistente: períodos livres e ocupados. 34
3.3 CSMA: Vazão x Tráfego. 37
3.4 M-CSMA-RC: Vazão x Tráfego. 41
3.5 M-CSMA-IC: Vazão x Tráfego, com poucos canais. 43
3.6 CSMA e M-CSMA-IC: Vazão x Tráfego. 45
3.7 CSMA e M-CSMA-IC-RC: Vazão x Tráfego. 48
3.8 CSMA e M-CSMA-IC-RC: Vazão x Tráfego. 49
3.9 CSMA: Atraso x Vazão. 52
3.10 CSMA e M-CSMA-IC: Atraso x Vazão. 55

4.1 Esquema desdobrado de uma M-LAN com N estações e filas 59
4.2 Diagrama de Transição de Estados 65
4.3 Fila $M/M/m/K/L$: Diagrama de Transição de Estados. . . 67
4.4 Fila $M/M/1/K/L$: Diagrama de Transição de Estados. . . 70

5.1 CSMA e M-CSMA-IC-RC: Atraso x Vazão. 84
5.2 Fila $M/M/1/M/M$: Número Médio de Pacotes. 85
5.3 Fila $M/M/1$: Tempo Médio de Espera x Taxa de Chegada . 87
5.4 Fila $M/M/1$: Atraso Médio do Pacote x Taxa de Chegada . 88

5.5 Fila $M/M/1$ com Prioridades: Tempo Médio de Espera, com 50% do tráfego para pacotes prioritários 90
5.6 Fila $M/M/1$ com Prioridades: Tempo Médio de Espera, com 80% do tráfego para pacotes prioritários 91
5.7 Fila $M/M/1$ com Prioridades: Tempo Médio de Espera, com 20% do tráfego para pacotes prioritários 92
5.8 M-CSMA-IC com a Fila $M/M/1$ com Prioridades: Atraso Médio do Pacote, com 50% do tráfego para pacotes prioritários 93
5.9 M-CSMA-IC com a Fila $M/M/1$ com Prioridades: Atraso Médio do Pacote, com 80% do tráfego para pacotes prioritários 95
5.10 M-CSMA-IC com a Fila $M/M/1$ com Prioridades: Atraso Médio do Pacote, com 20% do tráfego para pacotes prioritários 96
5.11 M-CSMA-IC com a Fila $M/M/1$ com Prioridades: Atraso Médio do Pacote, com 80% do tráfego para pacotes prioritários 97

A.1 M-CSMA-IC-RC: Vazão x Tráfego. 102
A.2 M-CSMA-IC-RC: Vazão x Tráfego, vazão comparada com pontos das curvas obtidas por Marsan. 103
A.3 M-CSMA-IC-RC: Atraso x Vazão, atraso comparado com pontos das curvas obtidas por Marsan. 105

B.1 M-CSMA-IC com colisões no receptor: Atraso x Vazão. . . 107
B.2 M-CSMA-IC-RC: diferença entre equações para atraso em relação à vazão. 108
B.3 M-CSMA-IC-RC: diferença entre equações para atraso calculado e estimado em relação à vazão. 109

Capítulo 1

Introdução

Desde o início da década de 70, o que se constata é um número crescente de pesquisas na área de redes utilizando da técnica de múltiplo acesso por radiodifusão (*broadcast*). Em particular, no final desta década e especificamente a partir de 1982, muitos trabalhos tratando da especificação, análise de desempenho, projetos e implementação de redes locais (LANs) foram publicados.

Em 1975, Kleinrock e Tobagi [4] propuseram e analisaram dois protocolos de acesso aleatório (CSMA[1]) com o objetivo de ver aumentada a eficiência dos protocolos até então existentes (ALOHA). Desta análise, ficou evidente a superioridade dos protocolos propostos quando tomados como índices de otimalidade a vazão e o atraso.

Em outro artigo [11], Tobagi apresentou um estudo sobre protocolos de múltiplo acesso para redes de pacotes. Neste artigo de revisão foram classificados os protocolos de acordo com a natureza estática, dinâmica ou adaptativa do mecanismo de alocação de faixa, de acordo com a natureza centralizada ou distribuída do processo de tomada de decisão e o grau de adaptabilidade do algorítmo conforme as necessidades do momento.

Da busca por protocolos mais eficientes com respeito aos tradicionais índices de otimalidade é que foi proposto por Marsan e Roffinella [6] um protocolo de múltiplo acesso para redes locais com comutação por pacotes utilizando múltiplos canais em um esquema *broadcast*. Como previsto pela Teoria de Informação com multi-usuários, a região da capacidade de canal viabiliza taxas de sinalização maiores que aquelas alcançadas em esquemas

[1] CSMA: Carrier Sense Multiple Access

tradicionais de um único par transmissor-receptor. Desse modo, melhoras significativas, em termos de desempenho, foram obtidas com os novos protocolos derivados do CSMA e CSMA-CD[2].

Cinco anos mais tarde, foi apresentado um novo artigo de revisão [9] contendo o estado-da-arte dos protocolos de múltiplo acesso. Como resultado desta coletânea de protocolos ficou claro o interesse crescente na comunicação por pacotes com tráfegos regulares (voz), além dos denominados rajadas (dados). Como consequência, surgiu a necessidade de se propor novos protocolos de múltiplo acesso para viabilizar a capacidade e características necessárias para se integrar aplicações diferentes em uma mesma rede, inclusive com serviços de comunicação em tempo real.

Devido a este fato, surgiu a motivação para se estudar mais profundamente os protocolos de múltiplo acesso ao canal associados a topologias adequadas a diversas aplicações. E, a partir deste estudo, ter condições de propor uma nova solução para o problema de troca de informações diversas entre usuários dentro de um ambiente restrito (uma rede local), associando prioridade a um protocolo de múltiplo acesso com multicanais por onde são transmitidos pacotes com mensagens que geram tanto tráfegos regulares quanto em rajadas. Assim, devido à dificuldade e escopo do problema a ser considerado, nos ateremos em progressivamente apresentar e analisar a proposta de inclusão de prioridade ao protocolo de múltiplo acesso com multicanais para a integração de voz e dados.

Pretendemos, desta maneira, apresentar uma análise de desempenho para uma solução adequada às necessidades atuais, tendo em vista uma futura interligação dos meios de comunição através de uma rede digital de serviços integrados (RDSI) que transmitirá os vários tipos de mensagens entre distâncias maiores, mas que para pequenas distâncias sua velocidade de transmissão ficará aquém da obtida através de uma rede local (LAN).

No Capítulo 2, faremos uma pequena revisão de redes locais, das principais topologias e técnicas de múltiplo acesso adequadas. Em seguida, no Capítulo 3, estudaremos a vazão e o atraso para sistemas de redes locais com os protocolos CSMA com multicanais, comparando-os com o protocolo CSMA. No Capítulo 4, apresentaremos a idéia de introduzir filas nas estações tanto para a transmissão como para a recepção aplicáveis em situações de transbordo no tráfego. Ainda no Capítulo 4, acrescentaremos

[2]CSMA-CD: CSMA *with Collision Detection*

prioridades às filas e consideraremos tráfego misto de pacotes (por exemplo, dados e voz) através da rede. As comparações dos diversos resultados obtidos anteriormente e neste trabalho serão apresentadas no Capítulo 5. E, finalmente, o Capítulo 6 tráz as conclusões.

Capítulo 2

Redes Locais: Protocolos e Topologias

2.1 Introdução

Neste capítulo, apresentaremos um resumo sobre redes locais. Com este objetivo, abordaremos as principais técnicas de acesso ao canal, procurando salientar suas vantagens, desvantagens e aplicações. Apresentaremos também, as principais topologias para redes locais e os protocolos de múltiplo acesso mais utilizados com as topologias de maior interesse.

2.2 Redes Locais

As redes locais são constituidas por um conjunto de estações (nós) interligadas através de um sistema de comunicação localmente distribuído. Este sistema de comunicação é composto de um arranjo topológico interligando os vários nós e de um conjunto de regras (protocolos) de forma a organizar a comunicação entre os nós.

Cada arquitetura de rede local possui características próprias, mas de maneira geral, podemos citar algumas características comuns às redes locais:

- Altas taxas de dados.
- Pequeno atraso de propagação.

- Distância relativamente pequena entre estações (Km).
- Meio de transmissão de baixo custo no caso de par trançado e cabo coaxial.
- Número grande e variável de equipamentos requerendo interconexão.

Tais características requerem redes de comunicação com topologias simples e interfaces de conexão, também simples e de baixo custo, de forma a proporcionar grande flexibilidade na acomodação da variabilidade do tráfego na rede e que atinja um nível desejável de confiabilidade.

As principais caraterísticas exploradas nas técnicas de múltiplo-acesso apropriadas a ambientes locais são: pequeno atraso de propagação e alta taxa de dados, que proporcionam um maior desempenho pela alta utilização do canal e pequeno atraso das mensagens.

Restrições, como simplicidade e baixo custo, conduzem à implantação de um único canal com faixa larga a ser compartilhado por usuários independentes.

Redes com transmissão em faixa larga utilizam-se da técnica de dividir a faixa de frequência disponível em vários canais separados, cada um capaz de transportar dados em elevadas taxas de transmissão.

A medida em que se multiplexa vários tipos de tráfego através de um sistema com faixa larga, se torna essencial que o protocolo de acesso tenha a habilidade de comportar, simultaneamente:

- tráfegos diferentes
- prioridades diferentes
- mensagens de comprimento variável
- limitações no atraso da mensagem

É, também, de grande importância garantir a correta operação dos protocolos com controle distribuído, assim como a robustez, ou seja, a prevenção contra a insensibilidade a erros que resultam em informação errada.

Os principais atributos que atuam como parâmetros de comparação nas redes locais são:

- custo: estações, interfaces e meio

- confiabilidade
- tempo de resposta = tempo de acesso + tempo de transmissão
- velocidade
- desempenho: capacidade efetiva de transmissão da rede
- facilidade de desenvolvimento
- modularidade
- reconfigurabilidade
- complexidade lógica
- facilidade de uso
- disponibilidade
- facilidade de manutenção
- dispersão geográfica
- etc.

O meio de transmissão para redes locais consistia, originalmente, no par trançado, cabo coaxial e alguns meios de transmissão sem fio. Avanços na tecnologia de fibras óticas criaram um potencial para redes locais que necessitam de ampla faixa de frequências para a sua viabilização.

As topologias e protocolos em anel, que requerem propagação unidirecional do sinal, incorporaram naturalmente a nova tecnologia. O barramento, entretanto, que inicialmente apresentava a vantagem da capacidade natural de radiodifusão ou bidirecionalidade, necessita restringir sua aplicação ao uso do barramento unidirecional para poder incorporar as fibras óticas de forma a atingir capacidade de radiodifusão de três maneiras diferentes:

1. através de dois barramentos separados com sinais propagando em direções opostas.
2. através de um barramento dobrado com receptores e sensores ligados ao canal de entrada e transmissores ao canal de saída.

3. através de um barramento duplamente dobrado, proporcionando canais de entrada e de saída.

Novos atributos para o projeto de redes surgiram com os altos produtos faixa-distância atingíveis com a fibra ótica. Dentre eles, está a velocidade lógica dos transmissores e receptores, assim como a necessidade de superar as disparidades da latência do sinal recebido e de garantir o mesmo nível nos receptores dos sistemas em barramento. Novas interfaces e sistemas foram criados para lidar com a baixa impedância das ligações unidirecionais do meio ótico e com seu alto produto faixa-distância[1].

Concomitantemente com o progresso da tecnologia ótica, comunicações envolvendo tráfegos integrados (voz, dados, imagens, etc) foram sendo desenvolvidas. Da mesma forma, foram sendo desenvolvidas novas topologias e protocolos para ILANs[2].

Durante este período, surgiu uma nova categoria de redes locais denominada M-LAN (*Multi-channel Multiple-access Local Area Network*), que faz uso de vários canais paralelos, com taxas moderadas, aos quais todas as estações estão conectadas. Estas redes atingem produtos faixa-distância muito elevados e evitam atrasos com o aumento do atraso de propagação normalizado[3], a, que é proporcional a distância. Ou seja, canais com velocidades moderadas, em paralelo, permitem o uso de faixa larga com alta taxa resultante, enquanto atingem distâncias maiores mantendo alto o número de estações conectadas à rede.

Em [6], foram propostos novos protocolos de acesso para esta nova categoria de redes locais: M-CSMA-RC, M-CSMA-IC, M-CSMA-CD-RC e M-CSMA-CD-IC (onde IC significa *Idle Choice* e RC, *Random Choice*). Os protocolos IC apresentaram melhor desempenho que os RC, e a característica CD, garantiu um melhor desempenho que o mesmo protocolo sem detectar colisões.

Estes novos protocolos, porém, foram analisados considerando-se apenas um tipo de mensagem compondo seu tráfego de informações. Alguns estudos posteriores trataram do problema de tráfego misto em M-LANs, mas

[1] Produto Faixa-Distância é a faixa atingível de um sinal transmitido por uma distância física e que pode ser recebido adequadamente sem a necessidade de repetidores.

[2] *Integrated LANs*

[3] a consiste no atraso de propagação fim-a-fim, τ, normalizado em relação ao tempo médio de transmissão de um pacote, T, ou seja, $a = \tau/T$.

ainda não se verificou - por parte dos autores - a aplicação de prioridades aos pacotes que exigem menor tempo de processamento.

Diante destas informações, partimos para a análise de desempenho de um sistema em barramento, dividido em M canais em paralelo e que se utiliza de um protocolo de acesso como o M-CSMA-IC, integrando os vários tipos de mensagens (como voz e dados, por exemplo). Para esta análise, consideramos um modelamento, baseado em Teoria de Filas com prioridades, para a obtenção de parâmetros tais como: o tempo médio de espera na fila, o número médio de usuários no sistema e na fila. Estes parâmetros nos permitem quantificar o desempenho da rede em questão e, consequentemente, compará-lo a outros resultados já conhecidos.

2.3 Protocolos de Múltiplo Acesso

Um protocolo de acesso é um conjunto de regras para o acesso ao meio usado para que seja possível manter um certo controle sobre a disciplina de comunicações através da rede.

Sempre que equipamentos caros, ou mesmo, recursos raros forem compartilhados por um grupo de usuários independentes e concorrentes, estes protocolos se fazem necessários. Isto se deve, principalmente, à necessidade de proporcionar um alto grau de conectividade entre os usuários com uma utilização eficiente. Um protocolo ideal é aquele que possibilita ao usuário ter a sensação de que o meio está totalmente dedicado às suas transmissões.

Os protocolos podem ser classificados como centralizados ou distruidos. Um protocolo é dito centralizado quando existe um equipamento central responsável pela sua coordenação. E distribuído, quando não existe tal equipamento e a coordenação é distribuída pelos diversos componentes da rede. Os protocolos ditribuidos são mais confiáveis que os centralizados, pois a falha do equipamento central poderia ocasionar a interrupção completa das transmissões.

Os usuários podem estar sujeitos a protocolos com conteção. Um protocolo com contenção permite que vários usuários, utilizando equipamentos diferentes, ou não, transmitam, simultaneamente, pelo mesmo meio de forma que uns podem interferir nas transmissões dos outros.

Por outro lado, quando apenas um usuário pode transmitir a cada instante (existindo, portanto, uma ordem de transmissão), os protocolos são

denominados sem contenção. A contenção é mais indicada para situações de baixo utilização, quando a probabilidade de interferência é menor, enquanto que a não contenção é adequada ao tráfego intenso quando a probabilidade de interferência é maior.

De acordo com Tobagi [11], os protocolos de múltiplo acesso podem ser divididos em classes como descritas a seguir:

2.3.1 Técnicas de Acesso com Alocação Fixa

Consistem em alocar uma parte da faixa total a cada um dos usuários de forma estática, independentemente de suas atividades. É bastante eficiente para tráfego regular (voz).

- FDMA (*Frequency Division Multiple Access*):

  Sua implementação é simples, não requerendo coordenação em tempo real entre os usuários. Baseia-se na pré-alocação da faixa por divisão de frequências diferentes entre usuários diferentes. Nesta divisão, a faixa fica sub-utilizada devido à necessidade de se deixar uma banda de resguardo (*guard band*) para a separação entre as diferentes frequências. Não permite flexibilidade na mudança de faixas.

- TDMA (*Time Division Multiple Access*):

  Consiste em alocar a faixa toda para cada usuário durante um período de tempo a ele reservado. Desta forma, permite uma boa conectividade entre os usuários, pelo mesmo canal, de forma síncrona. Sua implementação é mais complexa que a do FDMA, mas seu atraso de propagação é menor indicando um melhor desempenho.

  Algumas de suas desvantagens podem ser caracterizadas devido à necessidade de:

  - se ter conversores A/D[4] para as informações de tráfego contínuo (voz)
  - uma sincronização rápida na recepção das informações
  - um tempo de espera pelo período de tempo reservado à transmissão

[4] Conversores Analógico/Digital

- ATDMA (*Asynchronous TDMA* ou multiplexação estatística):

  Consiste, basicamente, em comutar a alocação do canal de um usuário para outro somente após o término da transmissão em andamento e para um usuário que tenha informações a serem transmitidas. Desta forma, o canal é alocado dinamicamente para os vários usuários de acordo com suas necessidades.

  Seu desempenho em sistemas de comunicação por pacotes corresponde a um sistema de filas com um único servidor e é o melhor que se pode atingir para demandas imprevisíveis. Devido à dificuldade de coordenação entre os usuários, este modo de multiplexação só é possível entre usuários que compartilham um único canal ponto-a-ponto.

- CDMA (*Code Division Multiple Access*)

  Permite sobreposição na transmissão de informação tanto no domínio do tempo quanto no domínio da frequência, devido ao "princípio" da ortogonalidade entre os diferentes códigos de sinalização em conjunto com filtros casados nos respectivos receptores. Desta forma, é possível a coexistência de vários usuários no sistema ocupando a mesma faixa, desde que usem códigos diferentes. Nas técnicas de acesso aleatório, onde os pacotes podem colidir, é possível separar, por captura, sinais sobrepostos no tempo atingindo conectividade e utilização eficiente do espectro.

2.3.2 Técnicas de Acesso Aleatório com Contenção

Nestas técnicas, a faixa toda fica disponível aos usuários como um único canal, ao qual usuários fazem acesso aleatoriamente. Colisões podem ocorrer e, para que o desempenho não seja degradado, pode-se sincronizar os usuários, permitir que o canal seja monitorado antes da transmissão ou ambos. É mais adequado para tráfegos com características de intermitência.

- ALOHA

  Uma das primeiras redes geograficamente distribuídas a ser projetada, e que utiliza técnicas hoje associadas às redes locais. O sistema, chamado por rede ALOHA, entrou em funcionamento em 1970, com o objetivo de permitir o acesso simples e de baixo custo a um grande número de terminais de um sistema computacional no Havaí.

A característica essencial da rede ALOHA é o uso de um canal de radiodifusão, via satélite, por todos os usuários, que competem pela sua utilização. Pacotes de dados são transmitidos ao mesmo tempo, usando informações de consistência de dados dos próprios pacotes.

Esta técnica apresenta como vantagens:

- simplicidade
- baixo custo
- adequação a pequenos tráfegos

E, como desvantagens:

- não existe limite máximo para o atraso
- não existe prioridade de acesso
- grande volume de tráfego de retransmissões e de colisões podem tornar a rede inoperante.

As duas versões, citadas a seguir, são teoricamente, aplicáveis aos ambientes: satélite, rádio e barramento comum local.

- ALOHA Pura: Comuta pacotes por radiodifusão com topologia em barramento. As transmissões são feitas a qualquer instante. Um sinal de reconhecimento de mensagem (ACK) é enviado no sentido contrário ao da mensagem indicando o recebimento correto. Caso isso não ocorra num certo intervalo de tempo pré-definido, a mensagem é retransmitida após um tempo aleatório contado a partir da expiração do prazo para recebimento do ACK. Sua vazão supera a do TDMA para baixo tráfego (abaixo de 18%). Muitas lacunas são geradas devido às colisões, que podem ocorrer em qualquer instante da transmissão. Durante as colisões, o canal fica inutilizado por um período de tempo no qual dois pacotes poderiam ser enviados, caracterizando, desta forma, a sub-utilização do canal.
- ALOHA Segmentada (*Slotted*):

  Os pacotes são transmitidos em segmentos (*slots*) de tempo com duração fixa, igual ao tempo de transmissão de um pacote. Os

usuários transmitem sempre no início de cada segmento fazendo com que, caso haja colisão, a sobreposição seja completa. Apresenta maior eficiência na utilização do canal com uma vazão máxima de 37% , e melhores resultados para tráfegos intermitentes sob a hipótese de que o número de usuários é extremamente grande.

- CSMA (*Carrier Sense Multiple Access*):

  Cada estação monitora o canal antes de transmitir, evitando colisões caso alguma transmissão esteja em andamento. Mas colisões podem ocorrer devido ao atraso de propagação do sinal no canal ser não-nulo e à possibilidade de tentativas de transmissão ocorrerem simultaneamente. Neste caso, as retransmissões são programadas para um certo tempo aleatório mais tarde.

  Existem, basicamente, dois tipos diferentes de CSMA quanto à política de retransmissão: p-persistente e não-persistente.

  - CSMA não-persistente

    Se o canal estiver livre, transmite, mas ao perceber o canal ocupado, o usuário programa sua transmissão para um certo tempo aleatório mais tarde. Sua vazão máxima chega a 80% .

  - CSMA p-persistente

    Se a estação perceber que o canal está ocupado, aguarda até que o mesmo esteja livre e, então, segue um dos seguintes procedimentos: ao perceber o canal livre, o usuário transmite com probabilidade p, ou programa sua transmissão para τ (τ = atraso de propagação) segundos mais tarde com probabilidade 1-p. Se o canal ainda estiver livre, repete este procedimento, caso contrário, o usuário age como se tivesse colidido e programa sua retransmissão para um certo tempo aleatório mais tarde.

    O CSMA puro é igual a um caso especial do CSMA p-persistente quando $p = 1$, onde a transmissão sempre ocorre após a transmissão em andamento e desde que tenha alguma estação com pacote pronto para transmitir. Neste caso, elimina-se o problema da sub-utilização do canal, pois desaparecem as lacunas. Em

contrapartida, muitos conflitos são gerados. A vazão máxima fica em 50% .

Para $p = 0.5$, a colisão pode ser evitada em 50% das vezes, em 25% das vezes deve haver colisão e em 25% das vezes pode-se introduzir lacunas [7].

- CSMA Segmentado (*Slotted*)

  O tempo é divido em mini-segmentos (mini-*slots*) de tamanho igual ao atraso de propagação máximo (τ) entre todos os pares. Logo, para se transmitir um pacote leva-se o tempo de vários mini-segmentos. Todos os terminais são sincronizados e forçados a transmitir no início de um mini-segmento. Se um pacote chega durante um mini-segmento, o terminal aguarda até o início do próximo mini-segmento e segue algum dos protocolos CSMA.

- CSMA-CD

  Neste protocolo, ao contrário dos demais protocolos tipo CSMA, o sinal de reconhecimento (ACK) não é enviado. As estações, ao transmitir um pacote, continuam a monitorar o canal e, no caso de colisão, são capazes de detectar e interromper suas transmissões liberando o canal. Enquanto isso, programam para retransmitir, após detectar o canal livre, ou segundo alguma das regras já mencionadas.

  O fato de não ter mensagem de reconhecimento faz com que o tráfego aumente elevando a vazão para 98% do ideal para situações com pouco tráfego e pequenas distâncias. Por outro lado, a presença de ruído mascara uma transmissão tida como com êxito. Isto é, uma mensagem transmitida e que não colida é tida como recebida corretamente mesmo que o ruído altere seu conteúdo, só que neste caso não haverá retransmissão.

  Trata-se, portanto, de um procedimento Datagrama Não-Confiável. Logo, este protocolo não é indicado para controle de processos críticos, mas sua utilização é de grande valia nas aplicações cujo retardo é importante e onde a fidedignidade não é tão rígida, como é o caso da voz [7].

  Para a transmissão de pacotes com dados que exijam muita fidedignidade, acrescenta-se o sinal de reconhecimento em detrimento do

tempo de transmissão.

- CSMA-CA (*CSMA with Collision Avoidance*):

  É a técnica mais utilizada nas redes de alta velocidade. Se o canal estiver livre, as estações seguem algum protocolo CSMA para obter acesso ao canal, podendo ocorrer colisões. Após cada transmissão, a rede entra em um modo onde as estações só podem começar a transmitir em intervalos de tempo a elas reservados. Ao findar uma transmissão, a estação alocada ao primeiro intervalo tem o direito de transmitir sem probabilidade de colisão. Se não o faz, o direito passa à estação alocada ao segundo intervalo e assim sucessivamente, até que ocorra uma transmissão, quando então todo o processo reinicia. Se todos os intervalos não são utilizados, a rede entra no estado livre para reinicialização [10].

- BTMA (*Busy Tone Multiple Access*):

  Possui comportamento e controle dinâmicos. Surgiu para solucionar o problema do terminal escondido em ambientes onde os usuários se comunicam com uma estação central (como é o caso da ALOHA). É feita uma divisão da faixa disponível em dois canais separados: um canal de sinal ocupado e outro de mensagens. Assim que a estação central detecta portadora no canal de mensagens, coloca um sinal de ocupado no outro canal. A dificuldade de se detectar a portadora no canal de mensagens quando se utiliza toda a faixa (como no CSMA) é mínima e, portanto, desprezada.

- SSMA (*Spread Spectrum Multiple Access*):

  É a forma mais comum de CDMA, onde a cada usuário é alocado um código diferente que é modulado na portadora e os dados são modulados baseados neles. Existem dois tipos mais comuns:

  - *Frequency-Hopped* SSMA: a frequência muda periodicamente de acordo com algum padrão conhecido.
  - *Phase-Hopped* SSMA: a portadora é modulada fasicamente pela sequência digital de dados e a de códigos.

É útil em várias aplicações: comunicações por satélite, *ground-radio* móvel, redes de comunicação de computadores. Usado na PRNET, devido à segurança, coexistência com outros sistemas e a habilidade de neutralizar os efeitos de multicaminhos. Beneficia-se do efeito de captura no SSMA assíncrono.

2.3.3 Técnicas de Acesso com Alocação por Demanda

Estas técnicas estão situadas entre as técnicas de acesso por alocação fixa e as de acesso aleatório. As primeiras apresentam capacidades menores em face a pequenos atrasos e as últimas devido a colisões. As técnicas de acesso por demanda requerem informações de controle explícitas sobre as necessidades das fontes de comunicação.

Em [11], estas técnicas foram classificadas segundo a natureza centralizada ou distribuída do controle. As técnicas de controle centralizado, são aquelas, nas quais, o processo de tomada de decisão é centralizado. Dentre elas, temos: os Sistemas com Circuito Orientado, os Sistemas *Polling*, *Polling* Adaptativo ou *Probing* e SRMA (*Split Channel Reservation Multiple Access*).

As técnicas de controle distribuído são aquelas em que todos os usuários executam um algorítmo distribuído baseado nas informações de controle trocadas entre eles, atingindo um aumento da confiabilidade e do desempenho. Alguns exemplos são: os Sistemas de Reserva ALOHA, FIFO (*First-In-First-Out*) e RR (*Round-Robin*), além do MSAP (*Mini-Slotted Alternating Priorities*). BRAM (*Broadcast Recognizing Access Method*). Algorítmo de Retransmissão em Árvore, Anel com Passagem de Permissão (*Token-Passing Ring*) e Anel com Inserção de Retardo (*Register-Insertion Ring*).

Já em [9], as técnicas de alocação por demanda foram consideradas como uma classe de esquemas, livres de colisão, que programam o acesso ao canal pelas múltiplas estações, segundo a natureza cíclica ou não do serviço.

A ordem de serviço cíclica é uma programação de estações em *round-robin* obtida através de um *token*, implícito ou não. Dentre as de controle não-cíclico, incluem-se: Anel de Segmentos (*Slotted Ring*) e Anel com Inserção de Retardo, que serão descritas mais adiante.

Esquemas de acesso por demanda são adequados a cargas entre médias e altas e, particularmente, para aplicações que requeiram tempo de acesso ao canal previsível e atraso de pacote limitado. Apresentando, entretanto,

um grande *overhead* quando os sistemas estão levemente carregados.

As técnicas de alocação por demanda, com serviço cíclico, foram apresentadas para três topologias: em anel, em barramento unidirecional (UBS[5]) e barramento bidirecional.

O protocolo de passagem de permissão pode ser implementado sobre as topologias em anel e em barramento bidirecional. Ainda no barramento bidirecional, se pode implementar, entre outros, os seguintes métodos de acesso: MSAP, BID, BRAM, SILENTNET, L-Expressnet, que constituem uma classe de *token* implícito, no acesso por demanda, denominada *Scheduling Access Delay*. Já o acesso ao barramento unidirecional pode ser de reserva (UBS-RR) ou *Attempt-and-Defer* (Expressnet, Fastnet, D-net, U-net e protocolos sem *token* (*tokenless*)).

2.3.4 Estratégias Adaptativas e Modos Mistos

Incluem as técnicas que misturam vários modos de acesso diferentes e as estratégias nas quais a escolha de um esquema de acesso é adaptativa às necessidades do usuário, com o objetivo de atingir desempenho próximo ao ótimo o tempo todo.

Até 1980, os esquemas conhecidos, entre outros, eram: URN, CSMA combinado com TDMA, SRUC (*Split Reservation Upon Collision*), PODA (*Priority Oriented Demand Assignment*) e MACS (*Mixed ALOHA Carrier Sense*) [11].

Ao longo da década de 80, foram surgindo propostas de novas técnicas de acesso que se utilizavam da estratégia adaptativa à intensidade de tráfego: Buzznet, CSMA-DP-LCP (CSMA-CD *with Dynamics Priorities and Low Collision Probability*), CSMA-DCR (CSMA *with Deterministic Contention Resolution*), entre outras, [9].

2.4 Topologias

Existem vários tipos de topologias para redes dentre as quais podemos destacar as topologias em estrela, em anel, barramento comum e em árvore. Em seguida, descreveremos cada uma delas.

[5] Unidirection Bus System

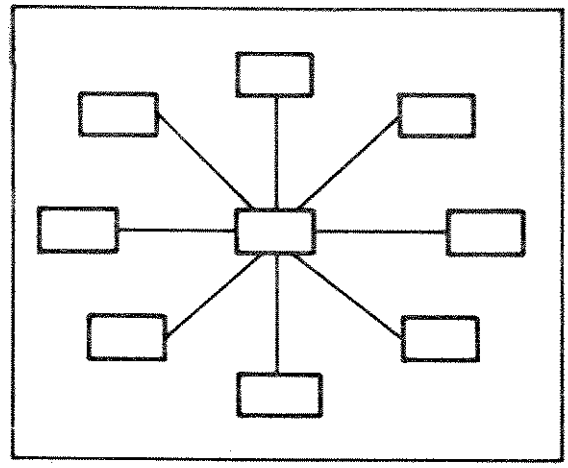

Figura 2.1: Topologia em Estrela.

2.4.1 Redes em Estrela (*Star*)

Cada um dos nós da rede está ligado a uma estação (ou nó) central, ativa ou passiva, através de enlaces dedicados, evitando problemas de roteamento e sequenciamento como é mostrado na Figura 2.1.

Quando a estação central é ativa, esta age como repetidor e como meio de acesso controlador. Por outro lado, quando a estação central é passiva, esta serve para ligar as N portas de entrada às N portas de saída.

Todas as mensagens que circulam pela rede devem passar pela estação central que na maioria das configurações, além da função de gerência das comunicações, tem também a função de processamento dos dados.

O gerenciamento das comunicações pode ser por chaveamento de pacotes ou de circuitos. Caso a estação central gerencie através de chaveamento de pacotes, ou seja, retransmitindo pacotes para seus destinos, tais estações necessitariam de *buffers* de armazenamento.

As vantagens desta topologia são a simplicidade funcional e ausência de roteamento. Suas principais desvantagens são:

- Confiabilidade: baixa, ou seja, alta probabilidade de falhas. A falha da estação central gera interrupção de todos os serviços. Tentativas para eliminar a falta de confiabilidade do nó central acabam tornando

complexo o sistema.

- Modularidade: permite número baixo de estações ligadas ao nó central, pois depende da capacidade de gerenciamento do mesmo.
- Custo: custo total alto, pois inclui o nó central. O custo total é dividido pelo número de estações, e como não são muitas, o preço por estação é alto.
- Desempenho: baixo, pois todas as mensagens têm que passar pelo nó central.

Esta topologia foi muito usada para automação de escritórios, onde o usuário já possuia, por exemplo, um PABX. Salvo casos deste tipo, esta topologia não tem sido muito difundida.

2.4.2 Redes em Anel (*Ring*)

Cada estação está interligada a exatamente duas outras, uma antecessora e uma sucessora (ver Figura 2.2). Embora teoricamente a transmissão possa ser feita nos dois sentidos, as implementações existentes têm utilizado apenas a transmissão unidirecional, principalmente devido à grande simplificação dos protocolos envolvidos, a não ser em casos que usem dois anéis em paralelo, um em cada direção.

- Vantagens:
  - Simplicidade do protocolo.
  - Simplicidade do hardware: interfaces simples nas estações.
  - Ordenação cíclica natural: elimina problemas de roteamento e sequenciamento.
  - Transmissão imediata, apenas um circuito de armazenamento.
  - Desempenho alto, possibilidade de mais de uma mensagem ser transmitida ao mesmo tempo.

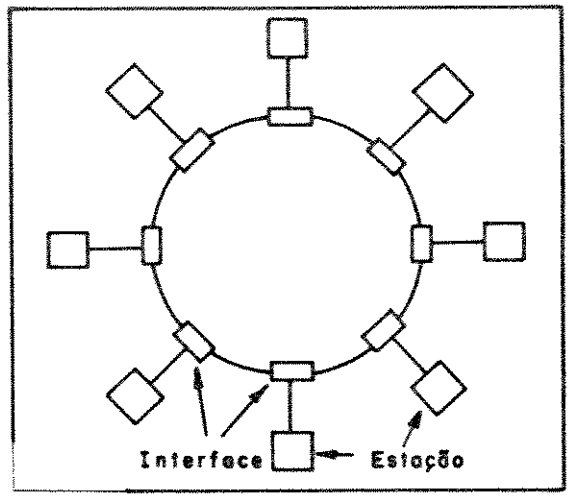

Figura 2.2: Topologia em Anel.

- Desvantagens:
  - Baixa confiabilidade decorrente da utilização de um meio ativo de interligação (repetidores). Diante de falhas em um dos repetidores, a interligação é descontinuada.
  - Tempo de transferência de mensagens é grande em relação ao barramento.
  - Pouca flexibilidade. Para ampliar a rede, inclusão de uma estação, é necessário refazer as conexões existentes entre as duas estações.

Várias propostas foram sugeridas para tentar solucionar o problema da confiabilidade. A primeira delas foi isolar o repetidor com defeito (*bypass*), comprometendo a intensidade do sinal. A degradação pode ser evitada, construindo-se concentradores de repetidores (*ring wiring concentrator*), como mostra a Figura 2.3 , permitindo o distanciamento entre as estações.

Pode-se, também, ter um anel duplo, triplo, ..., n-uplo a fim de recuperar 1, 2, ..., n-1 falhas, respectivamente. Esta configuração apresenta enlaces de reserva para evitar interrupções em caso de falha. Pode-se dispor de múltiplos anéis interligados para tentar simular um barramento virtual em termos de confiabilidade.

Figura 2.3: Concentrador de Repetidores.

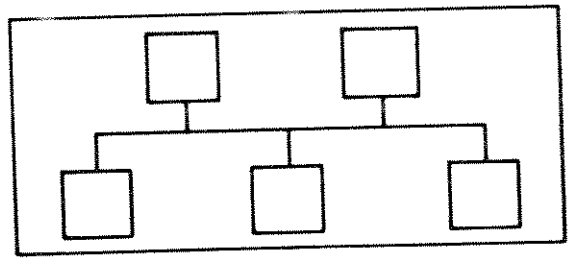

Figura 2.4: Topologia em Barramento.

2.4.3 Redes em Barramento (*Bus*)

Todas as estações estão conectadas a um único barramento, de forma que uma determinada transmissão é recebida por todas as demais estações (*broadcast communication*), eliminando problemas de armazenamento, roteamento e sequenciamento (ver Figura 2.4). Esta tem sido considerada a topologia mais adequada para redes locais, devido à sua inerente confiabilidade e capacidades naturais de radiodifusão, características bastante desejáveis para LANs.

- Vantagens:
  - Alta confiabilidade decorrente da utilização de um meio passivo na interligação dos equipamentos; a falha de um ou mais equipamentos não provoca descontinuidade na ligação entre os demais.
  - Pequeno tempo de transferência das mensagens.

  - Muita flexibilidade na expansão; seja qual for a localização desejada para o novo equipamento, nenhuma modificação nas ligações existentes se faz necessária.
  - Baixo custo das conexões.

- Desvantagens:
  - Protocolo mais complexo que para anéis.
  - Sendo o meio com transmissões tipo radiodifusão, apenas uma transmissão pode estar em andamento a cada instante.
  - Degeneração do sinal para distâncias superiores a L (L = distância máxima de transmissão pelo meio sem regeneração do sinal).
  - O custo ainda pode ser elevado, se o meio de transmissão for a fibra ótica.

Para se tentar melhorar o problema da degeneração pode-se utilizar barramentos múltiplos interligados por repetidores, introduzindo, porém, o problema da confiabilidade encontrado nos anéis. Supondo o pior caso, quando se tem um repetidor por estação, encontramos um barramento simulando um anel virtual.

2.4.4 Redes em Árvore (*Tree*)

É uma extensão do conceito de barramento, consistindo em vários barramentos interligados.

2.5 Protocolos e Topologias Associados

Dentre as topologias conhecidas, iremos salientar as em anel e em barramento por serem as mais adequadas para redes locais. A seguir, destacaremos os protocolos mais adequados para estas topologias:

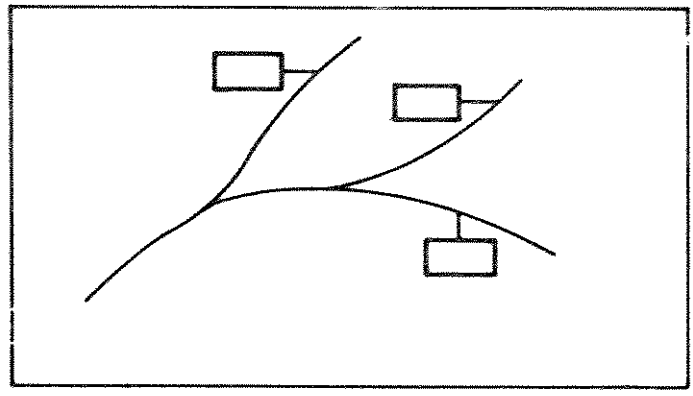

Figura 2.5: Topologia em Árvore.

2.5.1 Redes em Anel

- Passagem de Permissão (Padrão IEEE 802.5 - *Token Ring*) Consiste na passagem de um caracter de controle (*token*) entre os diversos equipamentos ligados ao anel. A estação que deseja transmitir, aguarda que o caracter chegue a ela, retira-o do anel (desde que este indique estar livre), transmite sua informação e o recoloca de volta ao anel de três maneiras diferentes: *single packet*, *single token* e *multiple token*.

  No *single packet* a estação devolve o caracter após o retorno completo da mensagem. A estação também pode recolocar o caracter de controle no anel após o retorno do mesmo imediatamente após a chegada do primeiro bit da mensagem enviada (*single token*). A estação pode, ainda, recolocar o caracter de controle após a transmissão (*multiple token*). O destinatário reconhece seu endereço e copia a informação no seu *buffer* receptor.

  Este protocolo não permite a utilização concorrente do anel por mais de uma estação, mas permite que mensagens de tamanho variável sejam transmitidas.

- Anel de Segmentos (*Slotted Ring*) Um número fixo de segmentos de mesmo tamanho são criados e utilizados para o transporte de pacotes. Cada segmento possui um bit indicador do seu estado: vazio ou cheio.

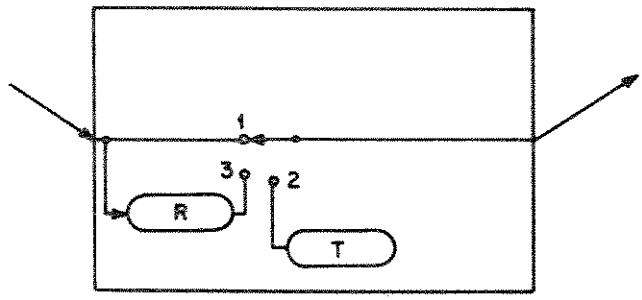

Figura 2.6: Registrador com Inserção de Retardo.

Qualquer estação com uma mensagem para ser transmitida divide-a em diversos pacotes do tamanho dos segmentos de transporte e ocupa o primeiro segmento vazio, alterando o seu bit indicador para cheio. Quando receber o segmento novamente, altera o bit indicador para vazio.

A determinação dos segmentos a serem utilizados pode ser feita por multiplexação no tempo (TDM), estatística ou por demanda. Com este protocolo, lacunas são criadas devido ao tempo de espera pelo início do segmento, mesmo que existam mensagens prontas para serem transmitidas nas estações.

- Inserção de Retardo (*Register-Insertion Ring*)

  Este protocolo se utiliza de um mecanismo de chaveamento que permite a conexão alternativa de dois registradores ao anel como está representado na Figura 2.6.

  Enquanto não há interesse em transmitir, a chave fica na posição 1 e com isto o registrador de transmissão(T) fica desconectado do circuito. Quando é necessário transmitir, uma mensagem é colocada no registrador T e a chave na posição 2, o que faz com que a mensagem seja transmitida.

  Durante o período de transmissão, os dados que forem recebidos serão armazenados no registrador de recepção (R). Quando a transmissão é completada, a chave é colocada na posição 3, os dados são colocados no anel com um retardo de uma mensagem. A chave permanece na

posição 3 até que a mensagem transmitida chegue ao registrador R. Neste momento, a chave retorna para a posição 1, o que faz com que a mensagem seja retirada do anel e, a partir de então, a interface está pronta para transmitir outra mensagem.

Este é um bom protocolo para cargas altas, porém sua confiabilidade é comprometida pelo atraso que cada estação adiciona ao enviar mensagens. O custo da tecnologia necessária para garantir a confiabilidade desejada é alto, desestimulando sua aplicação.

2.5.2 Redes em Barramento

O controle pode ser centralizado ou descentralizado e os protocolos podem pertencer a duas classes: uma que não permite colisão entre as mensagens transmitidas (Classe 1) e outra que permite (Classe 2). Dentre os protocolos de controle centralizado, destacaremos o *Polling*, da Classe 1 e dentre os protocolos descentralizados descreveremos os protocolos de Passagem de Permissão (Classe 1) e CSMA-CD (Classe 2).

- *Polling*

  Um controlador (que executa o *polling*) designa um a um a estação e o momento em que esta deve transmitir. Quando uma estação recebe autorização e tem uma mensagem para transmitir, transmite-a. Caso contrário, ele envia uma mensagem ao controlador para que este saiba que a estação em questão está ativa.

  Inconvenientes:

  - Existência de uma estação ativa responsável pelo controle, o que diminui a confiabilidade. No caso de falha desta estação, todas as transmissões ficam interrompidas. Algumas implementações utilizam um mecanismo que, ao perceber a falha do controlador, as estações nomeiam um novo controlador.
  - *Overhead* causado pelas mensagens necessárias à execução do *polling*.

- Passagem de Permissão (Padrão IEEE 802.4 - *Token Bus*)

  Existe uma permissão para acesso que circula entre as diversas estações de serviço. Uma estação pronta para transmitir e que recebe a permissão de acesso ao canal, pode enviar suas mensagens até não ter mais nenhuma pronta (serviço exaustivo) ou até um limite previamente estabelecido (serviço não exaustivo).

  Ao terminar a transmissão, a estação endereça a permissão de acesso à próxima estação que terá, então, a oportunidade de transmitir. É importante notar que, contrastando com um sistema de passagem de permissão em anel, a permissão de acesso no barramento deve ser endereçada, pois a ordem das estações neste sistema não precisa ser necessariamente a ordem das ligações físicas das estações ao barramento, sendo que esta ordem estabelece um ciclo lógico de acesso.

  item CSMA-CD (Padrão IEEE-802.3)

  É o mais conhecido esquema de acesso aleatório para sistemas em barramento. Sob este protocolo, cada estação monitora o canal antes de transmitir um quadro de informação para detectar tentativas de transmissão já em progresso. Se for detectada outra tentativa de transmissão, a estação aguardará até o final da transmissão em curso.

  Devido ao atraso de propagação no barramento ser não-nulo, duas ou mais estações podem tentar transmitir simultaneamente, mas os sinais interferirão uns nos outros, o que caracteriza uma colisão que será detectada por todas as estações. Quando uma colisão for detectada, a tentativa de transmissão será interrompida e a estação programará seu pacote para retransmissão, determinando um intervalo aleatório para retransmissão.

2.6 Conclusões

As redes em barramento são consideradas melhores devido à sua capacidade natural de radiodifusão e sua inerente confiabilidade (ligações passivas) que são características bastante desejáveis em redes locais. São também, mais simples de projetar.

Capítulo 3

M-LAN – Análise de Desempenho

3.1 Introdução

No Capítulo 2, foi feita uma revisão das técnicas de múltiplo acesso ao canal mais utilizados em redes locais. Apresentamos, de forma sintetizada, protocolos de múltiplo acesso, topologias e combinações de protocolos com topologias adequadas a redes locais, dentre os quais, o protocolo CSMA-CD mostrou-se o mais adequado para nossa aplicação, devido à sua inerente capacidade de radiodifusão com alta confiabilidade e tolerância a falhas.

A eficiência de protocolos tipo CSMA depende da razão entre o atraso de propagação fim-a-fim, que é proporcional à distância, e o tempo de transmissão do pacote .

A estrutura da primeira rede local construída foi em barramento passivo. Esta, porém, não acomoda eficientemente altas taxas de informação (da ordem de dezenas ou centenas de Mbps), grande número de usuários (alguns milhares) e longas distâncias (alguns kilômetros).

Outras estruturas surgiram satisfazendo as características acima, em detrimento de alguma outra. Na topologia em anel, por exemplo, a confiabilidade fica comprometida, pois longas distâncias exigem a utilização de vários repetidores ativos. Logo, quando a confiabilidade é importante na aplicação, arquiteturas passivas do tipo barramento são as mais recomendadas.

Algumas redes locais utilizando barramento unidirecional, como Fasnet

e Expressnet, são capazes de acomodar altas taxas de dados com transmissões por um único canal. Mas, com o uso de M canais em paralelo, pode-se ter as vantagens da faixa larga enquanto a velocidade dos canais é moderada. Uma grande vantagem das M-LANs é não requerer interfaces complexas, implicando na simplicidade de implantação além de diminuir os custos das instalações.

A arquitetura M-LAN, além de completamente modular (permitindo fácil manutenção e reparos), possui as características desejáveis como confiabilidade e tolerância a falhas.

Nosso objetivo, neste estudo, é analisar o desempenho de uma rede local multicanal com os protocolos de acesso tipo M-CSMA [6] com tráfego misto de pacotes de dados e de voz, por exemplo.

Da análise dos protocolos M-CSMA, observou-se uma vulnerabilidade na recepção de pacotes devido aos M canais em paralelo enviarem mensagens simultaneamente, podendo ocasionar colisões no receptor. Marsan e Roffinella [6], incluiram em seu trabalho, o cálculo da vazão para o protocolo M-CSMA-IC considerando colisões no receptor, denotado por M-CSMA-IC-RC[1] . Nosso interesse está na obtenção de expressões para o atraso desses pacotes.

A inclusão de filas nos receptores das estações, poderia evitar que pacotes que já se utilizaram do meio e, foram transmitidos com sucesso, necessitem ser retransmitidos devido à incapacidade de tratamento e armazenamento de todas as mensagens recebidas por um único receptor.

Como será considerado tráfego misto no modelamento, torna-se necessário considerar características individuais de cada tipo de tráfego. Desta forma, uma política que pode ser adotada é a de priorizar pacotes de voz, devido à exigência de transmissões em tempo real, em detrimento dos de dados.

Os pacotes de voz devem, então, assumir as primeiras posições da fila em ordem de chegada. Na prática, os pacotes de voz não podem ser retidos por muito tempo e devem, portanto, ser descartados e retransmitidos após um tempo aceitável de atraso. Esta situação, porém, não foi considerada nesta primeira abordagem. Assume-se que os pacotes de voz, que ao chegarem à fila, adquirem uma posição de forma que o tempo de retenção seja superior ao aceitável, são descartados antes de entrarem na fila.

[1] M-CSMA-IC-RC: M-CSMA-IC considerando colisões no receptor

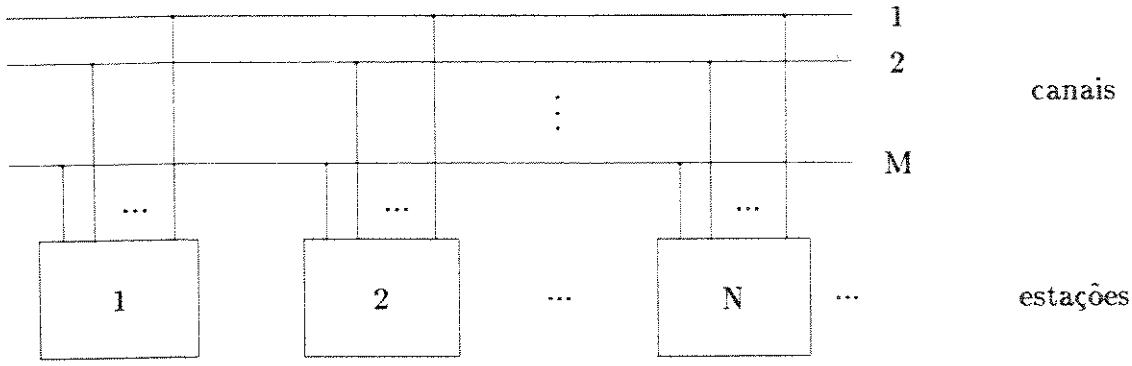

Figura 3.1: Arquitetura de uma M-LAN.

A partir de agora, desenvolveremos as expressões para a vazão e o atraso propostas em [6]. Para um maior esclarecimento, iniciaremos com o protocolo CSMA não-persistente que formou a base para os demais protocolos, para em seguida, abordarmos o estudo do atraso do pacote com colisões no receptor comparado com o atraso livre de colisões no receptor, em função da fila introduzida.

3.2 Rede local Multicanal

Uma rede local multicanal, ou simplesmente, M-LAN (Multichannel Local Area Network) é constituida de M canais do tipo radiodifusão em paralelo aos quais N estações estão conectadas através de M interfaces diferentes sendo uma para cada canal (ver Figura 3.1).

A faixa total disponível, W, é dividida em M subfaixas (W_i, com $i = 1, 2, \ldots, M$ delimitando os canais, tal que cada W_i pode assumir valores distintos. Mas por uma questão de simplicidade assumiremos que os W_i assumem valores iguais, tal que

$$W = \sum_{i=1}^{M} W_i = MW_i \tag{3.1}$$

Como iremos considerar que os pacotes têm o mesmo tamanho, o tempo

de transmissão pelo *i-ésimo* canal, T_i, satisfaz $T_iW_i = 1$ (transmissão digital).

Seja T_0, o tempo de transmissão de um pacote por um único canal de faixa W, então para $i = 1, 2, \ldots, M$:

$$T_i = MT_0$$

Seja A o atraso de propagação no meio, então o atraso normalizado, a_0, em relação a T_0 será $a_0 = A/T_0$. Por conseguinte, o atraso normalizado em relação a T_i será $a_i = A/T_i$. A partir destes resultados podemos obter a seguinte relação:

$$a_i = a_0 \frac{T_0}{T_i} = \frac{a_0}{M}$$

Por outro lado, considere a hipótese de população infinita no modelamento. Como consequência, teremos que o processo estocástico que melhor caracteriza o processo de chegada é o processo de Poisson com taxa λ pacotes por segundo.

Em condições de equilíbrio, a vazão normalizada em relação a T_0 é representada por $S = \lambda T_0$.

Seja λ_i o número de pacotes/segundo transmitidos com sucesso pelo *i-ésimo* canal, então a vazão do *i-ésimo* canal, S_i, será: $S_i = \lambda_i T_i$.

Como $\lambda = \sum_{i=1}^{M} \lambda_i$, podemos expressar a vazão total do sistema por:

$$S = T_0 \sum_{i=1}^{M} \frac{S_i}{T_i} \tag{3.2}$$

O tráfego total oferecido ao sistema, G, incluindo pacotes novos e reprogramados, é considerado poissoniano com taxa γ pacotes/seg. A expressão para o tráfego total normalizado em relação a T_0 é dada por $G = \gamma T_0$.

O tráfego oferecido a cada canal, G_i, tem seu comportamento definido pela política de escolha do canal. Se a escolha do canal para transmissão é feita de maneira aleatória, G_i tem distribuição poissoniana com taxa γP_i, onde P_i é a probabilidade de se escolher o canal i. Caso a escolha se dê dentre os canais previamente detectados livres, a distribuição de G_i, obviamente, não é poissoniana. Porém, esta suposição simplifica em muito o modelamento, pois G_i pode ser avaliado independentemente do tráfego em qualquer outro canal utilizando-se da expressão para um único canal:

$$G_i = \gamma P_i T_i$$

Como $\gamma = \frac{G}{T_0}$ e $\frac{T_i}{T_0} = M = \frac{W}{W_i}$, podemos expressar G_i por:

$$G_i = GP_i\frac{W}{W_i} \tag{3.3}$$

3.3 Análise da vazão

Nesta seção, será apresentado o desenvolvimento analítico da vazão segundo os protocolos de acesso M-CSMA, propostos em [6]. Para facilitar o entendimento, começaremos desenvolvendo o raciocínio usado por Kleinrock e Tobagi, em [4], na derivação da expressão da vazão em um único canal considerando o protocolo CSMA, uma vez que, o protocolo M-CSMA é uma extensão do protocolo CSMA.

3.3.1 Protocolo CSMA não-persistente

O protocolo CSMA foi proposto, inicialmente, para controlar conflitos na transmissão de pacotes através do acesso aleatório a um único canal de faixa larga por radiodifusão [4]. O controle baseia-se na monitoração do canal, previamente à transmissão, na tentativa de evitar colisões ao se tentar transmitir sobre uma transmissão já em andamento, destruindo ambas.

No caso de alguma estação não detectar a presença de portadora no canal, este é considerado livre e a transmissão começa imediatamente. Caso contrário, o canal é considerado ocupado com outra transmissão e a estação deve reprogramar para retransmitir após um atraso aleatório (situação de bloqueio). Devido ao atraso de propagação ser não-nulo, pacotes podem colidir. Da mesma forma que os bloqueados, os pacotes colididos são retransmitidos após um atraso aleatório.

A aceitação deste protocolo se verificou pela quantidade de aplicações e pelo número de trabalhos apresentados, posteriormente a sua publicação, contribuindo, ainda mais, na melhora de seu desempenho.

A seguir, iremos enumerar algumas características do sistema:

1. O acesso ao canal é feito aleatoriamente, de forma que o tráfego oferecido ao canal é poissoniano.

2. As estações monitoram o canal previamente à transmissão, na tentativa de evitar colisões através da detecção de portadora.

3. As estações não podem transmitir e receber pacotes de mensagens, simultaneamente.

4. Existe a necessidade de um sinal de reconhecimento (ACK), no sentido inverso ao do pacote, para informar o recebimento correto do mesmo.

5. As mensagens de reconhecimento são transmitidas por um canal reservado para este fim.

6. Há a necessidade de um mecanismo de retransmissão para os casos de colisão ocasionados devido ao atraso de propagação ser não-nulo.

7. Não se considera o ruído do canal.

8. Não são considerados o tempo de comutação do modo de transmissão para o de recepção e o tempo para detectar a portadora de alguma transmissão em curso.

9. Não se considera o efeito captura. Pacotes colididos são inutilizados.

10. O tamanho dos pacotes é fixo.

11. O atraso de propagação do meio é o mesmo para todos os pares fonte-destino.

12. O atraso de propagação é pequeno em relação ao tempo de transmissão do pacote.

13. A distinção entre os vários tipos de protocolo CSMA está na forma de retransmissão de pacotes colididos e daqueles que são bloqueados ao tentar fazer acesso ao canal.

Uma estação, seguindo o protocolo CSMA não-persistente, minimiza a interferência entre os pacotes, programando suas retransmissões para um certo tempo aleatório mais tarde.

Para o modelamento do tráfego, foram feitas as seguintes considerações:

- A fonte de tráfego é poissoniana.
- Infinitos usuários.
- Pacotes de tamanho fixo.
- Número médio de pacotes gerados por tempo de transmissão:

$$S = \lambda T$$

- Sob condições estacionárias, S é também a vazão.
- S também indica a utilização do canal, ou seja, $S = 1$ indica a utilização completa do canal. Porém, devido aos problemas de interferência inerentes aos modos de acesso aleatório, $S < 1$.
- Envio de mensagens de reconhecimento (ACK) necessário.
- O tráfego total oferecido inclui, além dos pacotes recém gerados, os programados para retransmissão devido a colisões e bloqueios, aumentando, desta forma, a taxa de tráfego oferecido por tempo de transmissão.

A seguir, listaremos as hipóteses assumidas em [4]:

- H1) O atraso médio de retransmissão ($\overline{X}$) é grande comparado com o tempo de transmissão de um pacote.
- H2) Os tempos entre as chegadas de pacotes (novos e programados para retransmissão) são independentes e exponencialmente distribuídos.
- H3) Os pacotes ACK são recebidos corretamente com probabilidade 1.
- H4) O tempo de processamento para realizar a verificação de paridade (*sumcheck*) e gerar o pacote ACK é desprezível.

Variáveis do Sistema:

- Taxa média de geração de pacotes: λ pacotes/segundo.
- Tempo de transmissão de um pacote: T segundos.
- Vazão do canal: $S = \lambda T$ pacotes por tempo de transmissão.
- Taxa de tráfego total oferecido: G pacotes por tempo de transmissão.
- Capacidade do canal: $C = Max_G\{S\}$.
- Atraso médio para retransmissão: $\overline{X}$ segundos.
- Atraso de propagação fim-a-fim: τ segundos.
- Atrasos normalizados em relação ao tempo de transmissão (T):

$$\begin{array}{lll} \text{Atraso de Retransmissão} & : & \delta = \overline{X}/T \\ \text{Atraso de Propagação} & : & a = \tau/T \end{array}$$

ANÁLISE DA VAZÃO:

Na Figura 3.2, considere o eixo do tempo sob o ponto de vista do transmissor. Ao monitorar o canal, o transmissor com pacotes prontos para a transmissão pode detectar o canal em 2 estados: livre ou ocupado. Porém, devido ao fato do atraso de propagação ser não-nulo, o transmissor pode detectar um canal livre enquanto, na verdade, ele está ocupado. Neste caso, ambos colidirão e deverão retransmitir.

Seja t, o momento em que uma estação, com algum pacote pronto para transmissão, monitora o canal e não detecta a presença de portadora (ou seja, o canal está livre, aparentemente) e seja a, o atraso de propagação normalizado em relação a T ($a = \tau/T$).

Se não houver nenhuma outra transmissão durante os próximos a segundos (entre t e $t + a$) significa que o pacote não sofreu colisão no canal. Qualquer outro pacote chegando ao canal durante o período vulnerável (entre t e $t + a$), não detectará a portadora e, apesar da primeira transmissão estar em andamento, transmitirá ocasionando uma colisão.

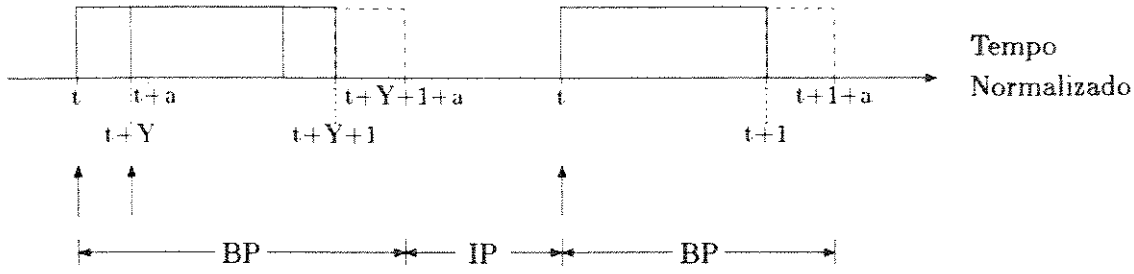

Figura 3.2: CSMA não-persistente: períodos livres, IPs, e ocupados, BPs.

Seja $t+Y$, o momento em que o último pacote chega ao canal durante o período vulnerável. A transmissão deste último pacote só será completada em $t+Y+1$ e, apenas a segundos mais tarde, o canal poderá ser detectado livre por todas as outras estações. Neste interim, qualquer estação que monitorar o canal, detectará a presença de outra transmissão em andamento e programará para retransmitir o pacote bloqueado após um certo tempo aleatório, X, segundo a distribuição de retransmissão.

Seja TP, o período de transmissão definido entre t e $t+Y+1+a$, podendo haver apenas uma transmissão com sucesso durante cada TP. É também definido como período ocupado, BP. Um período livre, IP, é definido como o intervalo entre dois TPs consecutivos.

A soma de um TP com um IP constitui um ciclo, cuja duração média é a soma da duração média dos períodos ocupados, $\overline{B}$, e livres, $\overline{I}$.

$$\overline{C} = \overline{B} + \overline{I}$$

Seja U, o tempo durante um ciclo em que o canal é usado sem conflitos. Seu valor médio, $\overline{U}$, indica a probabilidade de sucesso durante um certo período.

Segundo argumentos da Teoria de Renovação, a utilização média do canal é dada por:

$$S = \frac{\overline{U}}{\overline{C}} \tag{3.4}$$

Então, a probabilidade de um TP obter sucesso, é a probabilidade de que nenhum terminal (estação) transmita durante os primeiros a segundos do TP.

Como o tráfego, G, é considerado poissoniano com período vulnerável a, temos:

Probabilidade de sucesso: $\overline{U} = \epsilon^{-aG}$
Duração média dos períodos IP: $\overline{I} = \frac{1}{G}$
Duração média dos períodos BP: $\overline{B} = 1 + \overline{Y} + a$

Como consequência, obtemos a expressão para a duração média de um ciclo:

$$\overline{C} = \overline{I} + \overline{B} = \frac{1}{G} + 1 + \overline{Y} + a \tag{3.5}$$

Para encontrarmos $\overline{Y}$, utilizaremos a expressão para o valor esperado de uma variável aleatória contínua:

$$\overline{Y} \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} y f_Y(y)\, dy \tag{3.6}$$

Seja f_Y, a função densidade de probabilidade (pdf) dada por,

$$f_Y(y) \stackrel{\text{def}}{=} \frac{d}{dy} F_Y(y)$$

e, F_Y, a função distribuição de probabilidade (PDF) dada por,

$$F_Y(y) = P\{Y \leq y\}$$

Como a PDF representa a probabilidade de não ocorrerem chegadas no intervalo de comprimento $a - y$, então, para $y \leq a$:

$$F_Y(y) \stackrel{\text{def}}{=} \epsilon^{-G(a-y)}$$

Desta forma,

$$f_Y(y) = G\epsilon^{-G(a-y)} \tag{3.7}$$

Substituindo a eq. (3.7) na expressão para $\overline{Y}$, eq. (3.6), encontramos:

$$\overline{Y} = \int_0^a y\, G\, \epsilon^{-G(a-y)}\, dy = G\left[y\left(\frac{\epsilon^{-G(a-y)}}{G}\right)_0^a - \int_0^a \frac{\epsilon^{-G(a-y)}}{G}\, dy \right]$$

$$\overline{Y} = G\left[\frac{a}{G} - \frac{1}{G}\left(\frac{\epsilon^{-G(a-y)}}{G}\right)_0^a\right]$$

e, finalmente:

$$\overline{Y} = a - \frac{1}{G}\left(1 - \epsilon^{-aG}\right) \tag{3.8}$$

Podemos, agora, obter a expressão para o tempo médio dos BPs:

$$\overline{B} = 1 + \overline{Y} + a = 1 + 2a - \frac{1}{G}\left(1 - \epsilon^{-aG}\right) \tag{3.9}$$

Substituindo a eq. (3.8) na eq. (3.9), podemos encontrar a expressão para $\overline{C}$ em (3.5), ou seja,

$$\overline{C} = \overline{B} + \overline{I} = 1 + 2a - \frac{1}{G}(1 - \epsilon^{-aG}) + \frac{1}{G}$$

$$\overline{C} = \frac{G(1+2a) + \epsilon^{-aG}}{G}$$

e, consequentemente, encontrar a vazão pela eq. (3.4) cujo comportamento pode ser analisado pela Figura 3.3.

$$S = \frac{G\epsilon^{-aG}}{G(1+2a) + \epsilon^{-aG}} \tag{3.10}$$

3.3.2 Protocolos M-CSMA não-persistente

Os protocolos M-CSMA e M-CSMA-CD foram propostos com o objetivo de oferecer uma arquitetura de rede bastante confiável, como o barramento passivo, que proporcionasse uma alta capacidade de transmissão utilizando um conjunto de M barramentos de média a baixa taxa de transmissão, em paralelo, ao invés de um único canal com alta taxa de transmissão. Tais barramentos em paralelo não estão, necessariamente, separados fisicamente, mas podem ser obtidos pela divisão da faixa de frequência de uma única conexão física. Esta topologia oferece a confiabilidade característica de uma rede local com estrutura em barramento, enquanto proporciona a faixa larga que pode ser utilizada para transmissão de pacotes com a vantagem da tolerância a falhas.

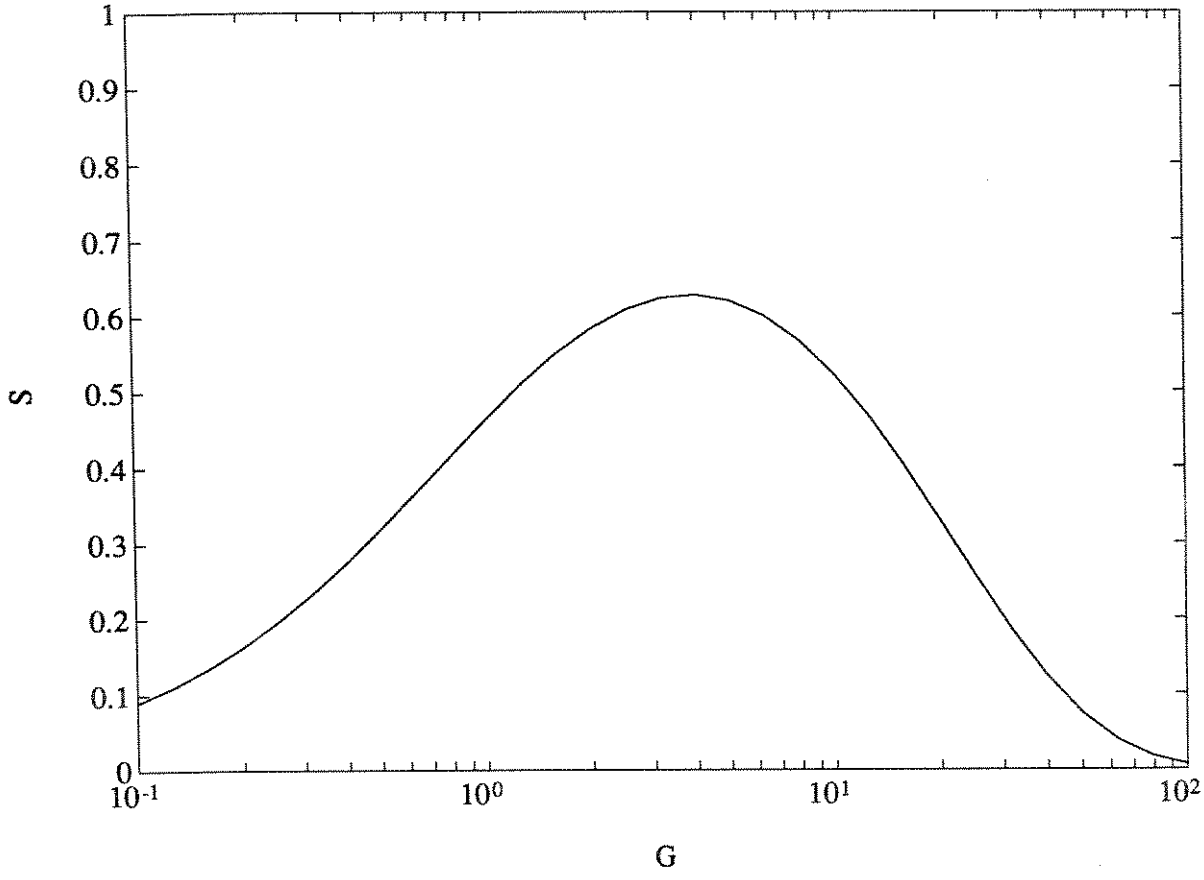

Figura 3.3: Vazão x Tráfego: CSMA, eq. (3.10), [4].

Para se utilizar de vários canais multiplexados na frequência, as estações devem estar equipadas com *modems* de alta velocidade e ágeis na frequência. Para conexões com poucos canais, uma alternativa seria a utilização de um conjunto de *modems* em paralelo, sintonizados em canais diferentes, não sendo uma solução economicamente atrativa para instalações com equipamentos terminais de baixo custo.

A extensão a sistemas de canais múltiplos implica na escolha, pela estação, do canal por onde o pacote será transmitido. Existem duas maneiras básicas para a escolha do canal: a primeira, considera a escolha aleatória do canal, previamente a monitoração (Protocolo M-CSMA-RC[2]) e, a segunda, escolhe um canal aleatoriamente dentre os monitorados e detectados livres (Protocolo M-CSMA-IC[3]).

Análise do desempenho destes protocolos mostraram-se melhores que a do CSMA para um único canal, pois, ao usar M canais em paralelo, obtem-se a possibilidade de utilização da faixa larga, enquanto se mantém uma alta razão entre o tempo de transmissão e o atraso de propagação.

A seguir, iremos enumerar algumas das características deste sistema de redes locais multicanais:

1. Apesar do acesso aos canais depender da política adotada para a escolha dos canais, podendo ser aleatória ou não, assume-se a distribuição poissoniana para o tráfego oferecido ao canal.

2. As estações são capazes de detectar a presença de portadora em cada um dos M canais, simultaneamente, permitindo o uso de protocolos tipo CSMA.

3. Cada estação é capaz de transmitir por um canal e receber por vários outros canais, simultaneamente.

4. Não se considera o tráfego de pacotes de reconhecimento (ACK).

5. Não se discute os efeitos das falhas na rede, ou seja, considera-se que os M canais estão disponíveis o tempo todo.

6. Cada interface é capaz de detectar colisões no canal em que estiver transmitindo, de forma que permite o uso do protocolo CSMA-CD para transmissão de pacotes em M-LANs.

[2] *M-CSMA with Random Choice.*
[3] *M-CSMA with Idle Choice.*

São válidas, também, as considerações feitas em [4] quanto: ao sinal ACK, mecanismos de retransmissão, ruídos no canal, tempos de comutação e detecção. Além dos efeitos de captura, do tamanho do pacote e atraso de propagação.

O Protocolo M-CSMA-RC

- O canal i é escolhido aleatoriamente com probabilidade P_i.
- O tráfego oferecido ao *i-ésimo* canal segue a distribuição de Poisson com taxa G_i pacotes/segundo, onde G_i é dado pela equação (3.3).
- A vazão do canal i, S_i, pode ser obtida a partir da expressão para a vazão no CSMA, equação (3.10) , substituindo G_i em G e a_i em a.

$$S_i = \frac{G_i \epsilon^{-a_i G_i}}{G_i(1+2a_i) + \epsilon^{-a_i G_i}}$$

Substituindo (3.3) em S_i, obtemos:

$$S_i = \frac{GP_iW/W_i \epsilon^{-a_i GP_i W/W_i}}{GP_iW(1+2a_i)/W_i + \epsilon^{-a_i GP_i W/W_i}}$$

e substituindo S_i na equação (3.2), temos:

$$S = T_0 \sum_{i=1}^{m} \frac{GP_iW/W_i \epsilon^{-a_i GP_i W/W_i}}{GP_iW(1+2a_i)/W_i + \epsilon^{-a_i GP_i W/W_i}} \frac{1}{T_i}$$

Como $\frac{W}{W_i} = \frac{T_i}{T_0}$, a expressão para a vazão segundo o protocolo M-CSMA-RC é:

$$S = \sum_{i=1}^{m} \frac{P_i \epsilon^{-a_i GWP_i/W_i}}{GWP_i(1+2a_i)/W_i + \epsilon^{-a_i GWP_i/W_i}} \qquad (3.11)$$

- O valor de P_i depende de G.[4] Para valores baixos de G, $P_{iotimo} \cong \frac{W_i}{W}$, enquanto que para valores altos de G, $P_{iotimo} \sim$ Uniforme.

[4] Exceto se $M = 2$, com $W_1 = W_2 = \frac{W}{2}$, $P_1 = P_2 = \frac{1}{2}$.

- Para M canais iguais, $W_i = \frac{W}{M}$, assim como $T_i = \mathbf{M}T_0$ e $a_i = a_0/\mathbf{M}$. Então, se $P_i = 1/M$, a vazão para o protocolo M-CSMA-RC (ver Figura 3.4) fica:

$$S = \frac{G\epsilon^{-a_0 G/M}}{G(1 + 2a_0/M) + \epsilon^{-a_0 G/M}} \tag{3.12}$$

que é a expressão (3.10) com faixa W e $a = \frac{a_0}{M}$.

- Com M canais em paralelo, a vazão e a capacidade são aumentadas por uma quantia equivalente à redução do atraso de propagação pelo número de canais.

O Protocolo M-CSMA-IC

Este protocolo foi desenvolvido para melhorar o desempenho obtido com o protocolo M-CSMA-RC. Resolveu-se restringir o conjunto de canais passíveis de serem escolhidos para transmissão, eliminando aqueles que já estivessem ocupados no momento em que alguma estação, com algum pacote pronto para transmissão, escolher um canal de forma aleatória. Assim, os pacotes só são programados para retransmissão quando todos os canais estiverem ocupados.

Como se considera que os M canais têm a mesma amplitude de faixa, W_i, a probabilidade de se escolher um canal dentre os M detectados livres é $1/M$.

Uma análise precisa do desempenho do protocolo seria bastante complexa de se obter. Diante desta dificuldade, assume-se que o tráfego oferecido aos M canais é poissoniano.

Estimativas para o desempenho foram obtidas através de duas técnicas, e seus resultados, verificados por simulação. A primeira, considerou os canais independentes entre si e, pela simulação, concluiu-se que quando os canais são considerados independentes e, portanto, o tráfego oferecido ao *i-ésimo* canal é poissoniano, a aproximação é muito boa para $M = 2$ canais. Porém, para valores maiores, os resultados são muito otimistas.

Assim sendo, um outro modelo, utilizado para a avaliação da vazão em sistemas com um grande número de canais disponíveis foi apresentado em [6].

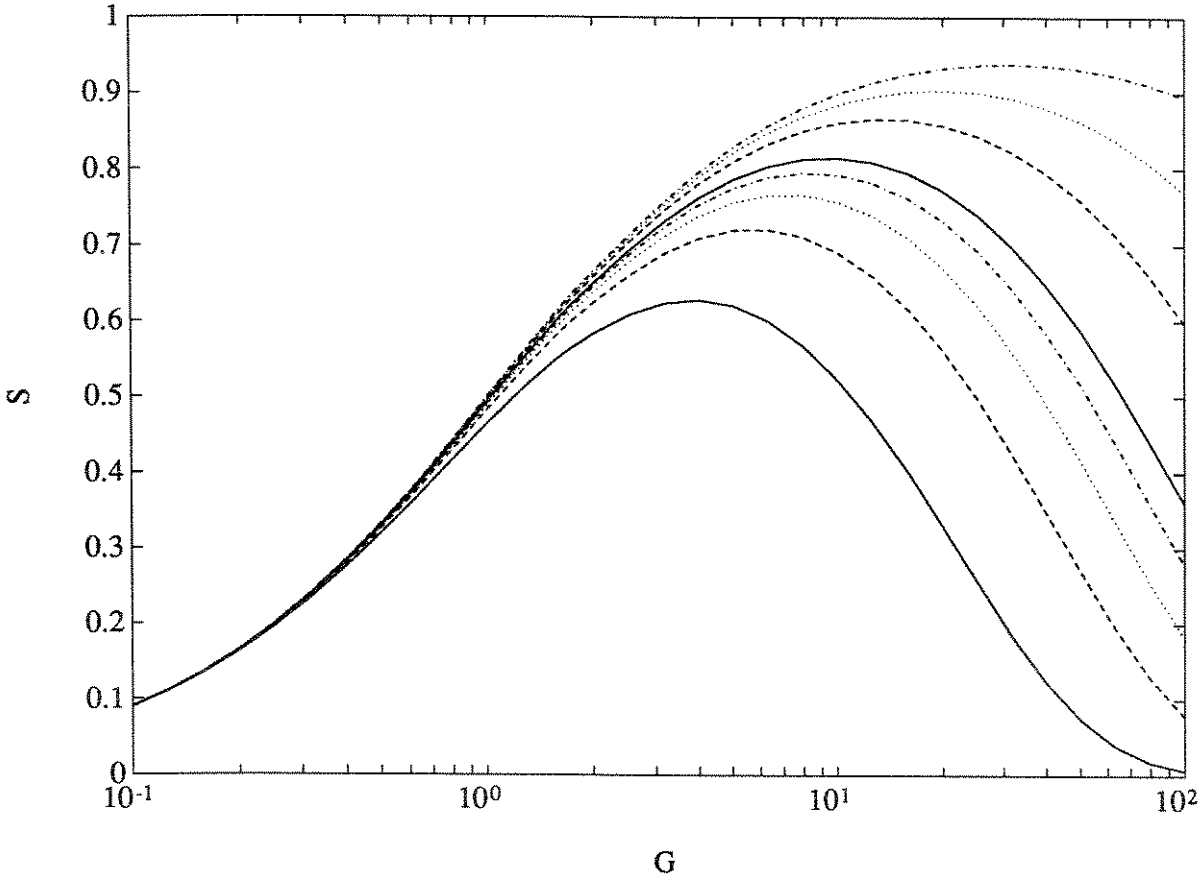

Figura 3.4: Vazão x Tráfego: M-CSMA-RC com $M = 1, 2, 3, 5, 10, 20, 50$, eq. (3.12),[6].

Na primeira estimatitiva, ao se considerar a independência entre os estados dos canais, chega-se a uma expressão para o tráfego oferecido ao *i-ésimo* canal através da expressão para um único canal, eq. (3.3), ou seja,

$$G_i = GT_i/T_0 P(\text{escolher i / i detectado livre})$$

Condicionando no número médio de canais livres, obtem-se:

$$G_i = GM \sum_{j=0}^{M-1} \frac{1}{j+1} C_{(M-1,j)} (1 - Pli)^{M-1-j} Pli^j$$

que é equivalente a:

$$G_i = G \sum_{j=1}^{M} C_{(M,j)} \frac{(1 - Pli)^{M-j} Pli^j}{Pli}$$

ou ainda,

$$G_i = G \frac{1 - (1 - Pli)^M}{Pli}$$

onde Pli, a probabilidade do canal escolhido estar livre, é dada por:

$$Pli = \frac{1 + a_i G_i}{G_i(2a_i + 1) + \epsilon^{-a_i G_i}}$$

A expressão para a vazão é obtida pela substituição de G_i em G na eq. (3.12), cujo comportamento pode ser observado na Figura 3.5).

$$S = S_i = \frac{G_i \epsilon^{-a_0 G_i/M}}{G_i(2a_0/M + 1) + \epsilon^{-a_0 G_i/M}} \tag{3.13}$$

Na segunda estimativa, consideram-se três estados possíveis para os canais: ocupado (B), livre de fato (I) e detectado livre quando está ocupado (V)[5].

Seja s', o conjunto de todos os estados possíveis no sistema por radiodifusão, cujos elementos, s_j, são *M-uplas* representando os estados dos M canais. Cada um dos elementos s_j possui i canais no estado I e v canais

[5]B, I e V vêm do Inglês: *Busy*, *Idle* e *Vulnerable*, respectivamente.

BIBLIOTECA CENTRAL

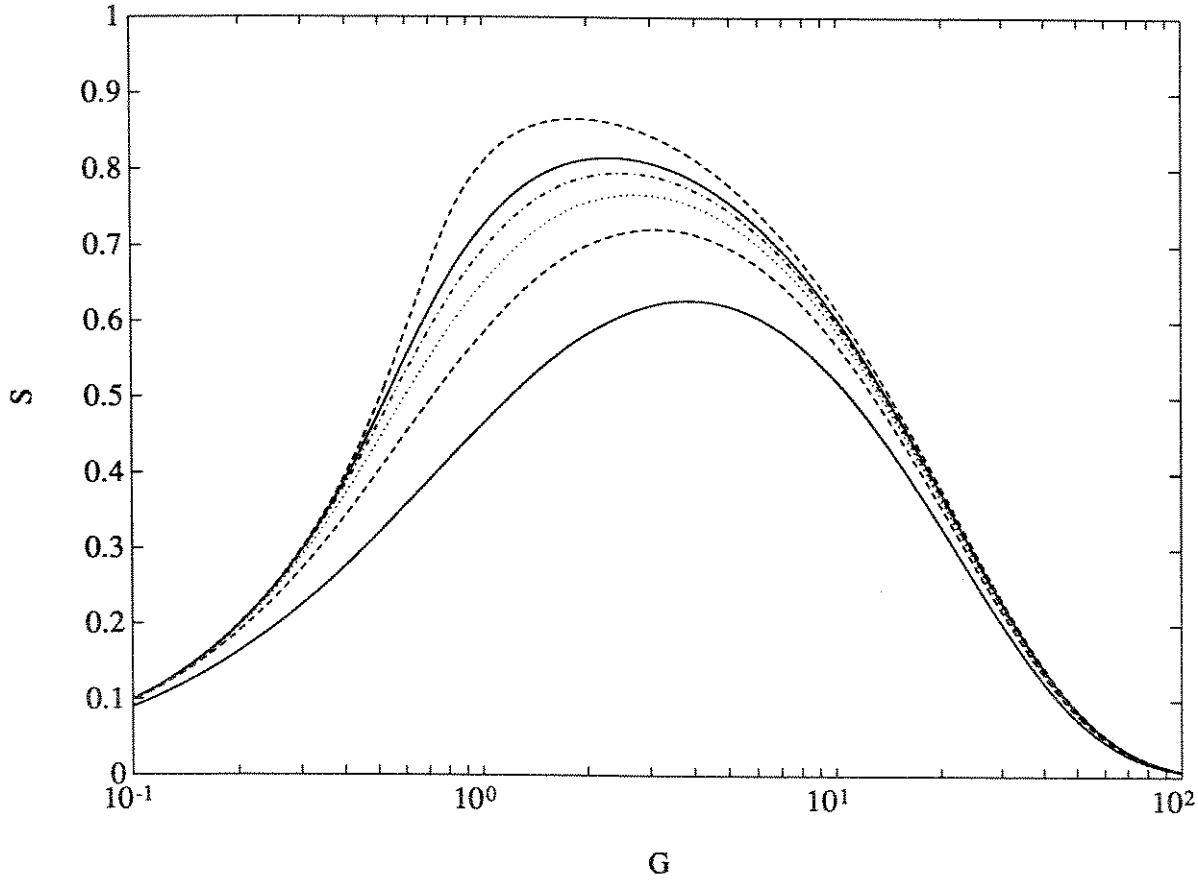

Figura 3.5: Vazão x Tráfego: M-CSMA-IC, com $M = 1, 2, 3, 5, 10$, eq. (3.13), [6].

no estado V. A vazão total pode ser avaliada como $S = GPs$, onde Ps, a probabilidade de um pacote ser transmitido com sucesso, é:

$$Ps = \sum_{s_j \in S'} \sum_{k=1}^{M} P\{\text{sucesso no } k\text{-}\acute{e}simo \text{ canal}/s_j\} P\{k/s_j\} P\{s_j\}$$

Considerando que:

- $P\{s_j\} = P_{iv}$, a probabilidade de ter i canais no estado I e v canais em V.
- $P\{k/s_j\} = \frac{1}{i+v}$, a probabilidade de escolher o canal k livre dentre os detectados livres.
- $P\{\text{sucesso no } k\text{-}\acute{e}simo \text{ canal}/s_j\} = \epsilon^{-a_0 G/(i+v)}$, a probabilidade de não haver transmissões nos $i + v$ canais durante o atraso de propagação do i-*ésimo* canal (a_i).

Observe que no M-CSMA-RC, um canal é escolhido dentre M. Como $a_i = a_0 T_0/T_i$ e $T_0/T_i = 1/M$, então $a_i = a_0/M$. Analogamente, o M-CSMA-IC escolhe um canal dentre $(i + v)$ detectados livres, então $a_i = a_0/(i + v)$. Logo, a vazão total é:

$$S = GPs = G \sum_{i=0}^{M} \sum_{v=0}^{M-i} \epsilon^{-a_0 G/(i+v)} \frac{i}{i+v} P_{iv} \tag{3.14}$$

Para M grande, o atraso de propagação no i-*ésimo* canal (a_0/M) é pequeno comparado ao tempo de transmissão. Pode-se, então, aproximar o modelo do protocolo desprezando o período vulnerável (V) na avaliação de P_{iv}. Assim, o sistema por radiodifusão com M canais pode ser visto como um sistema de perdas com M servidores. Este resultado, bastante conhecido em telefonia, tem P_M representando a fração de tempo em que todos os M servidores estão ocupados. Trata-se da Fórmula da Perda (ou Fórmula B) de Erlang:

$$P_M = \frac{\left(\frac{\lambda}{\mu}\right)^M / M!}{\sum_{k=0}^{M} \left(\frac{\lambda}{\mu}\right)^k / k!}$$

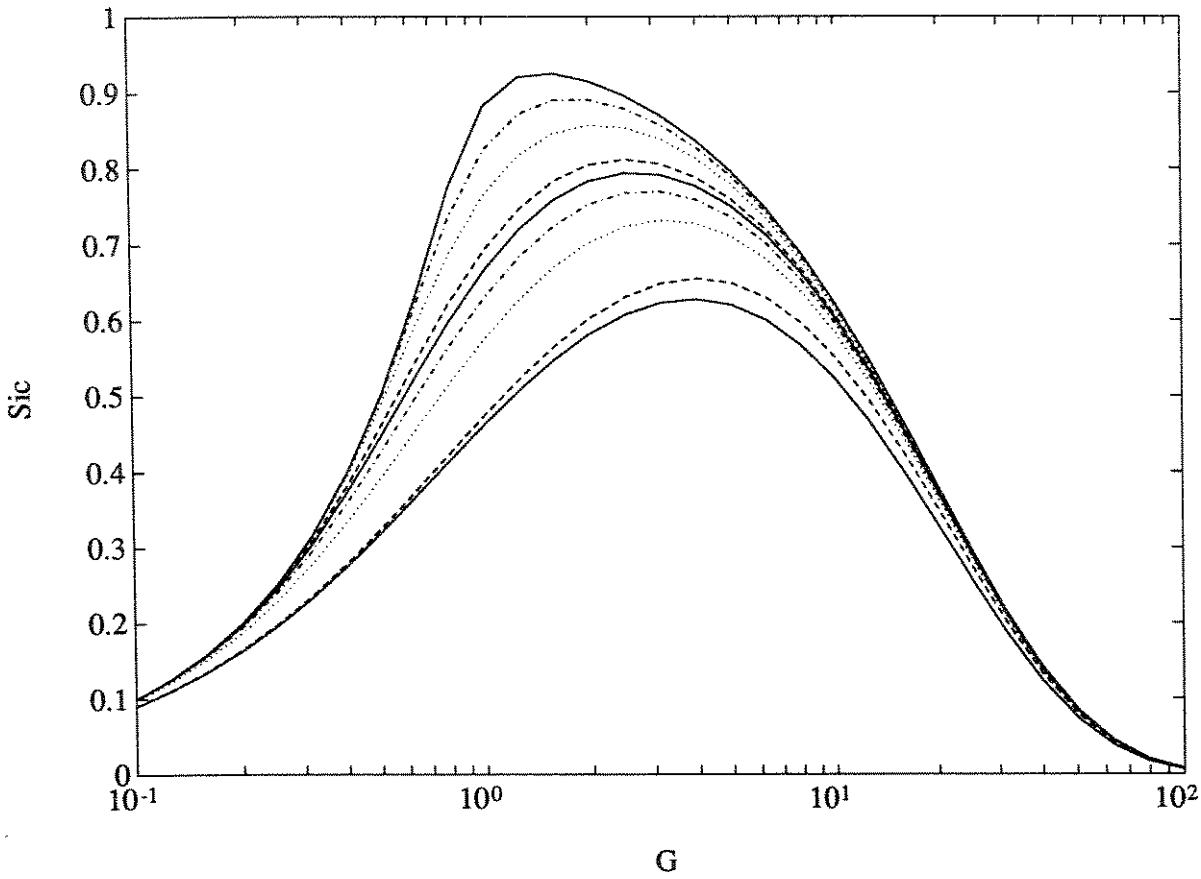

Figura 3.6: Vazão x Tráfego: CSMA e M-CSMA-IC com $M = 1, 2, 3, 4, 5, 10, 20, 50$, eq. (3.15), [6].

Portanto, a expressão para a probabilidade de se encontrar i canais livres, P_i, que é a probabilidade de se encontrar $M - i$ canais ocupados, P_{M-i}, pode ser obtida a partir da Fórmula da Perda de Erlang. Logo, a expressão para a vazão para o protocolo M-CSMA-IC considerando $v = 0$, ilustrada na Figura 3.6, é dada por:

$$S_{IC} = G \sum_{i=0}^{M} e^{-a_{ii} G/i} \frac{(GM)^{M-i}}{(M-i)!} \left(\frac{\sum_{k=0}^{M} (GM)^k}{k!} \right)^{-1} \qquad (3.15)$$

Resultados de simulações [6] mostram que a aproximação da expressão para a vazão é muito precisa para valores grandes de M ($M > 5$). Mas, leva a erros grandes para altas taxas de tráfego oferecido a um pequeno

número de canais, quando a probabilidade de um pacote ser gerado durante o período vulnerável é alta.

3.3.3 Protocolo M-CSMA-IC com colisões no receptor

Os protocolos M-CSMA permitem que as estações transmitam por um canal e, simultaneamente, recebam pacotes vindos de vários outros canais, podendo estar envolvidas em até M recepções ao mesmo tempo.

Os resultados para a vazão, obtidos até agora, assumiram que as estações eram capazes de tratar todos os pacotes que lá chegassem, de forma a perder apenas os pacotes que colidissem no canal.

Porém, para M grande esses resultados são muito otimistas. Uma multiplexação dos pacotes, vindos dos canais, pode ser necessária antes da distribuição ao usuário-destino. Neste caso, a presença de *buffers* compartilhados pode ser uma fonte de contenção entre pacotes e pode ocasionar perdas no receptor. Fenômeno denominado por "colisões no receptor".

Numa primeira abordagem, considerou-se que apenas n ($n < M$) receptores poderiam estar ativos numa mesma estação, num dado instante, ocasionando perdas. Considerou-se, ainda, que os pacotes colididos nos canais não acionariam os receptores.

Para se estimar a vazão com perdas devido às colisões no receptor, usou-se um modelo simplificado, com as seguintes características:

- Tráfego total oferecido poissoniano.
- Pacotes com estação fonte e destino atribuidas segundo a distruição uniforme.
- O número de estações com n receptores ativos quando i canais estão ocupados é denominado R_i, eq. (3.17).

A expressão para a vazão, considerando colisões no receptor, é a expressão para S_{IC} multiplicada pela probabilidade de não ocorrer bloqueio na recepção, ou seja,

$$S_{CR} = G \sum_{i=1}^{M} \epsilon^{-a_0 G/i} \frac{(GM)^{M-i}}{(M-i)!} \left(1 - \frac{R_i}{N}\right) \left(\sum_{k=0}^{M} \frac{(GM)^k}{k!}\right)^{-1} \tag{3.16}$$

Uma estimativa pessimista, considera R_i como sendo a parte inteira entre o número de canais ocupados, i, e o número de receptores por estação, n, ou seja, $R_i = \left[\frac{i}{n}\right]$. [6]

Uma estimativa mais precisa, considera R_i como sendo o número médio de estações com n receptores ativos quando i canais estão ocupados. Condicionando ao número de transmissões com sucesso, obteve-se:

$$R_i = \sum_{j=1}^{i} \sum_{k=1}^{N} k P_{kij} P'_{ij} \tag{3.17}$$

onde, P_{kij} é a probabilidade de k estações com n receptores ativos dados i canais ocupados e j transmissões com sucesso, e P'_{ij} é a probabilidade de j transmissões com sucesso dados i canais ocupados.

Para $n = 1$, devido à distribuição uniforme usada para escolher as estações destino, encontra-se:

$$P_{kij} = \frac{N! S(j,k)}{(N-k)! N^j} \quad ,$$

onde $S(j,k)$ são os Números de Stirling do Segundo Tipo [8], e representam o número de maneiras possíveis de se distribuir as j transmissões que obtiveram sucesso em k estações receptoras considerando que todas recebam pelo menos um pacote ($j \geq k$). Dessa forma, $S(j,k)$ são expressos por[7]:

$$S(j,k) = \frac{1}{k!} \sum_{r=0}^{k} (-1)^r C_{(k,k-r)} (k-r)^j \tag{3.18}$$

A análise do caso geral, onde n pacotes são tratados, simultaneamente, nos receptores não leva a uma expressão fechada para P_{kij}.

Quanto a P'_{ij}, pode ser estimado através da expressão:

$$P'_{ij} = C_{(i,j)} Pnc^j (1 - Pnc)^{i-j}$$

onde, Pnc é a probabilidade de um pacote não colidir, e é expressa por:

$$Pnc = \frac{S}{GPl}$$

[6] Neste caso, não se considera que colisões nos canais não ativam os receptores e, que alguns canais podem estar ocupados com pacotes que colidem no receptor.

[7] Note que, em termos genéricos, $C_{(n,k)}$ é a combinação de n elementos, k a k.

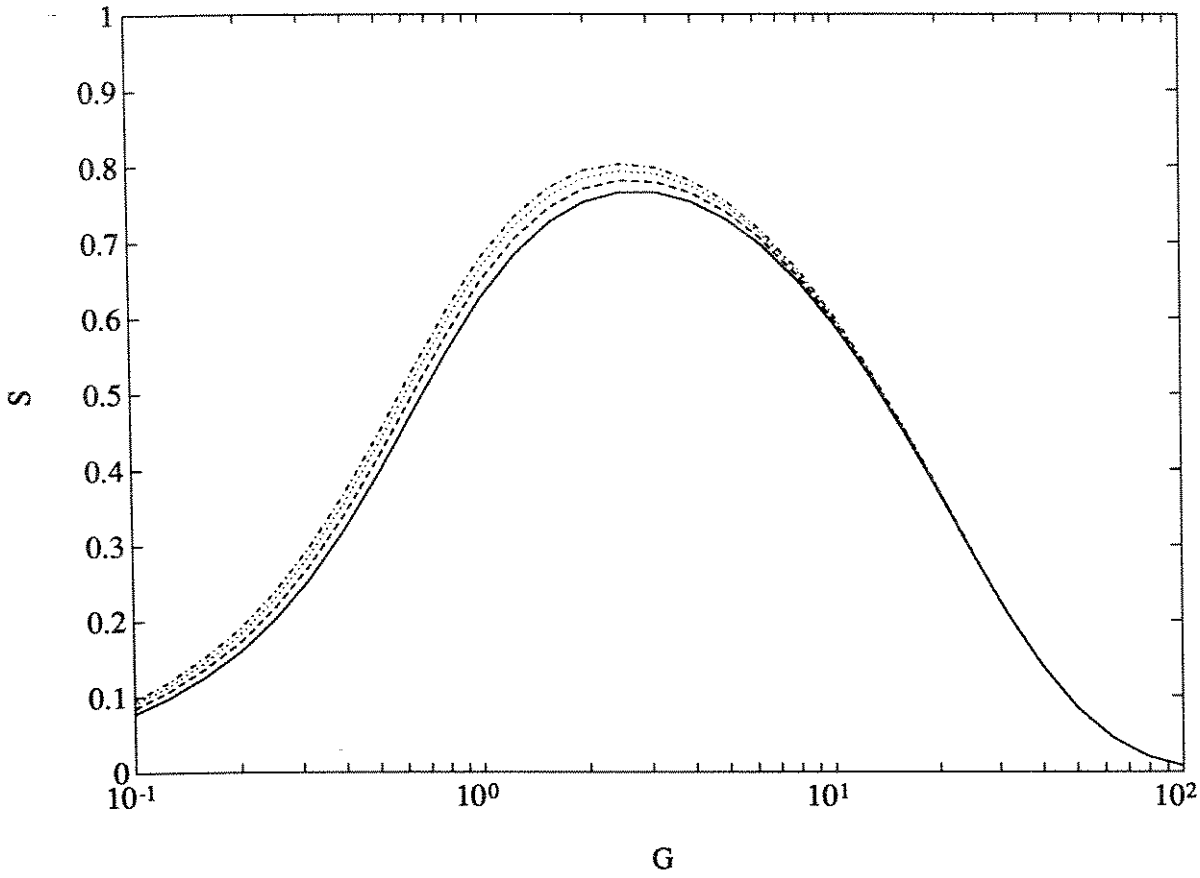

Figura 3.7: Vazão x Tráfego: CSMA e M-CSMA-IC-RC com $M = 5$ e $N = 20, 30, 50, 100$, eq.(3.16), [6].

e, Pl, a probabilidade de detectar pelo menos um canal livre, será dada pela eq. (3.23). A razão S/G pode ser obtida a partir de S_{IC}.

Os resultados obtidos, e validados através de simulação, mostram que, ao se aumentar o número de estações, obtem-se um aumento na vazão para um mesmo tráfego. Veja a Figura 3.7. Nota-se, no entanto, que para $N > 100$ não se observam aumentos significativos.

A Figura 3.8 mostra que a medida que o número de canais para uma dada população de usuários ($N = 100$) vai aumentando, não são observadas melhoras no desempenho como foi observado antes de se levar em consideração as colisões no receptor (Fig. 3.6).

Devido a uma proximidade maior deste modelamento aos sistemas reais,

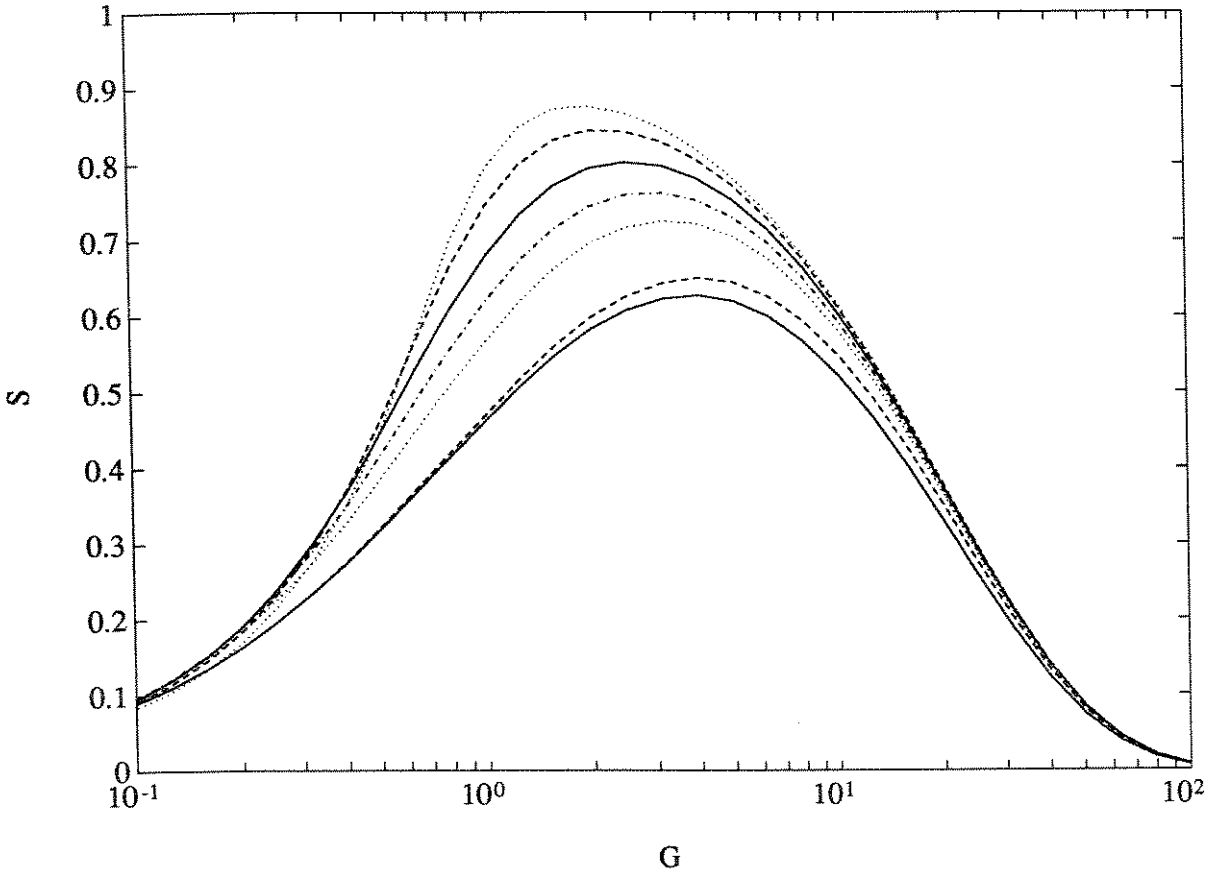

Figura 3.8: Vazão x Tráfego: CSMA e M-CSMA-IC-RC com $N = 100$ e $M = 1, 2, 3, 5, 10, 20$, eq.(3.16), [6].

sentimos a necessidade de aperfeiçoar o modelamento de maneira a evitar, na medida do possível, que colisões aconteçam nos receptores.

Com este intuito, introduziremos a idéia de adicionar uma fila em cada receptor com um *buffer* de armazenamento para os pacotes que chegarem enquanto a estação destino estiver ocupada com o tratamento de outro pacote. Não obteremos a vazão para este caso, mas é de se esperar que aumente, uma vez que os canais serão melhor utilizados. No entanto, o estudo avalia o atraso médio dos pacotes considerando a completa absorção dos pacotes pela fila na recepção comparado com o atraso médio dos pacotes podendo sofrer colisões no receptor.

3.3.4 Protocolos M-CSMA-CD não-persistentes

É sabido que o protocolo CSMA-CD oferece um melhor desempenho que o protocolo CSMA em redes locais, onde o atraso de propagação fim-a-fim é uma pequena fração do tempo de transmissão de um pacote. A extensão do protocolo CSMA-CD para um ambiente de rede multicanal (M-CSMA-CD) resulta numa melhora no desempenho ainda superior à obtida para o M-CSMA.

Este estudo, no entanto, se restringe à introdução de filas com prioridades para a transmissão e recepção de pacotes em estações com tráfego misto, segundo o protocolo M-CSMA, uma vez que não são conhecidos resultados para a vazão e o atraso do pacote considerando colisões nos receptores para o M-CSMA-CD.

As expressões para a vazão segundo os protocolos M-CSMA-CD-RC E M-CSMA-CD-IC podem ser encontradas em [6].

3.4 Análise do atraso do pacote

Nesta secão, desenvolveremos as expressões para o atraso médio de um pacote considerando, primeiramente, o protocolo CSMA e, em seguida, os protocolos M-CSMA-RC, M-CSMA-IC E M-CSMA-IC-RC.

3.4.1 CSMA

Seja D, o atraso médio do pacote definido em [4], como o tempo médio gasto a partir do momento em que um pacote é gerado até obter sucesso em

sua transmissão. Isto inclui, o atraso entre duas retransmissões sucessivas, R, para o caso do pacote ter sido bloqueado na entrada do canal e para o caso de ter sido transmitido e sofrido colisão no canal. Inclui, ainda, o tempo da transmissão que obteve sucesso, T, e seu atraso de propagação, τ.

O atraso médio para retransmissão pode ser representado por:

$$R = \begin{cases} T + 2\tau + Ta + \overline{X}, & \text{se houver colisão} \\ \overline{X}, & \text{se houver bloqueio} \end{cases}$$

Normalizando R em relação a T, temos:

$$R = \begin{cases} 1 + 2a + \alpha + \delta, & \text{se houver colisão} \\ \delta, & \text{se houver bloqueio} \end{cases}$$

onde, a, α e δ são o atraso de propagação (τ), o tempo de transmissão do pacote ACK (Ta) e o atraso de retransmissão ($\overline{X}$) normalizados em relação a T, respectivamente.

Seja Ps, a probabilidade de se obter sucesso na transmissão, então, $Ps = 1 - Pb$, onde $Pb = P\{bloqueio\}$. A não ocorrência de bloqueio durante um ciclo ocorre nos períodos livres e durante o atraso de propagação subsequente, então:

$$Ps = \frac{\overline{I} + a}{\overline{C}} = \frac{\frac{1}{G} + a}{\frac{1}{G} + 1 + a + \overline{Y}}$$

Substituindo o valor de $\overline{Y}$, eq. (3.8), obtemos a expressão para Ps:

$$Ps = \frac{1 + aG}{G(1 + 2a) + \epsilon^{-aG}} \tag{3.19}$$

Sob a condição de independência de tráfego, a taxa de transmissões efetivas (ou tráfego escoado pelo canal) é dada por:

$$H = G(1 - Pb) = GPs$$

E, o número médio de programações para retransmissão de um pacote é a razão entre a taxa de transmissões efetivas, H, e a vazão, S, menos 1, ou seja, o número médio de transmissões que detectaram o canal livre e que colidiram menos a que obteve sucesso: $\left(\frac{H}{S} - 1\right)$.

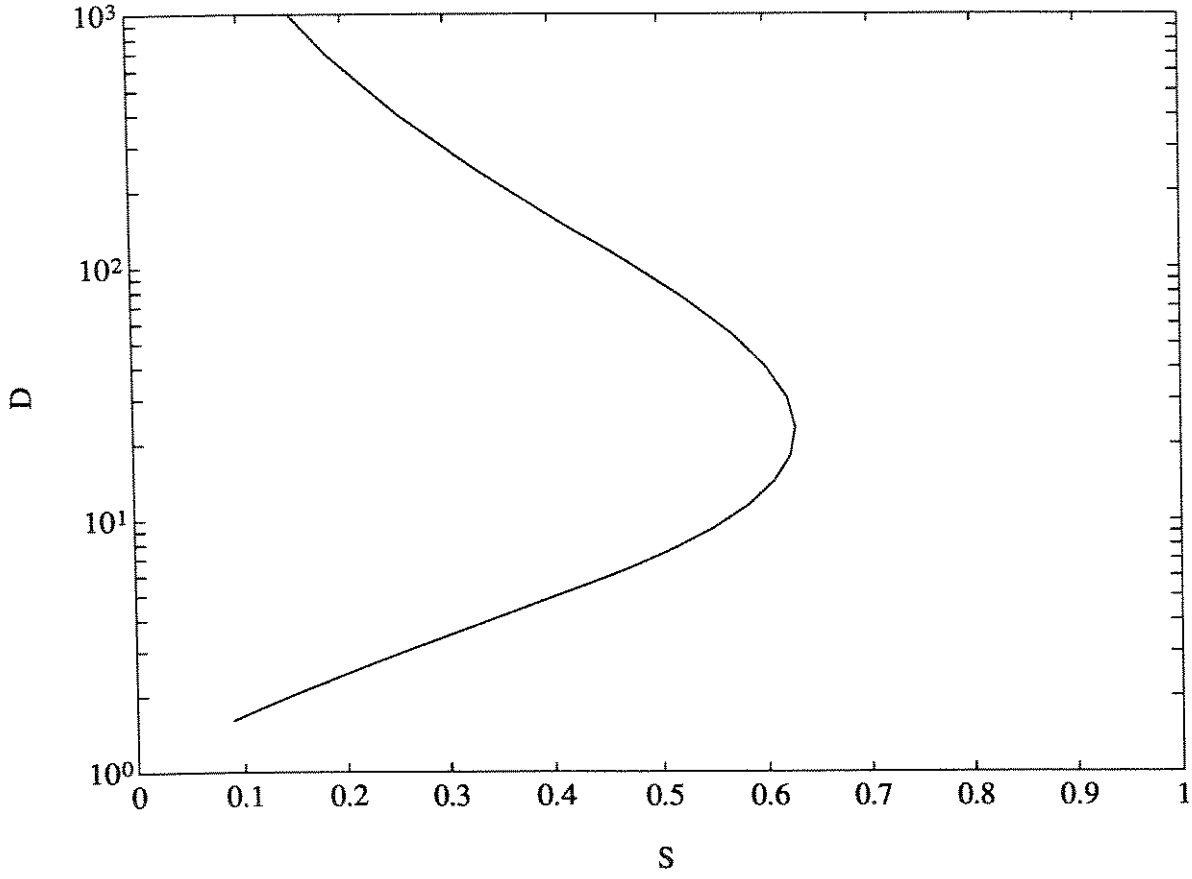

Figura 3.9: Atraso x Vazão: CSMA, eq.(3.20), [4].

Então, a expressão para o atraso médio do pacote (Figura 3.9) é dada por:

$$D = \left(\frac{H}{S} - 1\right)[1 + 2a + \alpha + \delta] + \left(\frac{G - H}{S}\right)\delta + 1 + a \qquad (3.20)$$

3.4.2 M-CSMA-RC

- A probabilidade de transmissão de um pacote com sucesso é: $Ps = \frac{S}{G}$.
- O número de programações para retransmissão de um pacote é uma variável aleatória com distribuição geométrica com média $\frac{G}{S}$, logo, o

número médio de programações para retransmissão de um pacote é $\frac{G}{S}$.

- O número médio de programações sem transmissões é $\frac{G}{S}(1 - Pli)$, onde: Pli é a probabilidade de detectar o canal escolhido livre. Ou seja, $Pli = 1 - Pb$, e Pb é a probabilidade do canal escolhido estar ocupado. Podemos encontrar a expressão para Pli, da mesma maneira em que $(1 - Pb)$ foi obtido no CSMA. Sabemos que $(1 - Pb) = (\overline{I} + a)/\overline{C}$, como $\overline{I} = \frac{1}{G}$ e $\overline{C} = \overline{I} + \overline{B} = \frac{1}{G} + 1 + a + \overline{Y}$, então

$$Pli = \frac{\frac{1}{G} + a}{\frac{1}{G} + 1 + a + \overline{Y}}$$

Substituindo (3.8), obtemos:

$$Pli = \frac{1 + aG}{G(1 + 2a) - \epsilon^{-aG}}$$

Substituindo a por a_i e G por G_i, temos:

$$Pli = \frac{1 + a_i G_i}{G_i(1 + 2a_i) - \epsilon^{-a_i G_i}} \tag{3.21}$$

- O número médio de colisões de um pacote é $\left(\frac{G}{S}Pli - 1\right)$, ou seja, o número médio de programações para retransmissão que detectaram o canal livre (quando este estava ocupado) menos a última tentativa que não sofre colisão no canal.

- Atraso médio de um pacote normalizado em relação a T_0, D_0, é então dado por:

$$D_0 = \left(\frac{G}{S}Pli - 1\right)(M + X_0 + \tau_0 + 2a_0) + \frac{G}{S}(1 - Pli)X_0 + M + a_0 \tag{3.22}$$

onde, X_0 e τ_0 são: o atraso de retransmissão e o tempo de transmissão do ACK, respectivamente. Ambos normalizados em relação a T_0.

3.4.3 M-CSMA-IC

O atraso médio normalizado em relação a T_0 para o protocolo M-CSMA-IC pode ser estimado como em D_0, eq. (3.22). Devemos, apenas, usar Pl, a probabilidade de detectar pelo menos um canal livre, ao invés de Pli, eq. (3.21). Supondo a independência entre os canais, $Pl = 1 - (1 - Pli)^M$ e considerando o modelamento feito baseado nos estados dos M canais, Pl pode ser estimado por:

$$Pl = 1 - \frac{(GM)^M}{M!}\left(1 - \frac{a_0}{M}\right)^M \left(\sum_{k=0}^{M} \frac{(GM)^k}{k!}\right)^{-1} \tag{3.23}$$

Assim, a expressão para o atraso médio do pacote normalizado em relação T_0 para este protocolo, é dada por (Figura 3.10):

$$D_{0IC} = \left(\frac{G}{S_{IC}} Pl - 1\right)(M + X_0 + \tau_0 + 2a_0) + \frac{G}{S_{IC}}(1 - Pl)X_0 + M + a_0 \tag{3.24}$$

3.4.4 M-CSMA-IC com colisões no receptor

Na expressão obtida para o atraso médio do pacote em [4] e reutilizada em [6], foi considerado o atraso do pacote devido às seguintes situações:

- Falta de canais livres no momento em que uma estação com algum pacote pronto deseja iniciar a transmissão.
- Colisões nos canais.

Marsan e Roffinella [6] obtiveram uma expressão para a vazão, S_{CR}, segundo o protocolo M-CSMA-IC considerando colisões no receptor, eq. (3.16). Para obter uma avaliação do atraso do pacote neste caso, propomos utilizar essa expressão para a vazão no cálculo do atraso médio do pacote segundo a expressão para o atraso médio pelo protocolo M-CSMA-IC, $D_{0\,IC}$, sem considerar colisões no receptor, eq. (3.24). Para tal, supomos que a detecção de colisão no receptor seja imediata. Como há uma queda na vazão, quando consideramos colisões no receptor, espera-se um aumento no atraso.

$$D_{0\,CR} = \left(\frac{G}{S_{CR}} Pl - 1\right)(M + X_0 + \tau_0 + 2a_0) + \frac{G}{S_{CR}}(1 - Pl)X_0 + M + a_0 \tag{3.25}$$

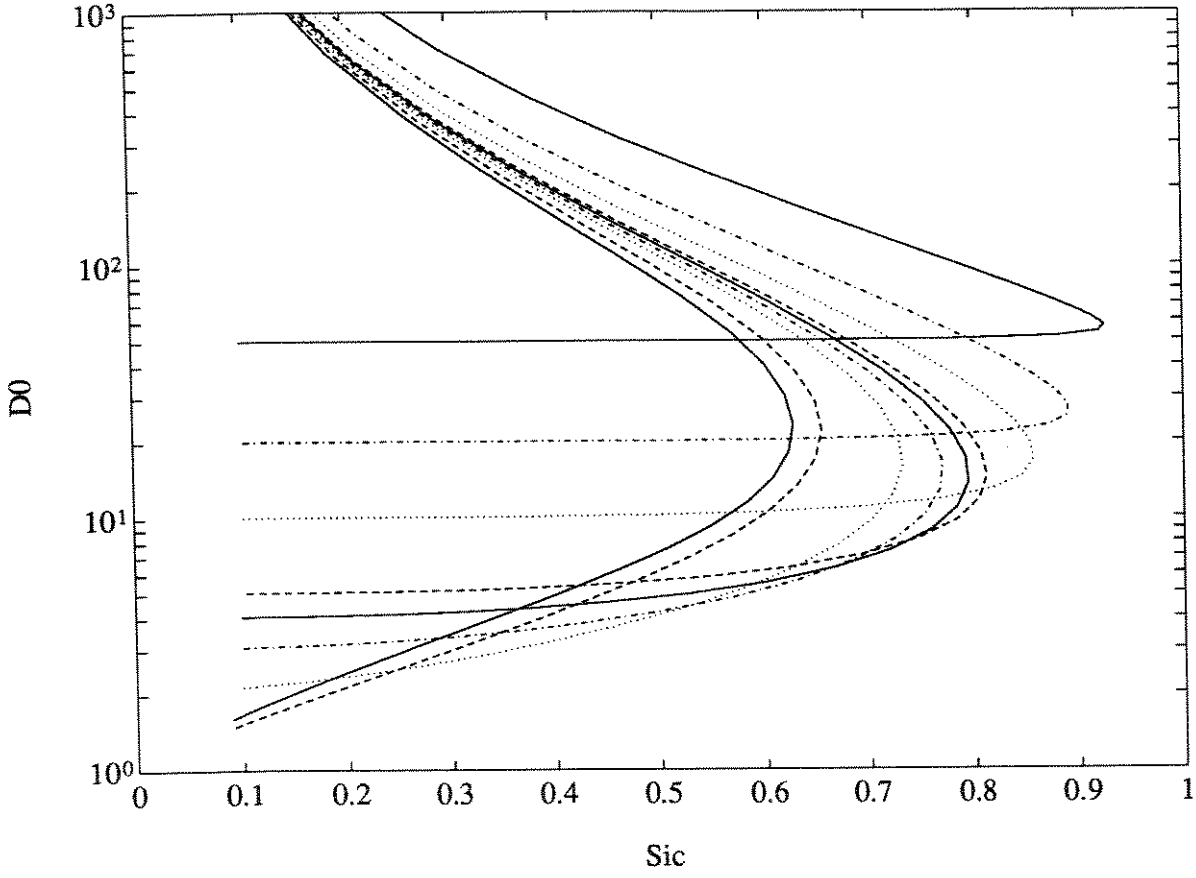

Figura 3.10: Atraso x Vazão: CSMA e M-CSMA-IC com $M = 1, 2, 3, 4, 5, 10, 20, 50$, eq.(3.24), [6].

Uma outra maneira de se calcular o atraso médio de um pacote com colisões no receptor, $D_{0\,CR}$, é através de $\overline{NR}_{CR}$, o número médio de retransmissões devido a colisões no receptor. Assumindo um comportamento markoviano para o tráfego que chega ao receptor, o número de retransmissões é uma variável aleatória com distribuição geométrica com média $1/Ps_{CR}$, onde Ps_{CR} é probabilidade de sucesso considerda, no caso, como sendo a razão entre a vazão com colisões no receptor, eq. (3.16) com a vazão sem considerar as colisões, eq. (3.15), ou seja, $Ps_{CR} = S_{CR}/S_{IC}$. Então, podemos estimar o atraso médio de um pacote por:

$$D_{0\,CR} = D_{0\,IC}\frac{1}{Ps_{CR}} = D_{0\,IC}\frac{S_{IC}}{S_{CR}} \tag{3.26}$$

Supondo, agora, que exista uma fila no receptor de uma dada estação, de maneira que todos os pacotes que chegam a esta estação podem aguardar pelo processamento em uma fila. Não existindo o atraso com a retransmissão do pacote colidido na recepção, computamos apenas o tempo médio de espera na fila. Portanto, para calcular o atraso médio do pacote neste sistema com fila, basta acrescentar o atraso médio de espera na fila da recepção ($\overline{w}^R$) na expressão para o atraso médio que não considera colisões no receptor.

$$D_0^R = D_{0\,IC} + \overline{w}^R \tag{3.27}$$

3.5 Conclusões

O conceito de M-LAN foi apresentado, incluindo o desenvolvimento das expressões utilizadas na análise da vazão para os protocolos CSMA, M-CSMA-RC, M-CSMA-IC e M-CSMA-IC com colisão nos receptores.

Também foi apresentado, o desenvolvimento das expressões para o atraso médio do pacote, segundo os protocolos CSMA, M-CSMA-RC e M-CSMA-IC. Além destes, introduzimos o cálculo do atraso para o protocolo M-CSMA-IC com colisões no receptor. Uma variação deste protocolo, introduz a idéia de filas tanto na chegada das mensagens à estação para serem transmitidas, como também na recepção após terem sido transmitidas com sucesso pelo meio.

A fim de que a transmissão de tráfego misto de pacotes seja possível, podemos associar prioridades aos pacotes com mensagens que exijam uma certa urgência no processamento, como é o caso da voz.

Assim, no próximo capítulo, abordaremos o modelamento de uma rede local multicanal sob circunstâncias de completa ocupação dos canais, de forma a armazenar os pacotes bloqueados em uma fila, tanto no acesso ao meio quanto no destino. Além disso, consideraremos o tráfego misto de mensagens, fazendo uso de filas com prioridade para situações de bloqueio.

Capítulo 4

Modelamento de uma M-LAN com Filas

4.1 Introdução

No Capítulo 3, analisamos a vazão e o atraso médio dos pacotes através de um sistema multicanal. Até então, o principal objetivo era o controle do acesso dos N usuários aos M canais através dos protocolos M-CSMA.

Em situações de completa ocupação do meio, os protocolos do tipo CSMA não-persistente propõem que as transmissões dos pacotes bloqueados sejam postergadas por um período de tempo aleatório. Mesmo que mensagens consigam um canal para seu encaminhamento e não colidam durante a transmissão, estas ainda podem encontrar o usuário-destino ocupado e serem, portanto, bloqueadas no destino. Isto acontece pois, na rede local em estudo, os usuários estão ligados entre si através de M canais paralelos e assíncronos, podendo ter até M transmissões simultâneas endereçadas a um mesmo usuário-destino, cujas capacidades de recepção são finitas. Uma forma de contornar esse problema pode estar no armazenamento dos pacotes bloqueados no momento da recepção.

Uma outra situação interessante, e que requer armazenamento, é quando o sistema se encontra sobrecarregado e, portanto, bloqueia o acesso de pacotes ao meio. Algum dispositivo de espera poderia ser acionado nas estações, nesse momento, armazenando os pacotes que chegam, ordenadamente. Uma outra possibilidade seria a adoção de um mecanismo central de controle dos pacotes oriundos de todas as estações do sistema,

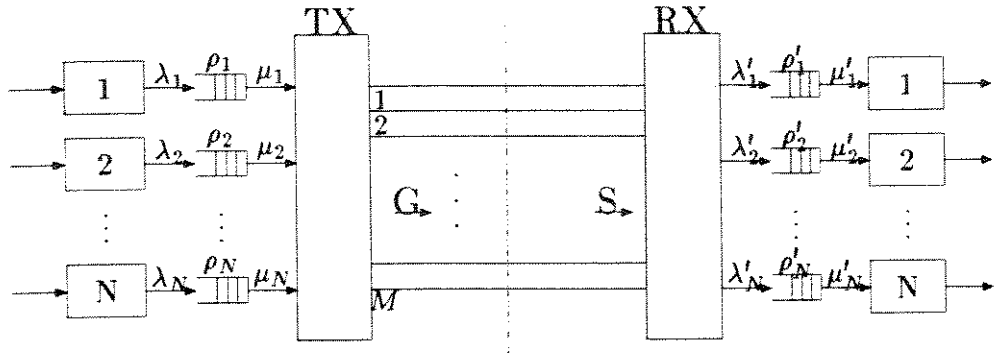

Figura 4.1: Esquema desdobrado de uma M-LAN com N estações e filas

armazenando-os enquanto aguardam pela liberação do meio.

Diante da necessidade de mecanismos de armazenamento e ordenação dos pacotes, a análise da vazão e atraso envolverá um suporte matemático baseado em Teoria de Filas.

Um esquema desdobrado do sistema composto por filas no transmissor e no receptor das estações conectadas entre si através de M canais paralelos, pode ser visto na Figura 4.1.

Neste trabalho, incluiremos filas nos receptores das estações para situações de bloqueio, ordenando os pacotes daquela estação segundo critérios próprios adequados à aplicação.

Em se tratando de tráfego misto de pacotes de voz e de dados, atribuiremos prioridades no tratamento dos pacotes de voz devido à exigência de processamento em tempo real. Esta prioridade, porém, pode ser adaptada às necessidades das estações individualmente.

4.2 Teoria de Filas

Os sistemas que envolvem Teoria de Filas representam uma classe de Sistemas de Fluxo. No nosso caso, o principal objetivo é a aplicação da teoria. Assim, não daremos ênfase a estrutura matemática do modelo.

Os Sistemas de Fluxo dividem-se em duas classes: os de Fluxo Estável, cuja quantidade de fluxo é previsível e constante, e os de Fluxo Instável (Aleatório ou Estocástico), onde os tempos de chegada de mensagens são incertos ou imprevisíveis, assim como a quantidade das mesmas.

A aleatoriedade (imprevisibilidade ou instabilidade) dos sistemas de Fluxo Instável leva a uma considerável complexidade para a solução e entendimento de tais sistemas. A maioria dos sistemas reais se enquadram nesta categoria.

Para o caso mais simples, o problema de fluxo através de um único servidor deixa de ser trivial, como no caso de fluxo determinístico (estável).

4.3 Especificação de Sistemas de Filas

Segundo [4], para se fazer a especificação de um Sistema de Filas deve-se identificar o Processo Estocástico que descreve o fluxo de chegada, a estrutura e a disciplina do sistema .

De maneira geral, o processo de chegada é descrito em termos da distribuição de probabilidade dos tempos entre as chegadas dos usuários, $A(t)$, onde $A(t) = P[\text{tempo entre chegadas} \leq t]$, e a demanda do canal (o número de partidas) é o tempo gasto pelos usuários nas instalações (tempo de serviço) com distribuição $B(x)$, onde $B(x) = P[\text{tempo de serviço} \leq x]$.

Algumas especificações adicionais com relação à estrutura e disciplina das instalações, são dadas a seguir:

- Capacidade (K = tamanho do *buffer*) de armazenamento disponível para espera. Em geral, é assumido K infinito.
- Número de estações de serviço disponíveis.
- Classificação dos usuários; neste caso, a distribuição entre chegadas $A(t)$ e a distribuição do tempo de serviço $B(x)$ são específicas de cada classe.
- Disciplina da fila; que descreve a ordem em que os usuários são servidos. Alguns exemplos são: First-Come-First-Serve (FCFS), Last-Come-First-Serve (LCFS) e ordem aleatória de serviço. Além destas, tem-se a possibilidade de estabelecer prioridades entre grupos de usuários.
- Disponibilidade do servidor, caso este esteja encarregado de outras funções.

Como medidas de desempenho e efetividade, temos:

- w, tempo de espera do usuário (na fila, sem ser atendido)
- n, número de usuários no sistema
- BP, duração de um período ocupado (serviço)
- IP, duração de um período ocioso (*idle*)
- w_0, tempo necessário para completar serviço em andamento.

Todas estas medidas são variáveis aleatórias, cuja Função Distribuição de Probabilidade (PDF) devemos encontrar. Como, em geral, a determinação da PDF é difícil e, em certos casos, impossível de ser obtida, recorre-se à determinação dos primeiros momentos da variável aleatória em consideração. Em sistemas bem comportados estatisticamente - isto é, convergência de variáveis aleatórias elementares descritivas implicando convergência em probabilidade - o primeiro e segundo momentos são os mais importantes.

Adotaremos a notação de Kendall para a especificação das filas, isto é, no formato $A/B/m/K/L/Z$, onde:

- A: descreve a distribuição do tempo entre chegadas
- B: descreve a distribuição do tempo de serviço
- m: denota o número de servidores, $m = 1, 2, \ldots, \infty$
- K: capacidade do *buffer* (tamanho máximo da fila de espera)
- L: representa o número de elementos da população que geram as chamadas ou mensagens
- Z: disciplina (FCFS, LCFS, SPT[1], SRPT[2], SEPT[3], SERPT[4], ...)

[1]SPT = *Shortest-Processing-Time-first.*
[2]SRPT = *Shortest-Remaining-Processing-Time-first.*
[3]SEPT = *Shortest-Expected-Processing-Time-first.*
[4]SERPT = *Shortest-Expected-Remaining-Processing-Time-first.*

As distribuições A e $B \in \{M, G, Er, D\}$, onde:

- M = Exponencial Negativa (Markoviana)
- G = Geral
- Er = Erlang
- D = Determinística

4.3.1 Identificação do Processo Estocástico

Ao analisarmos uma rede local que segue um protocolo com conteção, observamos que o acesso à rede acontece de forma aleatória no tempo. Desta forma, o processo que envolve o acesso à essa rede é um Processo Estocástico, onde o número de usuários (v. a. discreta) na rede num determinado instante de tempo descreve os estados de modo discreto. Este processo é usualmente chamado de Cadeia.

Como o acesso à rede (mudança de estado) pode se dar a qualquer instante no tempo, dizemos tratar-se de um Processo Contínuo no Tempo. Desta forma, a variável aleatória considerada é aquela que descreve o tempo em que o processo permanece num determinado estado.

Observamos ainda, que a transição de estados ocorre quando algum usuário entra na rede ou sai dela. Ou seja, a transição é de uma unidade e só depende do estado atual, o que nos leva a um Processo Markoviano. Trata-se, portanto, de uma Cadeia de Markov contínua no tempo. E, como as transições só ocorrem para um estado vizinho, E_{i+1} ou E_{i-1}, ao atual, E_i, onde i denota o tamanho da população naquele instante, nos referimos a uma classe especial de Cadeia de Markov conhecida por Processo de Nascimento e Morte.

Chamamos de nascimento à transição de E_i para E_{i+1} (chegada) e de morte à transição de E_i para E_{i-1} (partida). A taxa na qual os nascimentos ocorrem em uma população de tamanho i é denominada λ_i e a taxa de mortes, μ_i. Observe que λ_i e μ_i são independentes do tempo, dependentes apenas do estado E_i. Então, temos uma Cadeia de Markov Homogênea e contínua no tempo do tipo Nascimento e Morte.

O sistema mais simples a se considerar é um Sistema de Nascimento Puro onde assumimos $\mu_i = 0$ para todo i. Para simplificar ainda mais o

problema, assume-se $\lambda_i = \lambda$ para todo $i = 0, 1, 2, \ldots$, o que nos leva à distribuição de Poisson. Ou seja, um Processo de Nascimento Puro com taxa de nascimento (λ) constante, gera uma sequência de nascimentos que constituem um Processo de Poisson.

É também um Processo de Renovação, já que o interesse está em contar transições que acontecem em função do tempo. Considerando-se que os tempos entre as chegadas são variáveis aleatórias independentes e identicamente distribuídas (i.i.d.), o fluxo de chegadas forma um Processo de Renovação Estacionário. Em muitos casos, a soma de um número grande de processos de renovação estacionários e independentes (cada um com distribuição do tempo de renovação arbitrária) leva a um Processo de Poisson.

Vários processos naturais têm seus comportamentos significantemente modelados como processos de Poisson. Em Teoria de Filas, especificamente, muitas propriedades analíticas e probabilísticas (que simplificam em muito o problema) tornam-se válidas quando um processo de Poisson é identificado.

A chegada de usuários a alguma fila descreve um processo de Poisson com os seguintes parâmetros:

- λ = taxa média de chegadas por segundo.
- i = número de chegadas no intervalo de tempo de comprimento t.
- $P_i(t)$ = probabilidade de i chegadas em $(0, t)$, ou probabilidade de estar no estado i em t, ou ainda, probabilidade de ter i elementos.
- $E[i]$ = número médio de chegadas em $(0, t)$, $E[i] = \lambda t$.
- σ_i^2 = variância do número de chegadas, $\sigma_i^2 = \lambda t$.[5]

O tempo entre as chegadas adjacentes, $\tilde{t}$, de um processo poissoniano é uma variável aleatória importante para a descrição do processo. Entretanto, devido à propriedade Markoviana[6], fica difícil especificar quanto tempo o processo permaneceu em determinado estado. Restringem-se, dessa forma, as distribuições que podem descrever esta variável.

[5] Em um processo de Poisson, a média e a variância são idênticas e, neste caso, iguais a λt.

[6] A propriedade Markoviana diz que a história do processo está resumida no estado atual, ou seja, não existe memória.

As funções distribuição (PDF) e densidade (pdf) de probabilidade do tempo entre as chegadas são denotadas por $A(t)$ e $a(t)$, respectivamente. Como $A(t) = P[\tilde{t} \leq t]$, $A(t) = 1 - P[\tilde{t} > t]$. Mas, $P[\tilde{t} > t] = P_0(t)$ que é a probabilidade de não ocorrer nenhuma chegada em (0,t). Então, $A(t) = 1 - P_0(t) = 1 - e^{-\lambda t}$ para $t \geq 0$. Diferenciando, obtemos $a(t) = \lambda e^{-\lambda t}, t \geq 0$ que identifica a distribuição exponencial.

De fato, a exponencial é a única distribuição conhecida que não tem memória quando o tempo é contínuo, então a distribuição que descreve o tempo entre as transições é a exponencial.

Portanto, as distribuições do tempo entre as chegadas $A(t)$ e entre as partidas $B(x)$ da fila são exponenciais com parâmetros λ e μ, respectivamente, quando o sistema não está vazio. Ressaltamos que o tempo entre as partidas (ou mortes) corresponde ao tempo necessário para que o serviço seja completado. Logo, $B(x)$ = distribuição do tempo de serviço.

4.3.2 Identificação da Estrutura e Disciplina

Como já foi mencionado, estamos diante de dois tipos de problema envolvendo Teoria de Filas. No primeiro, M canais (servidores) estão sendo disputados por mensagens e, supondo haver algum tipo de controle das mensagens nas estações (apenas para situações de completa ocupação dos canais), as mensagens dos N usuários que fazem acesso aos canais e os encontram todos ocupados vão sendo armazenadas num *buffer* finito com capacidade para K $(K < N)$ mensagens.

O segundo problema de Teoria de Filas visa ordenar as chegadas de mensagens vindas dos M canais a um mesmo usuário-destino (único servidor, $m = 1$), enquanto o mesmo estiver ocupado, armanezando-as em um *buffer* de capacidade finita (K).

Em ambos os casos, podemos identificar dois tipos diferentes de mensagens: dados e voz, por exemplo, sendo que as transmissões e recepções das informações no formato de voz devem ser em tempo real e as demais podem aguardar algum tempo para serem reconhecidas. Diante desta distinção atribuiremos prioridades maiores às informações na forma de voz em relação às em forma de dados, o que pode resultar numa mudança de posições na fila do *buffer* de recepção de cada usuário.

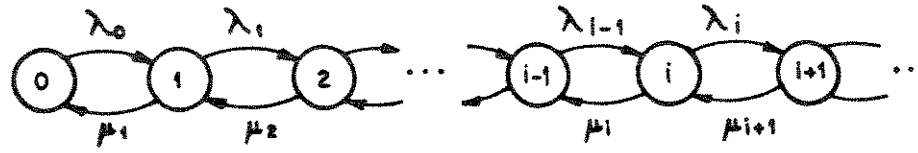

Figura 4.2: Diagrama de Transição de Estados

4.4 Modelamento de uma M-LAN com sobrecarga

A partir de agora, desenvolveremos as expressões para a probabilidade do sistema estar em determinado estado, apresentadas por Kleinrock em [2] para o caso da fila $M/M/m/K/L$ e da fila $M/M/1/K/L$.

O sistema $M/M/m/K/L$, por ser generalizado, é bastante complicado e se reduz a todos os casos mais simples (exceto para chegadas desencorajadas) a medida em que se permite a variação dos parâmetros do sistema.

Considera-se uma população finita de L usuários, cada um com parâmetro λ de chegada. O sistema tem m servidores, cada qual com taxa de serviço μ e capacidade de armazenamento finita de forma que o número de usuários no sistema (fila + serviço) seja, no máximo, igual a K, onde $m \leq K \leq L$.

Aqueles usuários, que na chegada, encontram o *buffer* com sua capacidade esgotada (K usuários no sistema) são rejeitados e devem ser retransmitidos.

De uma forma geral, o modelamento de filas se inicia pela identificação dos parâmetros que compõem o Diagrama de Transição de Estados (Figura 4.2).

As equações de equilíbrio são obtidas por inspeção do diagrama de estados para processos ergódicos ($\{\lambda_i/\mu_i\} < 1$). O fluxo deve ser conservado, ou seja, a taxa de fluxo que entra no estado E_i deve ser igual à taxa de fluxo que sai de E_i. ou seja,

Taxa de fluxo que entra em E_i = Taxa de fluxo que sai de E_i

$$\lambda_{i-1}p_{i-1} + \mu_{i+1}p_{i+1} = (\lambda_i + \mu_i)p_i \tag{4.1}$$

A solução para p_i na equação (4.1) pode ser dada em função de p_0 (p_i,

$i = 0$). Assim sendo,

Para $i = 0$, tem-se $p_1 = \frac{\lambda_0}{\mu_1} p_0$, pois $\mu_0 = \lambda_{-1} = p_{-1} = 0$,

para $i = 1$, tem-se $p_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} p_0$

Por indução, chega-se a

$$p_i = \frac{\lambda_0 \lambda_1 \ldots \lambda_{i-1}}{\mu_1 \mu_2 \ldots \mu_i} p_0$$

Então, para $i = 0, 1, 2, \ldots$, temos:

$$p_i = p_0 \prod_{j=0}^{i-1} \frac{\lambda_j}{\mu_{j+1}} \tag{4.2}$$

Da equação de conservação, $\sum_{i=0}^{\infty} p_i = 1$, podemos determinar p_0:

$$p_0 \left[1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} + \ldots + \frac{\lambda_0 \lambda_1 \ldots \lambda_{i-1}}{\mu_1 \mu_2 \ldots \mu_i} \right] = 1$$

$$p_0 = \frac{1}{1 + \sum_{i=1}^{\infty} \prod_{j=0}^{i-1} \frac{\lambda_j}{\mu_{j+1}}} \tag{4.3}$$

A seguir, iremos considerar as filas $M/M/m/K/L$ e $M/M/1/K/L$

4.4.1 Fila $M/M/m/K/L$

É um processo poissoniano com distribuição exponencial para o tempo entre as chegadas e para o tempo de serviço, com m servidores para L usuários com capacidade de armazenamento K, onde $m \leq K \leq L$.

Assumiremos que a taxa de chegada será descrita matematicamente por:

$$\lambda_i = \begin{cases} (L - i)\lambda & , \ 0 \leq i \leq K - 1 \\ 0 & , \ \text{caso contrário} \end{cases}$$

bem como a taxa de serviço por:

$$\mu_i = \begin{cases} i\mu & , \ 0 \leq i \leq m \\ m\mu & , \ i \geq m \end{cases}$$

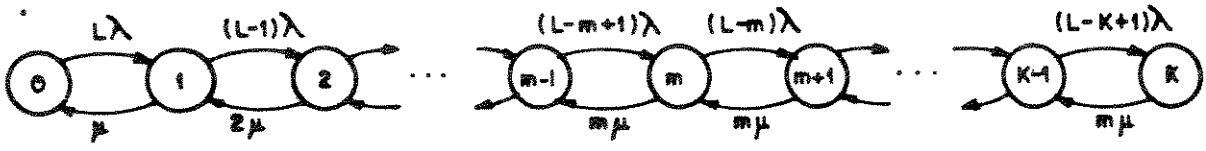

Figura 4.3: Fila $M/M/m/K/L$: Diagrama de Transição de Estados.

Como consequência da conservação de fluxo temos que as equações de equilíbrio podem ser obtidas do diagrama de estados correspondente e são dadas por:

$$1- \quad (L-i+1)\lambda p_{i-1} + (i+1)\mu p_{i+1} = [i\mu + (L-i)\lambda]p_i \quad , \quad 0 \leq i \leq m-1$$

$$2- \quad (L-i+1)\lambda p_{i-1} + m\mu p_{i+1} = [m\mu + (L-i)\lambda]p_i \quad , \quad m < i < K$$

A solução deste sistema de equações pode ser obtida recursivamente. Para a primeira equação (1), com i variando de 0 até $m-1$, podemos encontrar uma expressão genérica para a probabilidade de se encontrar i usuários no sistema no momento de chegada.

Para $i = 0$, como $p_{-1} = 0$, temos:

$$p_1 = L\frac{\lambda}{\mu}p_0$$

Para $i = 1$, temos:

$$L\lambda p_0 + 2\mu p_2 = [\mu + (L-1)\lambda]p_1$$

Isolando o termo com p_2 e substituindo p_1 pelo resultado obtido acima, chegamos a:

$$p_2 = L(L-1)\frac{\lambda^2}{2\mu^2}p_0$$

Para $i = 2$,

$$(L-1)\lambda p_1 + 3\mu p_3 = [2\mu + (L-2)\lambda]p_2$$

Isolando p_3 e substituindo os valores encontrados para p_1 e p_2, encontramos:

$$p_3 = L(L-1)(L-2)\frac{\lambda^3}{3.2\mu^3}p_0$$

Por indução, concluimos que:

$$p_i = \frac{L!}{(L-i)!i!}\left(\frac{\lambda}{\mu}\right)^i p_0 \quad , \quad 0 \leq i \leq m-1$$

Agora, considerando a segunda equação (2) com i assumindo valores entre m e K, temos, para $i = m$,

$$(L-m+1)\lambda p_{m-1} + m\mu p_{m+1} = [m\mu + (L-m)\lambda]p_m$$

Substituindo p_m e p_{m-1}, onde:

$$p_m = \frac{L!}{(L-m)!m!}\left(\frac{\lambda}{\mu}\right)^m p_0$$

e

$$p_{m-1} = \frac{L!}{(L-m+1)!(m-1)!}\left(\frac{\lambda}{\mu}\right)^{m-1} p_0$$

Isolando p_{m+1} e evidenciando p_0, chegamos a:

$$p_{m+1} = \frac{L!}{(L-m-1)!m!}\frac{\lambda^{m+1}}{m\mu^{m+1}}p_0$$

que é equivalente a:

$$p_{m+1} = \frac{(L-m)L!}{(L-m)!m!m}\left(\frac{\lambda}{\mu}\right)^{m+1} p_0$$

ou ainda, que:

$$p_{m+1} = \frac{(L-m)}{m}C_{(L,m)}\left(\frac{\lambda}{\mu}\right)^{m+1} p_0$$

Para $i = m+1$,

$$(L-m)\lambda p_m + m\mu p_{m+2} = [m\mu + (L-m-1)\lambda]p_{m+1}$$

Isolando p_{m+2} e substituindo a expressão para p_{m+1} e p_m, obtidas anteriormente, temos:

$$p_{m+2} = \frac{(L-m)(L-m-1)}{m^2}C_{(L,m)}\left(\frac{\lambda}{\mu}\right)^{m+2} p_0$$

Para $i = m + 2$,

$$(L - m - 1)\lambda p_{m+1} + m\mu p_{m+3} = [m\mu + (L - m - 2)\lambda]p_{m+2}$$

Substituindo p_{m+1} e p_{m+2} pelas expressões já obtidas, chegamos a:

$$p_{m+3} = \frac{(L-m)(L-m-1)(L-m-2)}{m^2 m} \frac{L!}{(L-m)!m!} \left(\frac{\lambda}{\mu}\right)^{m+3} p_0$$

que é equivalente a:

$$p_{m+3} = \frac{L!}{(L-m-3)!m!m^3} \left(\frac{\lambda}{\mu}\right)^{m+3} p_0$$

Da mesma forma que anteriormente, por indução chegamos a:

$$p_i = C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i \frac{i!}{m!} m^{m-i} p_0 \quad , \quad m \leq i \leq K$$

Desta forma, chegamos a expressão para p_i em acordo com [2], ou seja:

$$p_i = \begin{cases} C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i p_0 & , \ 0 \leq i \leq m-1 \\ C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i \frac{i!}{m!} m^{m-i} p_0 & , \ m \leq i \leq K \\ 0 & , \ \text{caso contrário} \end{cases} \tag{4.4}$$

Em [2], não consta o resultado para p_0. Assim, encontraremos a expressão para p_0 através da equação de conservação: $\sum_{i=0}^{\infty} p_i = 1$.

Substituindo p_i, segundo a eq. (4.4), na equação de conservação, obtemos:

$$\sum_{i=0}^{m-1} C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i p_0 + \sum_{i=m}^{K} C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i \frac{i!}{m!} m^{m-i} p_0 = 1$$

Isolando p_0, temos:

$$p_0 = \frac{1}{\left[\sum_{i=0}^{m-1} C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i + \frac{1}{m!} \sum_{i=m}^{K} C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i i! m^{m-i}\right]} \tag{4.5}$$

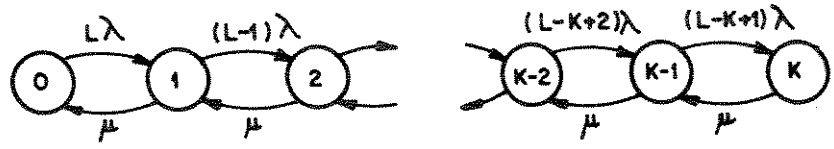

Figura 4.4: Fila $M/M/1/K/L$: Diagrama de Transição de Estados.

4.4.2 Fila $M/M/1/K/L$

É um processo poissoniano com distribuição exponencial para o tempo entre as chegadas e para o tempo de serviço, um servidor para L usuários ($L = M$ canais) com capacidade de armazenamento K, onde $K \leq L$.

A taxa de chegada será descrita por:

$$\lambda_i = \begin{cases} (L-i)\lambda & , \ 0 \leq i < K \\ 0 & , \ \text{caso contrário} \end{cases}$$

enquanto a taxa de serviço, por:

$$\mu_i = \mu$$

Como podemos observar no diagrama de estados (Figura 4.4), as equações de equilíbrio obtidas segundo a conservação de fluxo são dadas por:

$$\begin{cases} (L-i+1)\lambda p_{i-1} + \mu p_{i+1} = [\mu + (L-i)\lambda]p_i & , \ 0 \leq i \leq K \\ 0 & , \ i > K \end{cases}$$

A solução da equação de equilíbrio, pode ser obtida recursivamente, com i variando de 0 a K.

Para $i = 0$,

$$p_1 = L\frac{\lambda}{\mu}p_0$$

Para $i = 1$,

$$L\lambda p_0 + \mu p_2 = [\mu + (L-1)\lambda]p_1$$

Substituindo p_1, obtemos:

$$\mu p_2 = [\mu + (L-1)\lambda]L\frac{\lambda}{\mu}p_0 - L\lambda p_0$$

ou seja,

$$p_2 = L(L-1)\frac{\lambda^2}{\mu^2}p_0$$

Para $i = 2$, temos:

$$(L-1)\lambda p_1 + \mu p_3 = [\mu + (L-2)\lambda]p_2$$

Substituindo p_1 e p_2, encontramos p_3:

$$p_3 = L(L-1)(L-2)\frac{\lambda^3}{\mu^3}p_0$$

Por indução, chegamos a seguinte expressão para $i = K$,

$$p_K = (L-K+1)\frac{\lambda}{\mu}p_{K-1}$$

onde, $p_{K-1} = \left(\frac{\lambda}{\mu}\right)^{K-1} \frac{L!}{(L-K+1)!}p_0$, então,

$$p_K = \frac{L!}{(L-K)!}\left(\frac{\lambda}{\mu}\right)^K p_0$$

Assim, genericamente, temos:

$$p_i = \frac{L!}{(L-i)!}\left(\frac{\lambda}{\mu}\right)^i p_0 \tag{4.6}$$

A obtenção de p_0, vem através da substituição de (4.6) em $\sum_{i=0}^{\infty} p_i = 1$, logo:

$$p_0\left[1 + L\frac{\lambda}{\mu} + L(L-1)\left(\frac{\lambda}{\mu}\right)^2 + \ldots + \frac{L!}{(L-K)!}\left(\frac{\lambda}{\mu}\right)^K\right] = 1$$

$$p_0 = \frac{1}{\left[\sum_{j=0}^{K} \frac{L!}{(L-j)!}\left(\frac{\lambda}{\mu}\right)^j\right]} \tag{4.7}$$

Substituindo p_0 em p_i, temos:

$$\begin{cases} p_i = \frac{\frac{L!}{(L-i)!}\left(\frac{\lambda}{\mu}\right)^i}{1+\sum_{j=1}^{K}\frac{L!}{(L-j)!}\left(\frac{\lambda}{\mu}\right)^j} & , \ 0 \leq i \leq K \\ 0 & , \ \text{caso contrário} \end{cases} \tag{4.8}$$

Uma vez que uma fila com capacidade de armazenamento, K, bloqueia o acesso a qualquer dos $L - K$ usuários que não estão na fila, podemos deduzir que se $K = L$, nenhum usuário é bloqueado e as expressões obtidas nesse modelamento coincidem com as da fila $M/M/1//L$, em [2], onde a capacidade de armazenamento é considerada infinita.

4.5 Parâmetros de Desempenho

O número médio de usuários no sistema, $\overline{n}$, pode ser obtido através da expressão $\overline{n} = \sum_{i=0}^{\infty} i p_i$. Como $p_i = 0$ para $i > K$, então:

$$\overline{n} = \sum_{i=0}^{K} i p_i$$

Substituindo as expressões para p_i dadas pelas equações (4.4) e (4.8) obtemos, respectivamente, o número médio de usuários no sistema para:

- FILA $M/M/1$

$$\overline{n} = \frac{\rho}{(1-\rho)} \tag{4.9}$$

- FILA $M/M/m/K/L$:

$$\overline{n}_m = \frac{\sum_{i=0}^{m-1} i C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i + \sum_{i=m}^{K} i C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i \frac{i!}{m^i} m^{m-i}}{\sum_{i=0}^{m-1} C_{(L,i)} \left(\frac{\lambda}{\mu}\right)^i + \frac{1}{m!}\sum_{j=m}^{K} C_{(L,j)} \left(\frac{\lambda}{\mu}\right)^j j! m^{m-j}} \tag{4.10}$$

- FILA $M/M/1/K/L$:

$$\overline{n}_1 = \frac{\sum_{i=1}^{K} i \frac{L!}{(L-i)!} \left(\frac{\lambda}{\mu}\right)^i}{1 + \sum_{j=1}^{K} \frac{L!}{(L-j)!} \left(\frac{\lambda}{\mu}\right)^j} \tag{4.11}$$

A partir da interpretação direta da fila $M/M/m/K/L$, observamos tratar-se de uma generalização da fila $M/M/1/K/L$ quando $m = 1$. De fato, ao substituirmos m por 1 na expressão (4.10), chegamos, após algumas simplificações, à expressão obtida algebricamente para a fila $M/M/1/K/L$, eq. (4.11).

Segundo Little, para encontrarmos o tempo médio de espera para ser servido, $\overline{w}$, subtrai-se o tempo médio de serviço, $\overline{x}$, do tempo médio total no sistema, T, e como $\overline{n} = \lambda T$, temos que:

$$\overline{w} = T - \overline{x} = \overline{n}/\lambda - 1/\mu \tag{4.12}$$

4.5.1 Atraso com o protocolo M-CSMA-IC com filas

Seja uma rede local multicanal que segue o protocolo M-CSMA-IC. Uma fila do tipo $M/M/1/K/L$ é acionada assim que pacotes transmitidos com sucesso pelo meio não encontrem a estação-destino com o receptor apto a receber pacotes novos. Neste caso, o receptor é o servidor (único) e o tamanho da população é o número de canais, $L = M$.

Uma outra fila na transmissão, a $M/M/m/K/L$, pode ser acionada independentemente da fila na recepção para armazenar ordenadamente pacotes bloqueados ao tentar fazer acesso a um dos M canais. Neste caso, temos $m = M$ servidores com $L = N$ estações gerando tráfego.

Como $\rho = \lambda/\mu$, podemos representar $\overline{w}$ de duas maneiras:

$$\overline{w}^R = \overline{n}_1/\lambda - \rho/\lambda = (\overline{n}_1 - \rho)/\lambda \tag{4.13}$$

para o tempo médio de espera nas filas localizadas nos receptores das estações, ou ainda,

$$\overline{w}^T = (\overline{n}_m/\rho - 1)/\mu \tag{4.14}$$

representando o tempo médio de espera na fila aguardando pela liberação do meio.[7]

Incorporando o tempo médio de espera nas filas ao atraso médio do pacote através dos M canais, D_{0IC}, obtemos três expressões para o atraso médio do pacote:

[7] As expressões (4.13) e (4.14) baseiam-se na eq. (4.12). A diferença entre as duas reside na facilidade de obtenção de parâmetros necessários aos calculos de cada uma delas, em nosso sistema.

A primeira, considera apenas a fila da estação ao transmitir, ou seja,

$$D_0^T = D_{0IC} + \overline{w}^T \tag{4.15}$$

A segunda, semelhante à primeira, considera apenas a fila na recepção, pois supõe a inexistência da necessidade de retransmissões, uma vez que a fila da recepção absorve todos os pacotes lá bloqueados.

$$D_0^R = D_{0IC} + \overline{w}^R \tag{4.16}$$

E, por fim, considerando as duas filas simultaneamente.

$$D_0^{RT} = D_{0IC} + \overline{w}^R + \overline{w}^T \tag{4.17}$$

4.6 Filas com Prioridades

Como já foi mencionado, as filas podem ser classificadas segundo a disciplina para escolha do próximo usuário a ser servido. Esta decisão pode se basear em alguma medida relacionada com o tempo de chegadas dos usuários à fila (FCFS, LCLS e ordem de serviço aleatória), com o tempo corrido ou remanescente de serviço dos usuários (SJF[8], LJF[9], SRPT, ...), ou ainda, segundo alguma função do tipo de usuário (prioridades) [3].

Nosso interesse está em atribuir prioridades aos diferentes tipos de informação: voz, dados, imagem, etc., de forma que as informações na forma de voz sejam processadas sem demora. As demais podem sofrer um pequeno atraso quando houver falta de canais para a transmissão do volume de informações desejado. Temos, neste caso, um sistema com duas classes de prioridades entre os usuários (P = 2).

Devido à complexidade de modelamentos que incluem prioridades, aplicaremos o conceito de prioridades ao mais simples dos sistemas: a fila $M/M/1$.

A fila $M/M/1$, ou seja, fila com único servidor com um processo de chegada poissoniano e uma distribuição exponencial para o tempo de serviço, é um processo de nascimento e morte com coeficientes constantes e forma um dos mais simples modelos de sistemas de filas.

[8] *SJF = Shortest-Job-First*
[9] *LJF = Longhest-Job-First*

A distribuição do tempo entre as chegadas é exponencial, da mesma forma que a do tempo entre as mortes (serviço), desde que a taxa de mortes seja constante ($\mu_i = \mu$, para $i = 1, 2, \ldots$).

A fim de expressar parâmetros para um sistema de filas com prioridades, tais como: taxas de chegada e de serviço, ou mesmo a fração do tempo em que o servidor está ocupado, iremos definir algumas variáveis.

Seja $p = 1, 2, \ldots, P$, então:

λ_p = taxa de chegada dos usuários de prioridade p segundo a distribuição de Poisson.

$\overline{x}_p$ = tempo médio de serviço do usuário da classe p.

ρ_p = fração do tempo em que o servidor está ocupado com usuários da classe p ($\rho_p < 1$).

Desta forma, a taxa total de chegadas é a soma das taxas de chegada de cada classe de prioridade, ou seja,

$$\lambda = \sum_{p=1}^{P} \lambda_p$$

E, a taxa média de serviço, a soma das taxas médias de serviço de cada classe de prioridade proporcionalmente à sua taxa de chegada,

$$\overline{x} = \sum_{p=1}^{P} \left(\frac{\lambda_p}{\lambda} \right) \overline{x}_p$$

A fração do tempo em que o servidor está ocupado, ρ, é o produto da taxa de chegadas pela taxa média de partidas (serviço), $\rho = \lambda \overline{x}, \rho < 1$.

Substituindo λ e $\overline{x}$, temos

$$\rho = \sum_{p=1}^{P} \lambda_p \overline{x}_p = \sum_{p=1}^{P} \rho_p$$

Consideraremos que o sistema é não preemptivo, ou seja, uma vez que algum usuário começa a receber serviço não pode ser interrompido por outro usuário. Quando a decisão estiver entre usuários de mesma classe, então a disciplina a ser adotada é FCFS.

Para a facilidade de realização do modelamento proposto iremos utilizar a seguinte notação:

$\overline{w}_p$ =tempo médio de espera para usuários da classe p

T_p =tempo total médio no sistema para usuários da classe p
De forma que, segundo Little,

$$T_p = \overline{w}_p + \overline{x}_p$$

O tempo de espera, w_p, de um usuário pode ser decomposto em três partes:

1. Atraso devido ao usuário em serviço (se houver) no momento de sua chegada, w_0.

   Utilizaremos o conceito de vida residual[10] para expressar w_0, o atraso médio do usuário que chega e encontra outro em serviço. Ou seja, w_0 é o tempo residual médio de serviço vezes a probabilidade de se encontrar um usuário sendo atendido.

   Como o processo é de Poisson, a fração do tempo em que o servidor está ocupado com usuários do grupo p, ρ_p, expressa a probabilidade de se encontrar algum usuário do grupo em questão, em serviço. E, o tempo médio residual de serviço é dado pelo segundo momento do tempo de serviço, $\overline{x_p^2}$, dividido por duas vezes o primeiro momento do tempo de serviço do usuário da classe de prioridade p, $\overline{x}_p$. Assim,

$$w_0 = \sum_{p=1}^{P} \rho_p \frac{\overline{x_p^2}}{2\overline{x}_p}$$

   Como $\rho_p = \lambda_p \overline{x}_p$, w_0 também pode ser expresso como:

$$w_0 = \sum_{p=1}^{P} \lambda_p \frac{\overline{x_p^2}}{2} \tag{4.18}$$

   Supondo que o tempo médio de serviço é constante para todas as classes de prioridades, $\overline{x_p} = \overline{x}$ com $p = 1, 2, \ldots$, então, o tempo de serviço de qualquer usuário é uma variável aleatória com distribuição

[10] Entende-se por vida residual, o tempo remanescente de serviço de um usuário a partir do momento em que um novo usuário chega e o encontra em serviço. [3]

exponencial e taxa μ, e seu segundo momento é dado por $\overline{x_p^2} = \overline{x^2} = 2/\mu^2$. Então,

$$w_0 = \sum_{p=1}^{P} \lambda_p/\mu^2$$

2. Atraso devido aos usuários na fila quando da sua chegada e que recebem serviço primeiro (w_N).

   O tempo necessário para que os N_{ip} usuários do grupo i na fila, quando o usuário do grupo p chega, sejam servidos é o produto entre o valor médio $\overline{N}_{ip}$ e o tempo médio de serviço de um membro do grupo i. Mas, o tempo de serviço de qualquer membro do grupo i é retirado de forma independente de $B_i(x)$[11], então:

$$w_N = \sum_{i=1}^{P} \overline{x}_i \overline{N}_{ip} \tag{4.19}$$

3. Atraso devido aos usuários que chegam a fila após sua chegada com prioridades maiores, w_M.

   O tempo necessário para os M_{ip} usuários do grupo i que chegam ao sistema enquanto o usuário p está na fila e que recebem serviço antes dele serem atendidos é o produto entre o valor médio de M_{ip} e o tempo médio de serviço de um membro do grupo i, $\overline{x}_i$. Então,

$$w_M = \sum_{i=1}^{P} \overline{x}_i \overline{M}_{ip} \tag{4.20}$$

   Desta forma, o tempo total médio de atraso do usuário do grupo p, $p = 1, 2, \ldots, P$, é:

$$\overline{w}_p = w_0 + w_N + w_M \tag{4.21}$$

   Substituindo (4.19) e (4.20) na expressão para $\overline{w}_p$, eq. (4.21), obtemos:

$$\overline{w}_p = w_0 + \sum_{i=1}^{P} \overline{x}_i (\overline{N}_{ip} + \overline{M}_{ip})$$

[11] $B_i(x)$ = distribuição do tempo de serviço dos usuários do grupo i

Estamos considerando um sistema com prioridades, no qual as mensagens vão sendo colocadas nos grupos a que pertencem a medida em que vão chegando ao canal para transmissão. Uma mensagem do grupo p que chega, entra na fila atrás das mensagens do mesmo grupo e dos grupos de maior prioridade, e na frente do grupo $p-1$ e de grupos de menor prioridade. Estamos nos referindo ao sistema de filas com prioridades HOL (*Head-Of-the-Line*).

Segundo Little, existirão em média $\lambda_i w_i$ mensagens do grupo i na fila quando a mensagem do grupo p chegar,

$$\overline{N}_{ip} = \lambda_i \overline{w}_i, \quad \text{para } i = p, p+1, \ldots, P$$

As mensagens dos grupos de prioridades $p+1, p+2, \ldots, P$ que chegam ao sistema enquanto a mensagem do grupo p estiver na fila, recebem serviço antes dele. Como p gasta em média $\overline{w}_p$ segundos na fila e como cada processo de chegada é independente do tamanho da fila, temos que em média $\lambda_i \overline{w}_p$ mensagens do grupo i chegando enquanto p está na fila, ou seja:

$$\overline{M}_{ip} = \lambda_i \overline{w}_p, \quad \text{para } i = p+1, p+2, \ldots, P$$

Logo, para um sistema de filas com prioridades sem preempção a equação para $\overline{w}_p$ torna-se:

$$\overline{w}_p = w_0 + \sum_{i=p}^{P} \lambda_i \overline{w}_i \overline{x}_i + \sum_{i=p+1}^{P} \lambda_i \overline{w}_p \overline{x}_i$$

Isolando $\overline{w}_p$, obtemos:

$$\overline{w}_p \left[1 - \lambda_p \overline{x}_p - \sum_{i=p+1}^{P} \lambda_i \overline{x}_i \right] = w_0 + \sum_{i=p+1}^{P} \lambda_i \overline{w}_i \overline{x}_i$$

Mas, $\rho_p = \lambda_p \overline{x}_p$, então:

$$\overline{w}_p = \frac{w_0 + \sum_{i=p+1}^{P} \rho_i \overline{w}_i}{1 - \sum_{i=p}^{P} \rho_i}$$

Particularizando para $p = P$, temos que $\overline{w}_P = w_0/(1 - \rho_P)$

Resolvendo de forma recursiva, teremos:
Para $p = P - 1$,

$$\overline{w}_{P-1} = \frac{w_0 + \rho_P \overline{w}_P}{1 - \rho_{P-1} - \rho_P}$$

e, assim por diante, até que genericamente, para qualquer p, temos:

$$\overline{w}_p = \frac{w_0}{(1 - \rho_{p+1} - \rho_{p+2} - \ldots - \rho_P)(1 - \rho_p - \rho_{p+1} \ldots - \rho_P)}$$

Fazendo $\sigma_p = \sum_{i=p}^{P} \rho_i$, para $p = 1, 2, \ldots, P$, temos:

$$\overline{w}_p = \frac{w_0}{(1 - \sigma_{p+1})(1 - \sigma_p)} \tag{4.22}$$

4.6.1 M-CSMA com prioridades

Lembrando que estamos considerando o tráfego misto de pacotes, surge a necessidade de atribuirmos prioridade aos pacotes que exigem tempo real de processamento, como é o caso da voz. Neste caso, o cálculo do atraso é feito da mesma maneira que no sistema com fila e tráfego homogêneo, acrescentando ao atraso do pacote no meio, $D_{0\,IC}$, o tempo médio de espera na fila de recepção do usuário de prioridade p, $\overline{w}_p^R$,

$$D_{0\,p}^R = D_{0\,IC} + \overline{w}_p^R \tag{4.23}$$

onde, $\overline{w}_p^R$ é dado por:

$$\overline{w}_p^R = \frac{w_0^R}{(1 - \sigma_{p+1})(1 - \sigma_p)} \tag{4.24}$$

e, w_0^R por:

$$w_0^R = \frac{\rho_1}{\lambda_1} + \frac{\rho_2}{\lambda_2}$$

Da mesma forma que incluimos uma fila com prioridades nos receptores das estações, seria interessante incluir uma outra fila na chegada de mensagens às estações. Esta facilidade adicional, beneficia o acesso dos usuários ao meio, aos quais, o tratamento recebido pelos seus pacotes fica transparente.

Assim, os pacotes que chegarem a uma determinada estação para serem transmitidos e não encontrarem canais livres no momento da tentativa de transmissão, são redirecionados a uma fila de espera, e armazenados, para tráfego de pacotes homogêneo, em ordem de chegada. O atraso médio do pacote com fila só na transmissão, para o protocolo M-CSMA-IC, foi obtido na seção 4.5.1: Como os pacotes de voz exigem transmissão em tempo real, iremos atribuir maior prioridade a eles que aos demais tipos de pacote. Para a obtenção das expressões para o atraso médio do pacote, nesta situação, basta substituir $\overline{w_p}$ em $\overline{w}$, na equação (4.15), ou seja:

$$D_{0\,p}^T = D_{0\,IC} + \overline{w}_p^T \tag{4.25}$$

onde,

$$\overline{w}_p^T = \frac{w_0^T}{(1-\sigma_{p+1})(1-\sigma_p)}$$

e, $w_0^T = (\rho_1 + \rho_2)/\mu$.

Agora, considerando as duas filas com prioridades, simultaneamente, obtemos a seguinte expressão para o atraso médio do pacote em um sistema com um único canal:

$$D_{0\,p}^{RT} = D_{0\,IC} + \overline{w}_p^R + \overline{w}_p^T \tag{4.26}$$

4.7 Conclusões

Neste capítulo, procuramos justificar a utilização de Teoria de Filas para o modelamento do acesso do usuário à uma rede local com M canais.

Incluimos o desenvolvimento do modelamento de dois tipos de filas: a primeira com $m = M$ servidores representando canais paralelos para a transmissão independente de mensagens e a segunda levando em consideração o conflito gerado na recepção de mensagens vindas de M usuários (canais) para um mesmo servidor (usuário-destino). Ambas com chegadas poissonianas e serviço exponencial, além de população e capacidade de armazenamento finitas.

Deste modelamento, encontramos as expressões que representam parâmetros de desempenho para posterior análise e comparação dos mesmos com resultados já conhecidos. Para isso, atribuiremos valores às variáveis, tais como as taxas de chegada e de serviço, segundo um padrão que possibilite a interação de uma rede local multicanal integrada (M-ILAN) à Rede Digital de Serviços Integrados (RDSI).

Nosso próximo passo é a obtenção de curvas ilustrando o comportamento das diversas expressões obtidas para o atraso médio do pacote. Levaremos em consideração as diversas combinações de atraso no meio mais filas, na recepção e transmissão, com e sem prioridades.

Os resultados obtidos e os valores dos parâmetros utilizados nos cálculos serão apresentados no capítulo de resultados.

Capítulo 5

Resultados

5.1 Introdução

Neste capítulo, traremos os resultados da análise do atraso do pacote através de um sistema de rede local com M canais em paralelo e que segue o protocolo M-CSMA-IC [6], com algumas variações.

Primeiramente, analisaremos as curvas obtidas para o atraso médio do pacote calculado através da equação (3.25) para o protocolo M-CSMA-IC considerando colisões no receptor, Figura 5.1.

Em seguida, analisaremos o comportamento do atraso médio do pacote quando não ocorrem colisões no receptor devido à inclusão de uma fila em cada uma das estações para armazenar os pacotes que chegam enquanto a estação está ocupada com o processamento de um outro pacote.

Como o tratamento de tráfego misto de pacotes é um problema atual e de grande interesse, principalmente em LANs, procuramos considerar tipos diferentes de pacote - voz e dados, por exemplo - sendo transmitidos pelo sistema, de forma que, a necessidade de se distinguir o tratamento entre os tipos de pacotes se torna evidente.

Para que fosse possível armazenar, por algum tempo, informações que exigiam um atraso muito pequeno na transmissão, decidimos priorizar alguns pacotes. Os pacotes de voz exigem tempo real de processamento para poderem manter o nível de bom entendimento. Por outro lado, os pacotes de dados não são prejudicados com o atraso no envio. Assim, podemos priorizar os pacotes de voz para que sejam atendidos na frente dos pacotes de dados. Porém, dependendo do tipo de informação que os pacotes de dados

carregam, pode ser imprescindível a sua imediata transmissão. Neste caso, inverte-se a prioridade.

Podemos, ainda, ter interesse na convivência de pacotes de imagens ao vivo, conjuntamente com imagens sem a necessidade de tempo real de processamento, assim como, imagens paradas, figuras, gráficos, etc.. Neste caso, priorizam-se os pacotes a serem transmitidos em tempo real. Existem, enfim, várias aplicações e a atribuição da maior prioridade fica a cargo de cada estação, individualmente.

Desta forma, concluiremos com a apresentação da análise do atraso médio do pacote acrescido do tempo médio de espera na fila $M/M/1$ com prioridades que anexamos ao receptor das estações.

5.2 Atraso Médio do Pacote: M-CSMA-IC-RC

Como pode ser observado na Figura 5.1, o comportamento do atraso calculado pela equação (3.25), para o protocolo M-CSMA-IC considerando colisões no receptor, não difere muito do atraso calculado pela equação (3.24), para o mesmo protocolo sem considerar as colisões que, porventura, venham a acontecer no receptor, Figura 3.10. A diferença entre os protocolos é mais pronunciada para baixos valores da vazão, e para um grande número de canais.

Desenvolvemos uma outra expressão para o atraso do pacote considerando colisões no receptor, eq. (3.26). Os resultados se aproximam dos resultados obtidos pela eq. (3.25). Uma comparação mais detalhada entre as duas expressões é apresentada no Apêndice B.

Como j ' mencionamos, anteriormente, detectamos uma divergência entre os resultados obtidos no cálculo da vazão pela equação (3.16) e os resultados publicados em [6]. Por consequência, os resultados para o atraso do pacote também divergem pois o atraso é função da vazão.

Uma comparação entre os dois resultados é apresentada no Apêndice B através das Figuras A.2, para a vazão e A.3, para o atraso do pacote.

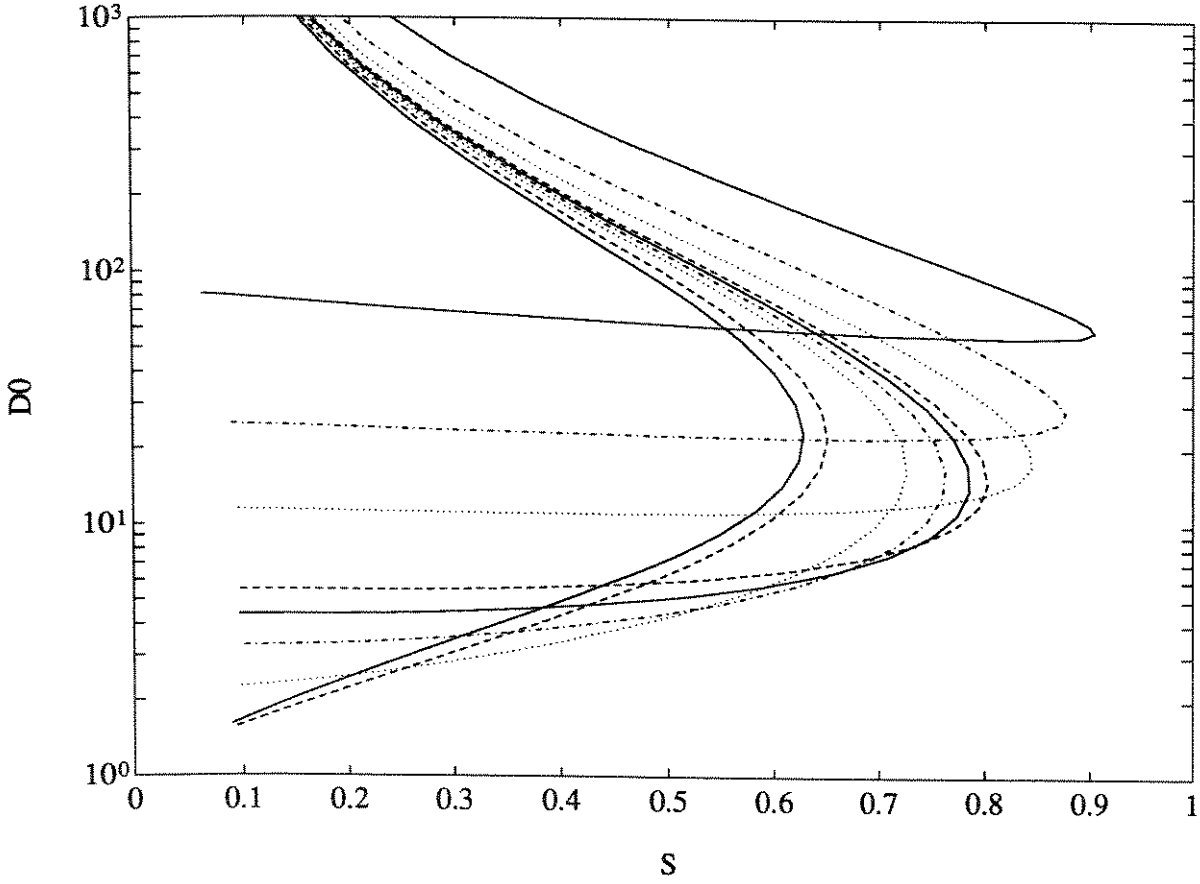

Figura 5.1: Atraso x Vazão: CSMA e M-CSMA-IC-RC, com $N = 100$ e $M = 1, 2, 3, 4, 5, 10, 20, 50$, eq.(3.25).

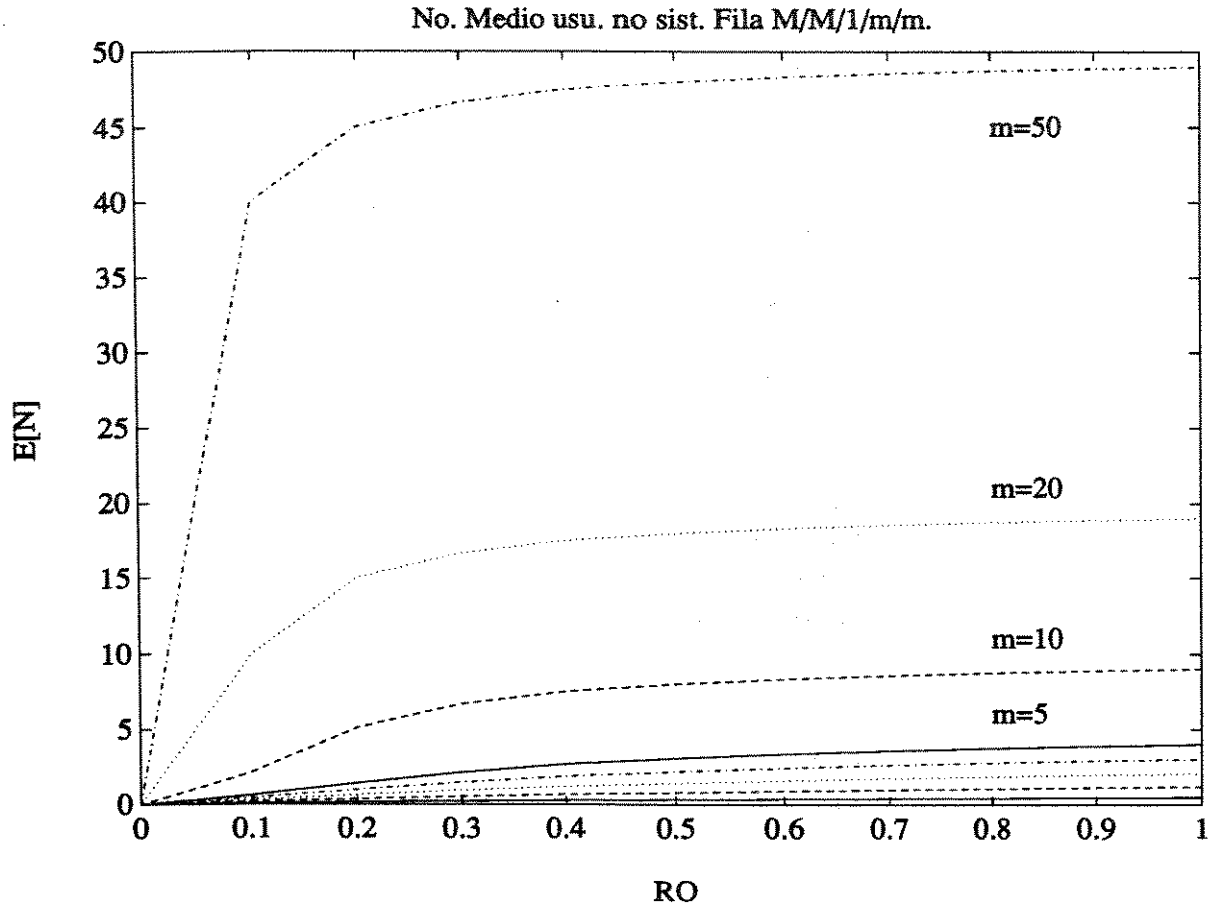

Figura 5.2: Fila $M/M/1/M/M$: Número Médio de Pacotes, eq. (4.11), com $M = 1, 2, 3, 4, 5, 10, 20, 50$.

5.3 M-CSMA-IC com Fila $M/M/1/K/L$ no Receptor

Como descrevemos no Capítulo 4, o cálculo do tempo médio de espera por serviço depende do número médio de usuários, $\overline{n}_1$, na fila com um servidor, eq. (4.11).

Para darmos uma noção do comportamento desta variável, apresentamos a Figura 5.2 com o número médio de usuários na fila $M/M/1/K/L$, com $K = L$ assumindo os valores , conforme ρ varia entre 0 e 1.

O protocolo M-CSMA-IC-RC aumenta o atraso da mensagem em relação ao M-CSMA-IC. Porém, os pacotes colididos, e que devem ser retransmi-

tidos, aumentam desnecessariamente o tráfego oferecido ao meio, uma vez que pacotes colididos no receptor já foram transmitidos com sucesso através do meio. Para que o canal não sofra esse tipo de utilização ineficiente, incluimos uma fila nos receptores. A fila acarreta um acréscimo pequeno ao atraso do pacote referente ao tempo de espera na fila, mas garante a recepção de todos os pacotes que foram transmitidos com sucesso.

Por esse ser um sistema mais próximo de um problema real, acreditamos estar contribuindo para a melhoria do desempenho de sistemas com problemas semelhantes.

Para que possamos dar continuidade a análise dos resultados se faz necessária a definição de alguns parâmetros, como a taxa de pacotes que chega a fila do receptor, por exemplo.

Em [4], o tráfego total oferecido ao sistema, G, foi considerado poissoniano com taxa γ pacotes que chegam (entre novos e reprogramados para transmissão) por tempo de transmissão, T_0.

$$G = \gamma T_0$$

Seguindo o mesmo raciocínio, concluimos que a taxa total de pacotes que sai do sistema, γ', vem da vazão normalizada em relação T_0, ou seja,

$$\gamma' = S_{IC}/T_0$$

Como supomos uniformidade na distribuição dos pacotes que saem dos canais em direção às estações destino, a taxa de chegada à fila do receptor de cada uma das N estações pode ser considerada $\lambda' = \gamma'/N$, em média. Entretanto, para situações de pior caso, consideramos que todo o volume de tráfego escoado se dirige a uma única fila. Assim, assumimos que a taxa de chegada à qualquer das N filas, λ', pode chegar a ser o tráfego escoado por tempo de transmissão, ou seja, a vazão do sistema normalizada em relação a T_0. Mantendo a normalização dos parâmetros, obtemos:

$$\lambda' = S_{IC}$$

Considerando que pacotes de uma mesma mensagem podem entrar no sistema M-canal por qualquer dos M canais, independentemente, é razoável a suposição feita com relação ao comportamento poissoniano da taxa de chegada.

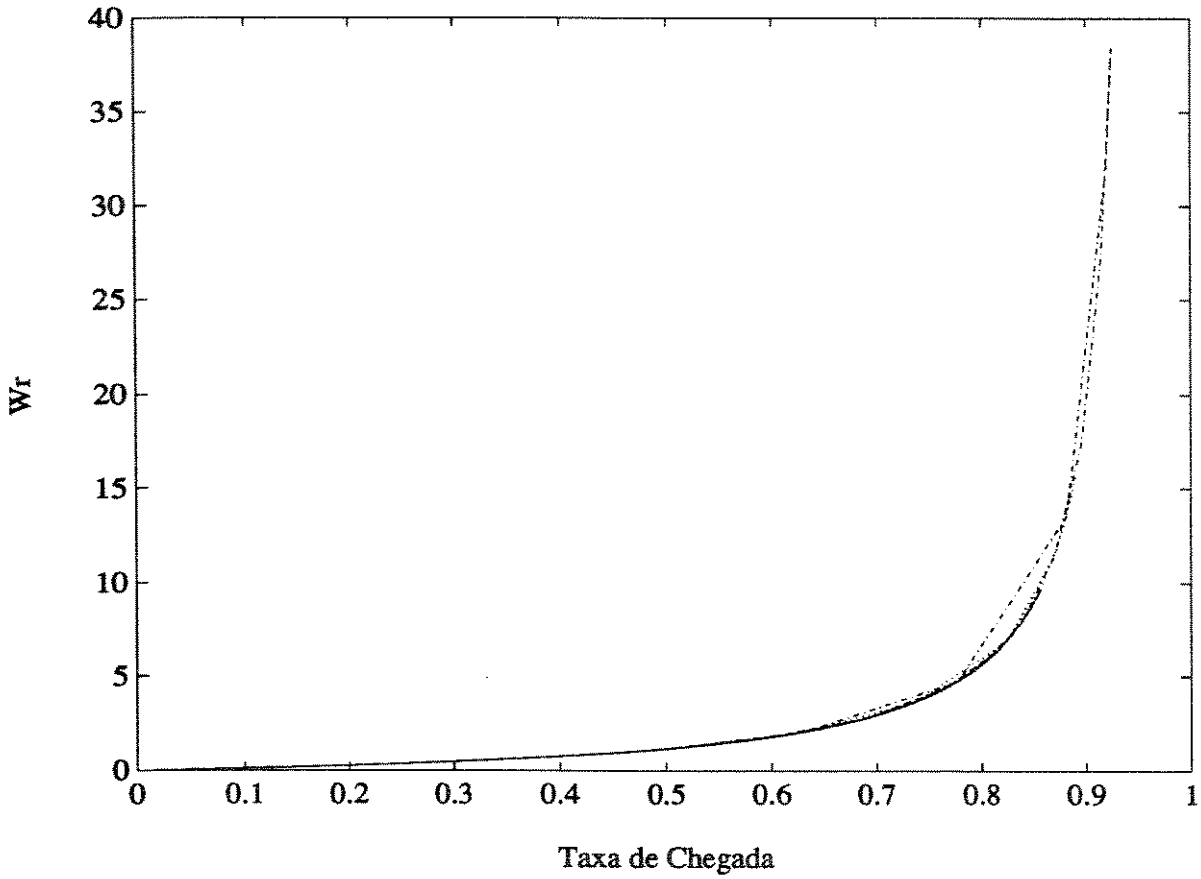

Figura 5.3: Fila $M/M/1$:Tempo Médio de Espera x λ', eq.(4.13).

O tempo médio de espera na fila é obtido através da equação 4.13. O número médio de pacotes no sistema obtido para a fila $M/M/1$ é apresentado na Figura 5.3. Como podemos observar, a curva assume um comportamento assintótico a medida em que a taxa de chegada se aproxima de taxa de serviço.

Na Figura 5.4, trazemos as curvas para o atraso total do pacote consistindo na soma do atraso no meio, segundo o protocolo M-CSMA-IC, com o tempo médio de espera na fila $M/M/1$, que considera infinito o tamanho da população.

Podemos notar que a inclusão de uma fila tipo $M/M/1$ no receptor das estações aumenta o atraso do pacote, com relação ao protocolo M-CSMA-IC sem considerar colisões no receptor, a medida que o sistema se

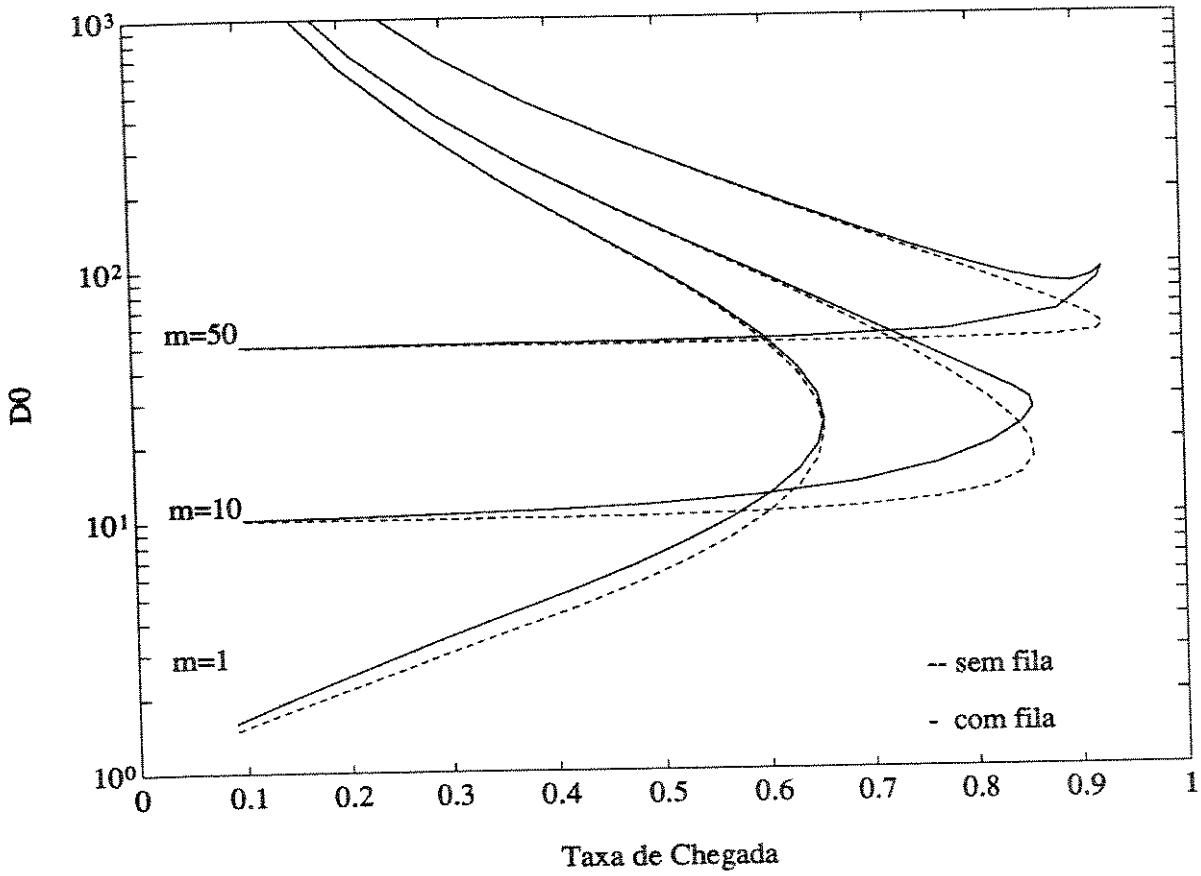

Figura 5.4: Fila $M/M/1$: Atraso Médio do Pacote x λ', eq. (4.16).

aproxima da sua capacidade máxima. No entanto, o atraso com a fila não é superior ao atraso para o mesmo protocolo quando consideramos as colisões no receptor.

Com a inclusão da fila pode-se melhorar a utilização dos canais, pois, não mais retransmitem pacotes que colidem no receptor mesmo após terem sido transmitidos com êxito através do meio.

O aumento no atraso do pacote pode ser minimizado a medida em assumimos uma taxa de serviço maior. Tal efeito é semelhante ao que demonstraremos para a fila $M/M/1$ com prioridades, na Figura 5.11.

5.4 M-CSMA-IC com Fila $M/M/1$ com Prioridades

Nesta seção, ilustraremos o atraso médio do pacote num sistema com filas nos receptores, do tipo $M/M/1$ com prioridades diferentes para os diferentes tipos de pacote considerados na análise. Para tal, precisamos definir alguns dos parâmetros da expressão que calcula, $\overline{w}_p$, o tempo médio de espera na fila, equação (4.22).

A equação para $\overline{w}_p$ depende, basicamente, da taxa média de chegada à fila, λ_p, e da taxa média de serviço, μ_p, onde $p = 1, 2, \ldots, P$ classes de prioridade.

No nosso estudo, consideramos duas classes de prioridades, portanto $P = 2$. A obtenção de λ'_1 e λ'_2 é análoga à descrição feita na seção 5.3, diferente apenas na identificação da proporção da taxa média de chegada (vazão do sistema) que corresponde aos pacotes do tipo 1, classe de menor prioridade e aos pacotes do tipo 2, classe mais prioritária.

Para permitir uma maior abrangência, consideramos três proporções da taxa total de chegada, λ', para representar λ'_2, a taxa média de chegada dos pacotes mais prioritários, tipo 2.

No primeiro caso, foi considerado λ'_2 como sendo 50% da vazão total, $\lambda'_2 = 0.5\lambda'$. Consequentemente, $\lambda'_1 = 0.5\lambda'$. No segundo caso, consideramos o volume do tráfego escoado do tipo 2, 80% da vazão total e, no terceiro caso, 20%.

Para a obtenção do fator de utilização da fila, ρ_i, fixamos a taxa média de serviço, μ, em 0.95, por ser um valor próximo ao máximo valor da vazão normalizada, porém maior, nos permitindo analisar o atraso em uma fila

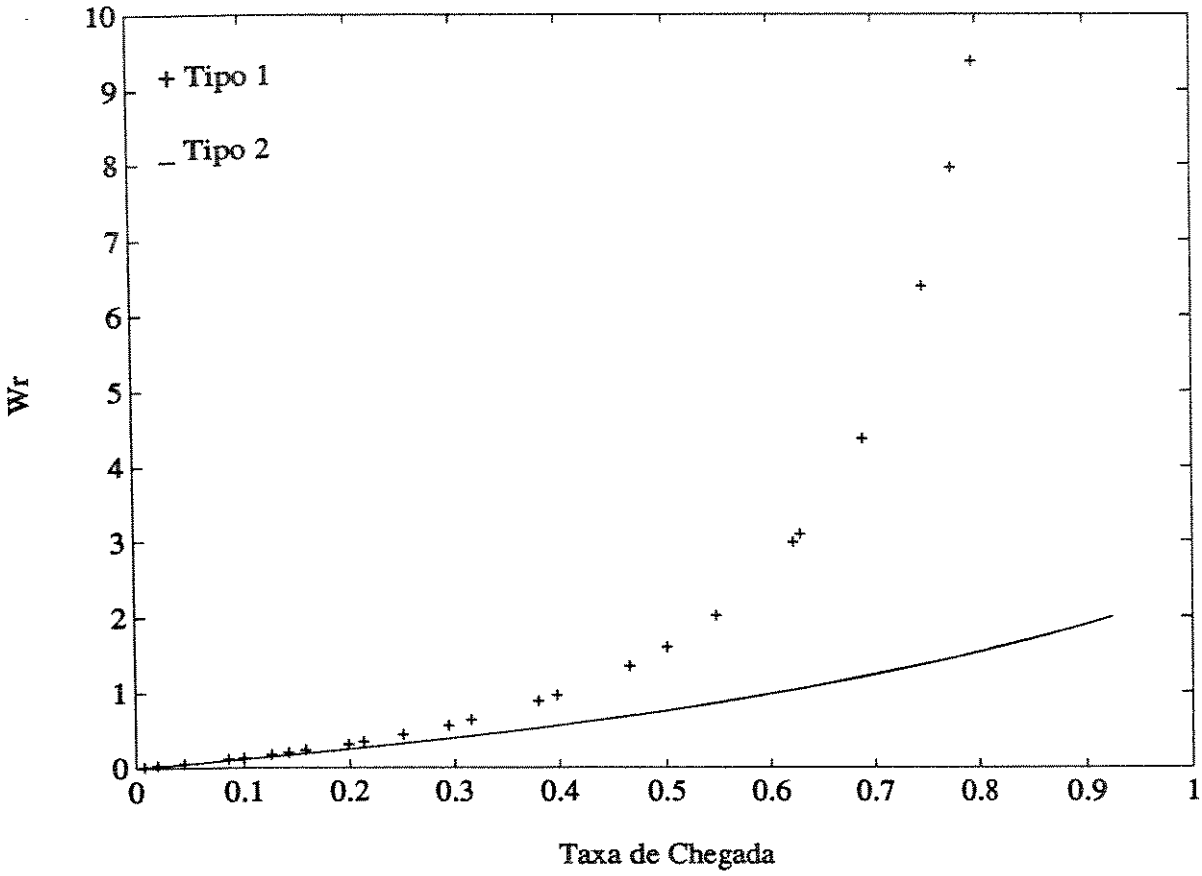

Figura 5.5: Fila $M/M/1$ com Prioridades: Tempo Médio de Espera x λ', $\lambda'_2 = 0.5\lambda'$, $\mu = 0.95$, $M = 50$, eq. (4.22).

com o *buffer* carregado de pacotes a espera de serviço.

Nas Figuras 5.5, 5.6 e 5.7, podemos observar o comportamento da dependência do tempo médio de espera na fila em função à taxa de chegada para os três casos mencionados.

No primeiro caso, quando as taxas de chegada de pacotes tipo 1 e tipo 2 são equivalentes, notamos que os pacotes do tipo 1 tem seu tempo de espera crescendo assintoticamente a medida que a sua taxa de chegada se aproxima da taxa de serviço. Por outro lado, os pacotes de tipo 2 mantém um comportamento linear e com atrasos bem menores que os atrasos sofridos pelos pacotes do tipo 1.

A medida que aumentamos a proporção de pacotes do tipo 2 em relação

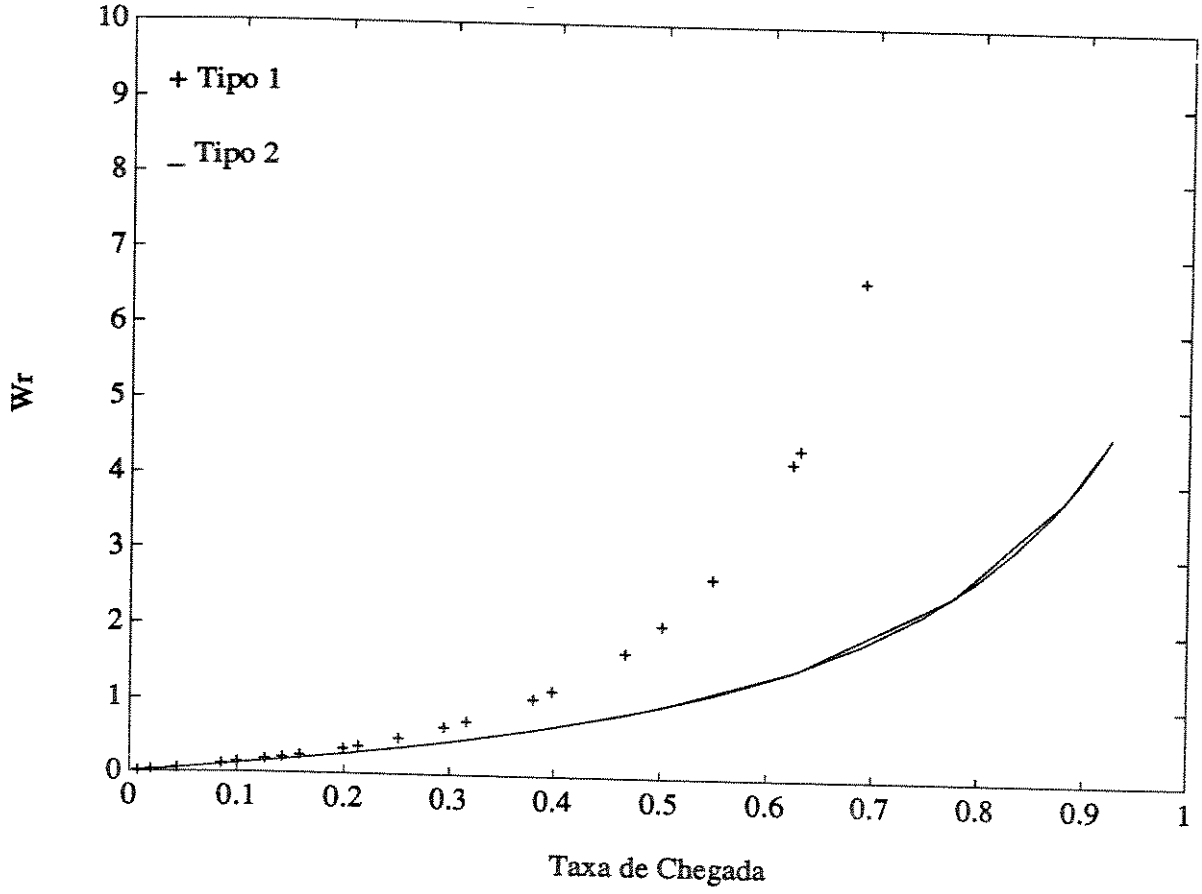

Figura 5.6: Fila $M/M/1$ com Prioridades: Tempo Médio de Espera x λ', $\lambda'_2 = 0.8\lambda'$, $\mu = 0.95$, $M = 50$, eq. (4.22).

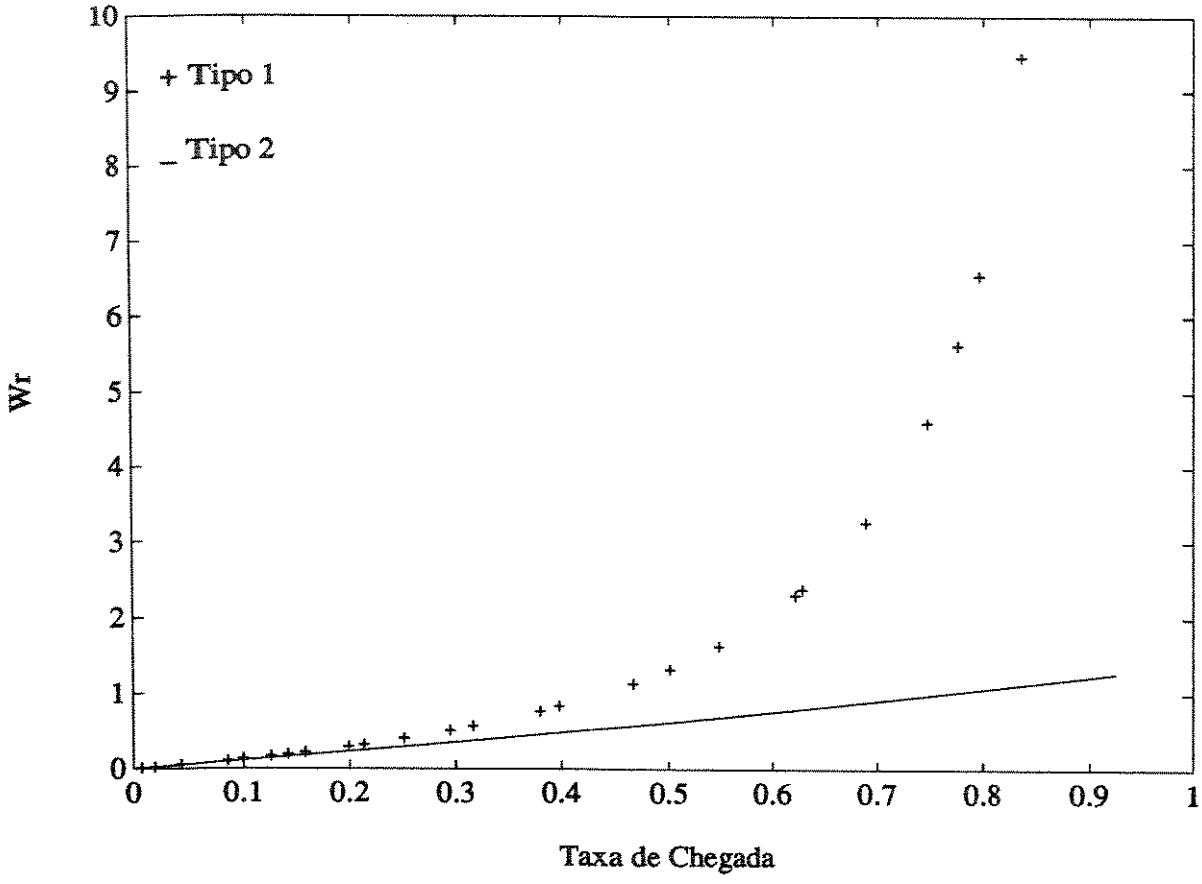

Figura 5.7: Fila $M/M/1$ com Prioridades: Tempo Médio de Espera x λ', $\lambda'_2 = 0.2\lambda'$, $\mu = 0.95$, $M = 50$, eq. (4.22).

a λ', como pode ser observado na Figura 5.6 para $\lambda'_2 = 0.8\lambda'$, notamos mais de 100% de aumento no tempo de espera dos pacotes mais prioritários, pois disputam o servidor com mais pacotes de mesma prioridade. Como era de se esperar, os pacotes do tipo 1 também têm seu desempenho prejudicado, apresentando um aumento no tempo de espera rápido, para um valor mais baixo da taxa total de chegada que quando havia equivalência de tráfego.

No terceiro caso, quando a taxa de chegada dos pacotes tipo 2 é 20% da taxa total, observamos pouca mudança no comportamento do tempo de espera dos pacotes com menor prioridade. No entanto, observamos uma pequena diminuição no tempo de espera dos usuários mais prioritários.

A partir de agora analisaremos o atraso médio do pacote no sistema

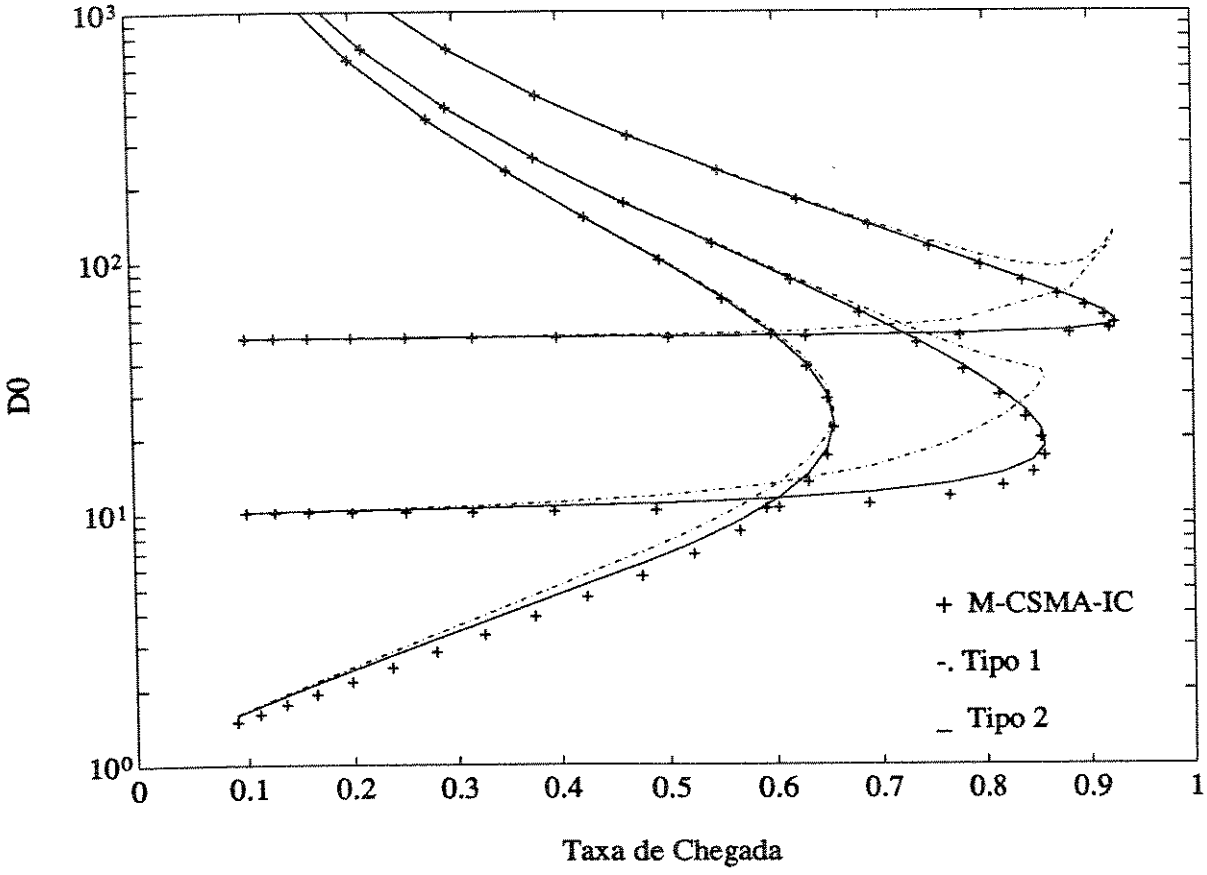

Figura 5.8: Protocolo M-CSMA-IC com a Fila $M/M/1$ com Prioridades: Atraso x Vazão, $\lambda_2' = 0.5\lambda'$, $\mu = 0.95$, $M = 1, 10, 50$, eq. (4.23).

que consiste de N estações conectadas a M canais paralelos e assíncronos e com uma fila no receptor de cada estação.

As Figuras 5.8, 5.9 e 5.10, comparam os atrasos obtidos segundo o protocolo M-CSMA-IC sem considerar colisões no receptor e sem fila com o mesmo protocolo, mas com a adição do tempo de espera na fila com prioridades no receptor, para os três casos, com: 50%, 80% e 20% da taxa total de chegada à fila, de pacotes do tipo 2, conservando os valores também para μ e M.

No gráfico da Figura 5.8, podemos observar o efeito do maior tempo de espera dos pacotes do tipo 1 acrescido ao atraso do meio, segundo o

protocolo M-CSMA-IC sem considerar colisões no receptor, eq. 4.23.

Notamos que, quando o tempo médio de espera na fila assume aquele comportamento assintótico que observamos na Figura 5.5, o atraso médio do pacote no sistema como um todo, aumenta bem acima do atraso para pacotes do tipo 2, sobretudo para altas taxas de chegada e maior número de canais. No entanto, isso só acontece na saturaçao do sistema, quando a vazão começa a diminuir por ter um tráfego acima do tolerável. Essa situação de saturação do sistema é, em geral, controlada de forma a não permitir que o sistema entre em colapso.

Os pacotes de maior prioridade apresentam um pequeno aumento no atraso do pacote em relação ao atraso pelo protocolo M-CSMA-IC sem considerar colisões nos receptores, tendendo a se igualar a medida que aumenta o número de canais.

No caso em que a taxa de chegada dos pacotes tipo 2 é 80% da taxa total, percebemos um maior atraso para os pacotes mais prioritários em relação ao atraso analisado quando havia equivalência de tráfego. Como já haviamos observado no gráfico da Figura 5.6, o tempo de espera dos pacotes menos prioritários assume um comportamento assintótico para uma taxa total de chegada de pacotes mais baixa, ocasionando o atraso mais pronunciado que podemos observar no gráfico da Figura 5.9.

Ao assumirmos uma proporção de 20% de tráfego tipo 2 e 80% de tráfego tipo 1, notamos um comportamento para os pacotes de maior prioridade semelhante ao obtido quando havia equivalência de tráfego. Porém, para pacotes do tipo 1 observa-se uma dimunição do atraso em relação aos outros casos apresentados.

Um valor baixo para a taxa de serviço do sistema foi assumido para que fosse possível a formação de filas que gerassem atrasos significantes no resultado final. Porém, se aumentarmos o valor de μ observamos uma melhora no sistema graças a uma diminuição no atraso médio do pacote. Na Figura 5.11, podemos verificar essa melhoria no desempenho dos pacotes do tipo 1 e tipo 2, quando a proporção de 80% de pacotes prioritários é prejudicial nos dois casos.

E, para valores cada vez maiores de μ, mais semelhante é o comportamento do atraso do sistema com a fila com o do sistema sem a fila calculado pelo protocolo M-CSMA-IC.

Vale ressaltar que o atraso obtido com o protocolo M-CSMA-IC-RC, que

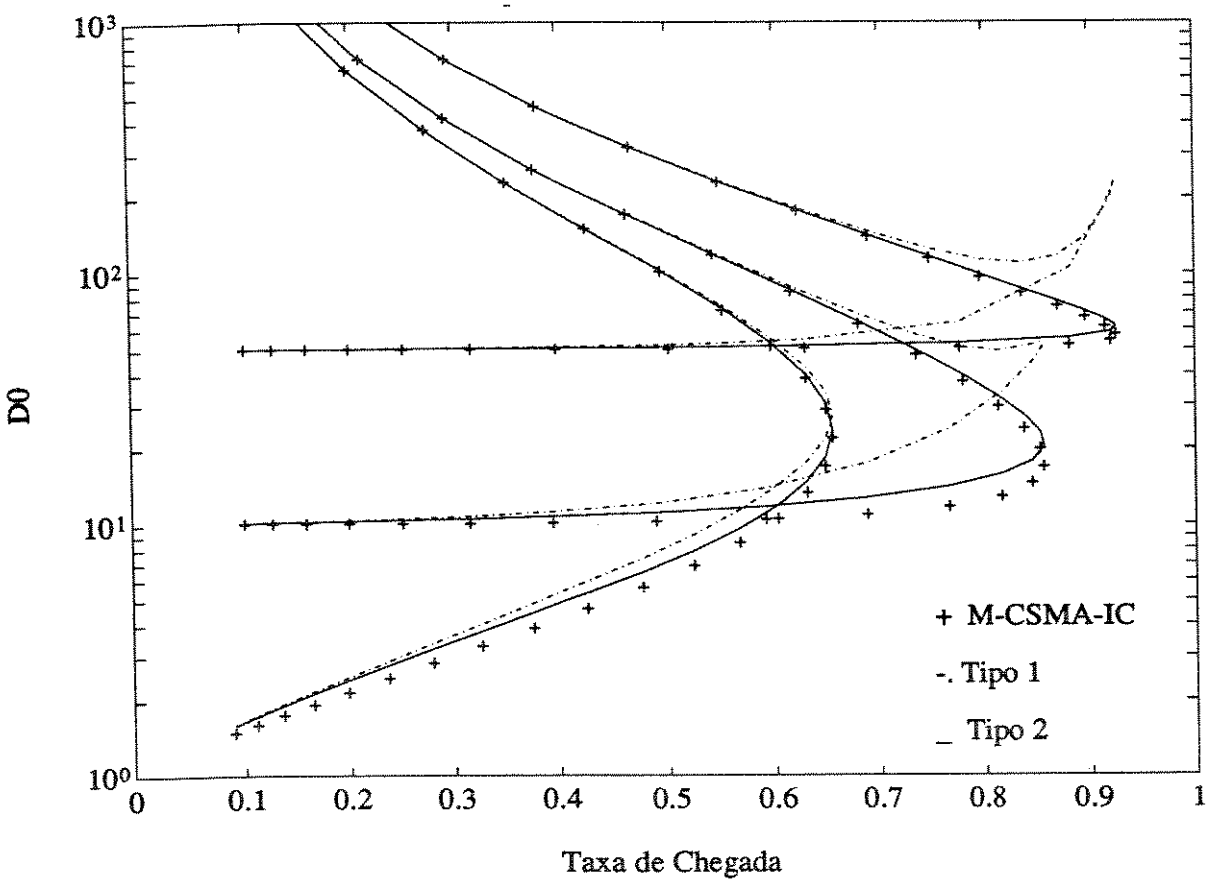

Figura 5.9: Protocolo M-CSMA-IC com a Fila $M/M/1$ com Prioridades: Atraso x Vazão, $\lambda_2' = 0.8\lambda'$. $\mu = 0.95$, $M = 1, 10, 50$, eq. (4.23).

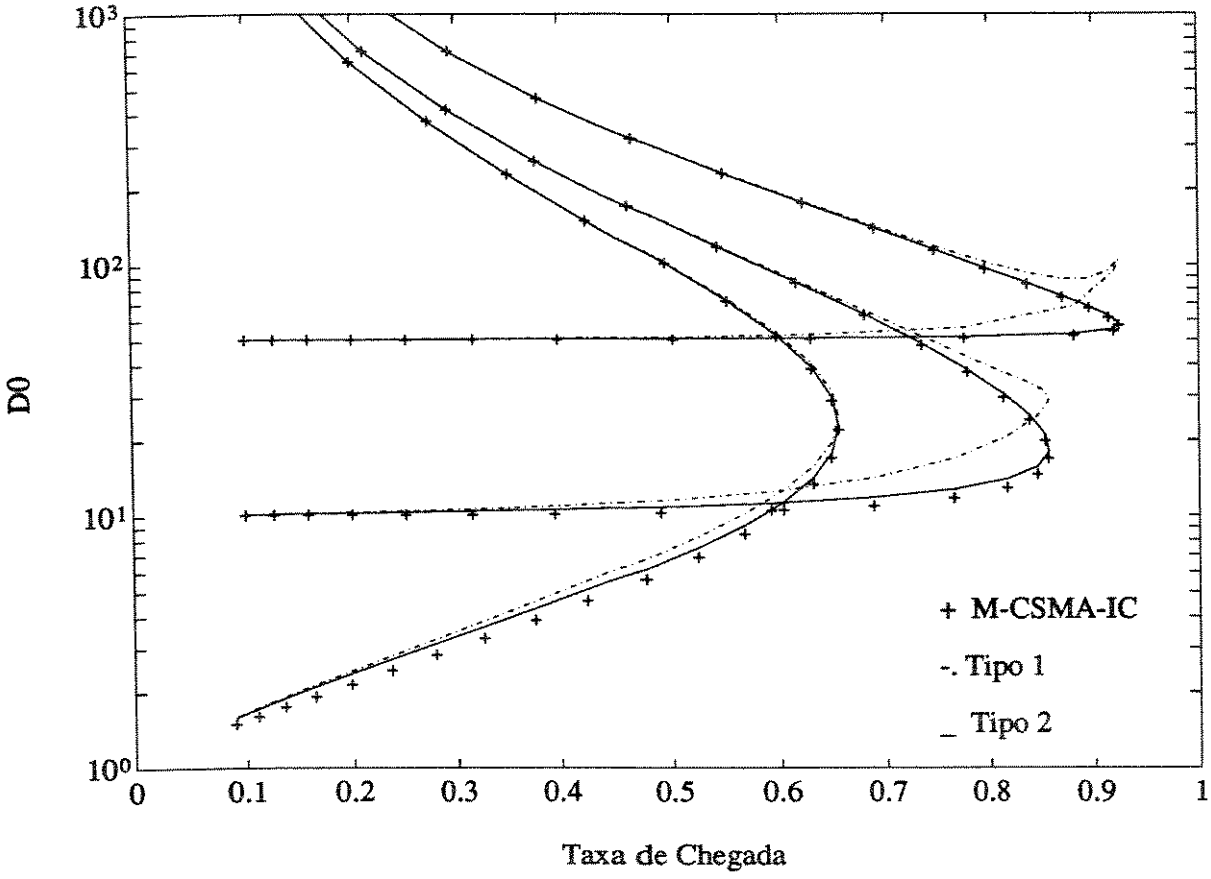

Figura 5.10: Protocolo M-CSMA-IC sem e com a Fila $M/M/1$ com Prioridades: Atraso x Vazão. $\lambda_2' = 0.2\lambda'$, $\mu = 0.95$, para $M = 1, 10, 50$, eq. (4.23).

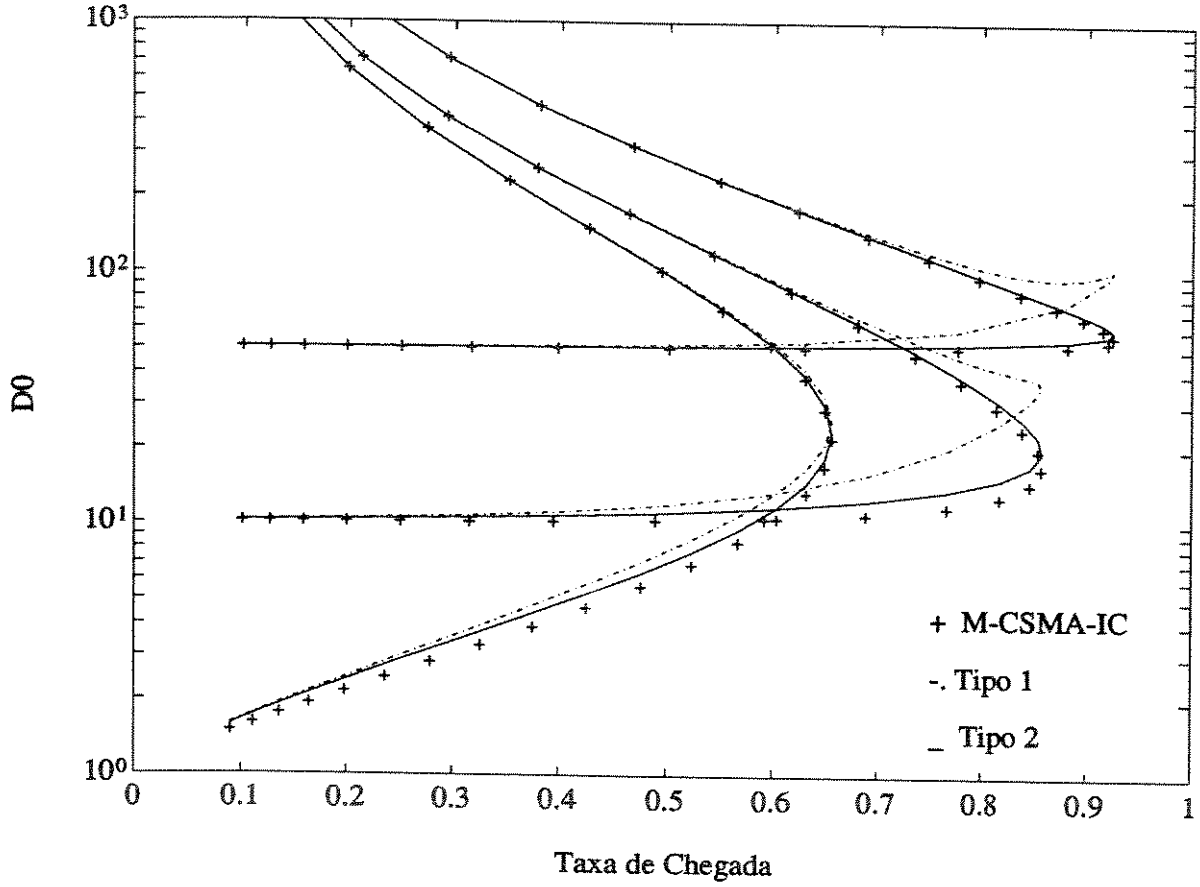

Figura 5.11: Protocolo M-CSMA-IC sem e com a Fila $M/M/1$ com Prioridades: Atraso x Vazão, $\lambda_2' = 0.8\lambda'$, $\mu = 1$, $M = 1, 10, 50$, eq. (4.23).

considera colisões no receptor, se mostrou maior que os atrasos obtidos com a adição da fila no sistema com o protocolo sem considerar as colisões, para valores de λ' até 0.6, aproximadamente. Na maioria dos casos, apenas os pacotes menos prioritários apresentam um maior atraso quando λ' assume valores acima de 0.6.

Para baixos valores de M, notamos que o atraso pelo protocolo M-CSMA-IC-RC é inferior ao obtido pelo protocolo M-CSMA-IC com filas com prioridade. Esse fato deve ser atribuido à imprecisão da expressão (3.15) para a vazão com valores baixos de M, $M < 5$.

Capítulo 6

Conclusões Finais

No primeiro capítulo, fizemos um pequeno histórico sobre redes locais desde o seu surgimento até os rumos de integração destas LANs com redes de acesso público, como a RDSI, que transmite diversos tipos de informação por pacotes. Para um melhor embasamento, passamos a revisar o protocolos de acesso a redes locais mais interessantes, à nossa aplicação, no Capítulo 2.

Com o objetivo de analisar o desempenho de redes locais multicanais desenvolvemos, no Capítulo 3, os procedimentos para a obtenção de alguns parâmetros, tais como a vazão e o atraso dos pacotes segundo os protocolos CSMA não-persistente e M-CSMA.

Já no Capítulo 4, visto que a inclusão de filas nos receptores e, possivelmente, nos transmissores das estações se fez conveniente para se tentar evitar a perda de pacotes por bloqueio, desenvolvemos as expressões para o tempo médio de espera nas filas $M/M/m/K/L$ e $M/M/1/K/L$, além de considerarmos o atraso médio do pacote na fila $M/M/1$ com prioridades.

No Capítulo 5, apresentamos os resultados mais importantes que obtivemos com o modelamento e que mostram as contribuições de se incluir filas para evitar bloqueios de mensagens que fluem por uma M-LAN tratando, inclusive, de tráfego misto.

Resultados interessantes, mostram que o pacote sofre um acréscimo no atraso praticamente desconsiderável, em relação ao protocolo M-CSMA-IC sem considerar colisões no receptor, e, na maioria das situações, apresenta um atraso para o pacote inferior ao obtido pelo mesmo protocolo considerando colisões no receptor.

Neste trabalho, não analisamos o comportamento do atraso médio do pacote com a inclusão de filas no transmissor das estações, apesar de termos modelado um sistema com filas tanto no receptor como também no transmissor das estações. Uma análise da fila no transmissor pode ser facilmente implementada dependendo, apenas, da definição dos parâmetros adequados. A fila no transmissor pode ser interessante sobretudo quando lidamos com prioridades e tráfego misto de pacotes.

A partir dos resultados aqui obtidos, outros modelamentos podem ser desenvolvidos como, por exemplo, a adoção de outro tipo de fila que represente as situações reais, com ainda mais fidelidade. Pode-se considerar outras distribuições para os tempos entre as chegadas e partidas da fila. Ou ainda, um sistema que considere a perda de pacotes por parte das filas no receptor e seu atraso de retransmissão.

No caso de tráfego misto de pacotes, pode-se considerar o atraso que pacotes, que exigem tempo real de processamento, podem estar sujeitos devido ao bloqueio, por parte da fila, quando o tempo de espera for superior ao aceitável para o tipo de mensagem no pacote bloqueado.

Apêndice A

M-CSMA-IC-RC: Divergência na Vazão

Durante a reprodução dos resultados apresentados em [6], houve uma divergência no que diz respeito à vazão segundo o protocolo M-CSMA-IC considerando colisões no receptor, apesar de chegarmos à mesma expressão com nosso modelamento.

Nas curvas obtidas para a vazão, segundo o protocolo M-CSMA-IC, sem considerar colisões no receptor, era observado um aumento da vazão com o tráfego a medida em que um número cada vez maior de canais era alocado. Estas curvas podem ser examinadas na Figura 3.10.

Porém, quando colisões foram consideradas na recepção dos pacotes, o aumento do tráfego não gerou mais um aumento na vazão como vinha se verificando. A Figura A.1 ilustra essa queda na vazão para $M = 20$.

Por algum motivo desconhecido, até o momento, as curvas reproduzidas para a expressão (3.25), como podemos observar através da Figura 5.1, não apresentam essa mesma característica interessante e que tanto motivou o desenvolvimento desse trabalho.

Como o compto da vazão pelo protocolo M-CSMA-IC-RC difere do cálculo da vazão pelo protocolo M-CSMA-IC, apenas pela probabilidade de colisão no receptor, imaginamos estar aí o motivo da divergência.

Analisando com mais detalhes o cálculo do número médio de estações com todos os receptores ativos quando i canais estão ocupados, R_i, observamos um possível erro de arredondamento devido à precisão requerida no cálculo dos Números de Stirling do Segundo Tipo, $S(j,k)$. A equação (3.18)

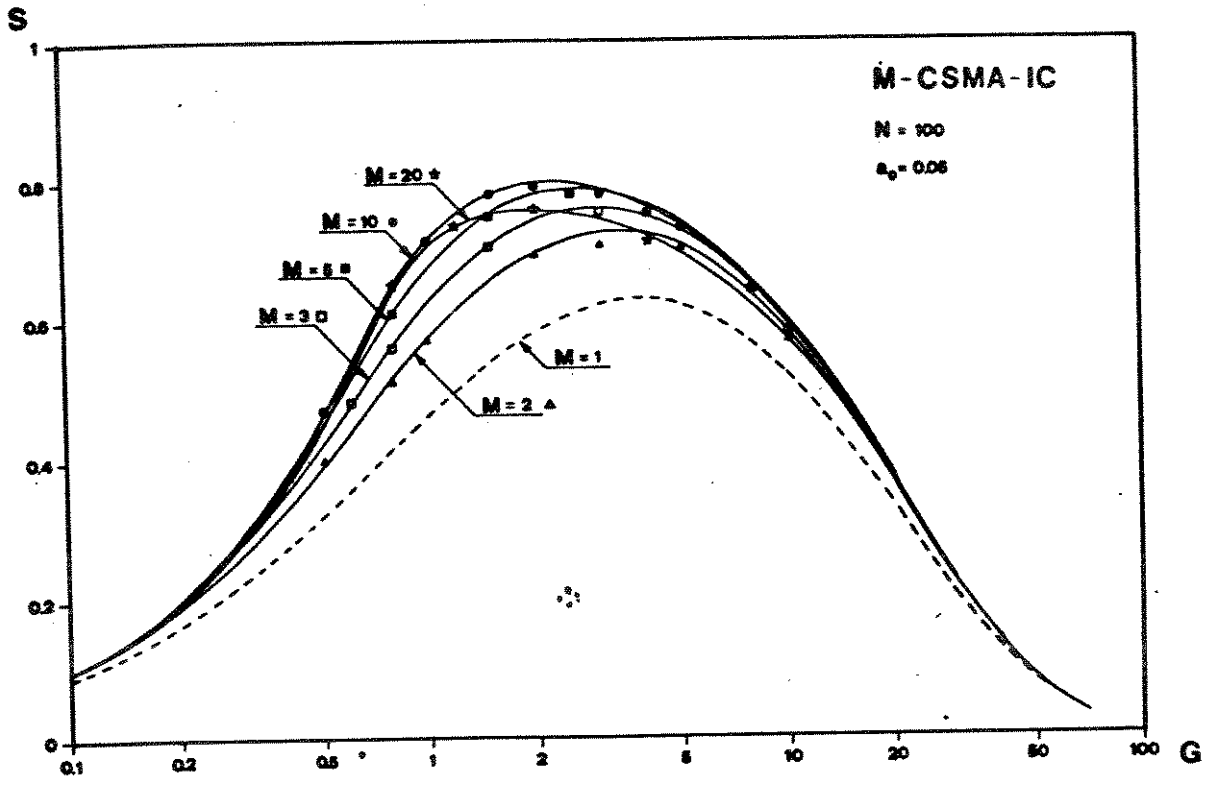

Figura A.1: Vazão x Tráfego: M-CSMA-IC-RC com $N = 100$ estações para $M = 1, 2, 3, 5, 10, 20$ canais e $a_0 = 0.05$, [4]

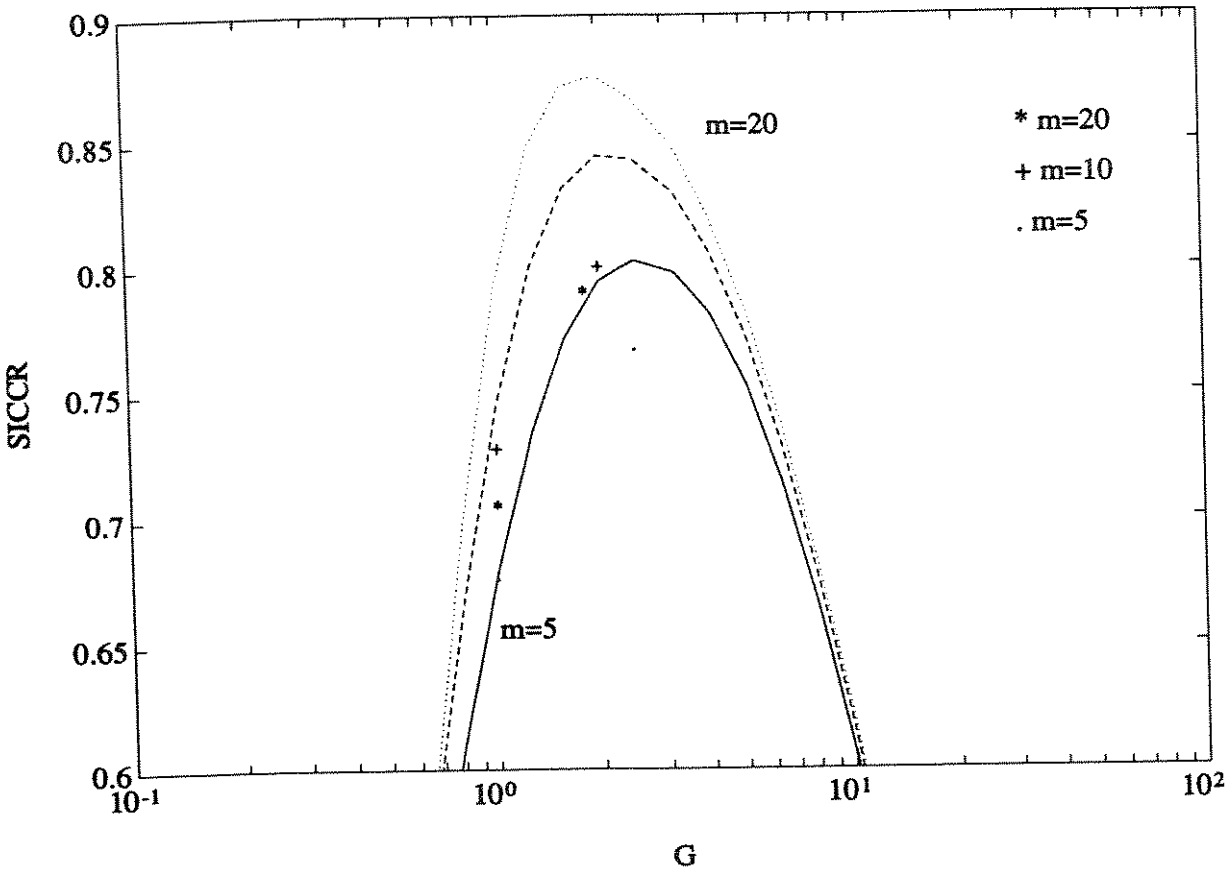

Figura A.2: Vazão x Tráfego: M-CSMA-IC-RC através da eq. (3.16) e pontos
da Fig. A.1.

gera números com até 50 dígitos significativos, sendo, portanto, difícil lidar com tamanha precisão sem gerar tais problemas com arredondamentos.

Tentativas de solucionar, ou mesmo minimizar, tal diferença foram feitas sem a obtenção de melhoras significativas no resultado final.

Assim, a título de comparação, apresentaremos os resultados obtidos segundo a equação (3.16) com alguns pontos das curvas publicadas em [6], na Figura A.2.

O atraso médio do pacote com esse mesmo protocolo, e que não foi apresentado em [6], foi calculado segundo a equação (3.25) para os pontos citados e são apresentados na Tabela A.1 e ilustrados na Figura A.3 com-

as curvas calculadas para a vazão segundo a expressão 3.16.

G	Siccr	D0iccr
1.0	0.467	5.7532
1.0	0.578	5.1235
1.0	0.631	5.5869
1.0	0.672	7.4399
1.0	0.728	12.4960
1.0	0.711	25.6122
1.8	0.756	31.8789
2.0	0.800	18.3026
2.5	0.761	15.6720
2.7	0.767	13.9073
3.1	0.722	15.9121
3.9	0.632	22.2940

Tabela A.1: Tráfego e Vazão: pontos da Fig. A.1 e Atraso, calculado pela eq. (3.25).

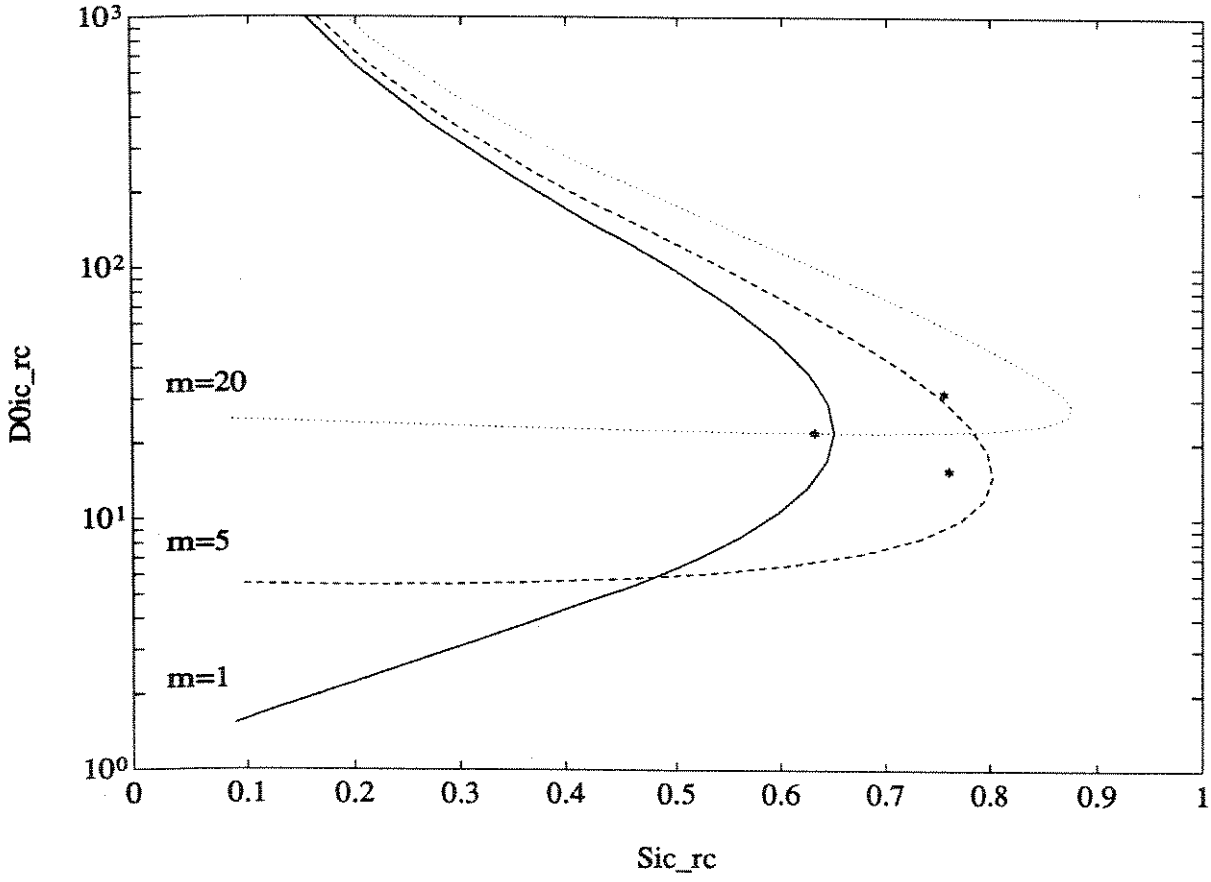

Figura A.3: Atraso x Vazão: M-CSMA-IC-RC através da eq. (3.16) e pontos da Fig. A.1.

Apêndice B

M-CSMA-IC-RC: Expressões para o Atraso

No Capítulo 3, apresentamos duas expressões representando o atraso médio do pacote através de um sistema com M canais paralelos, e que segue o protocolo M-CSMA-IC-RC, que considera colisões nos receptores. As curvas obtidas através da equação (3.26) são bastante semelhantes às curvas obtidas segundo a equação 3.25, que foram apresentadas na Figura 5.1.

Uma comparação entre as expressões (3.24) e (3.25), pode ser vista na Figura B.1 com $M = 1, 20e50$ canais.

Podemos dizer que o valor estimado para o atraso médio do pacote através da equação (3.26) é uma boa aproximação do mesmo valor obtido segundo o desenvolvimento em [6].

Para que possamos analisar com mais detalhes, apresentaremos na Figura B.2, a diferença entre o atraso obtido segundo o desenvolvimento de [6], e que se baseou em [4], eq. (3.25), e o atraso estimado pela equação (3.26) relacionados com a vazão segundo o protocolo M-CSMA-IC-RC, eq. (3.16).

Na Figura B.3, podemos observar a diferença ponto-a-ponto relacionada com o tráfego total oferecido ao canal, G, que varia de 0.1 a 100.

Como podemos observar, o atraso segundo a equação (3.25) é maior que o estimado pela equação (3.26), e quanto mais canais tiver o sistema, maior será a diferença entre eles, que vale 2.5 para $M = 50$, aproximadamente 1, para $M = 20$, e possui valores abaixo de 0.5 para 10 ou menos canais. Notamos, ainda, que a diferença é maior quanto menor for o tráfego.

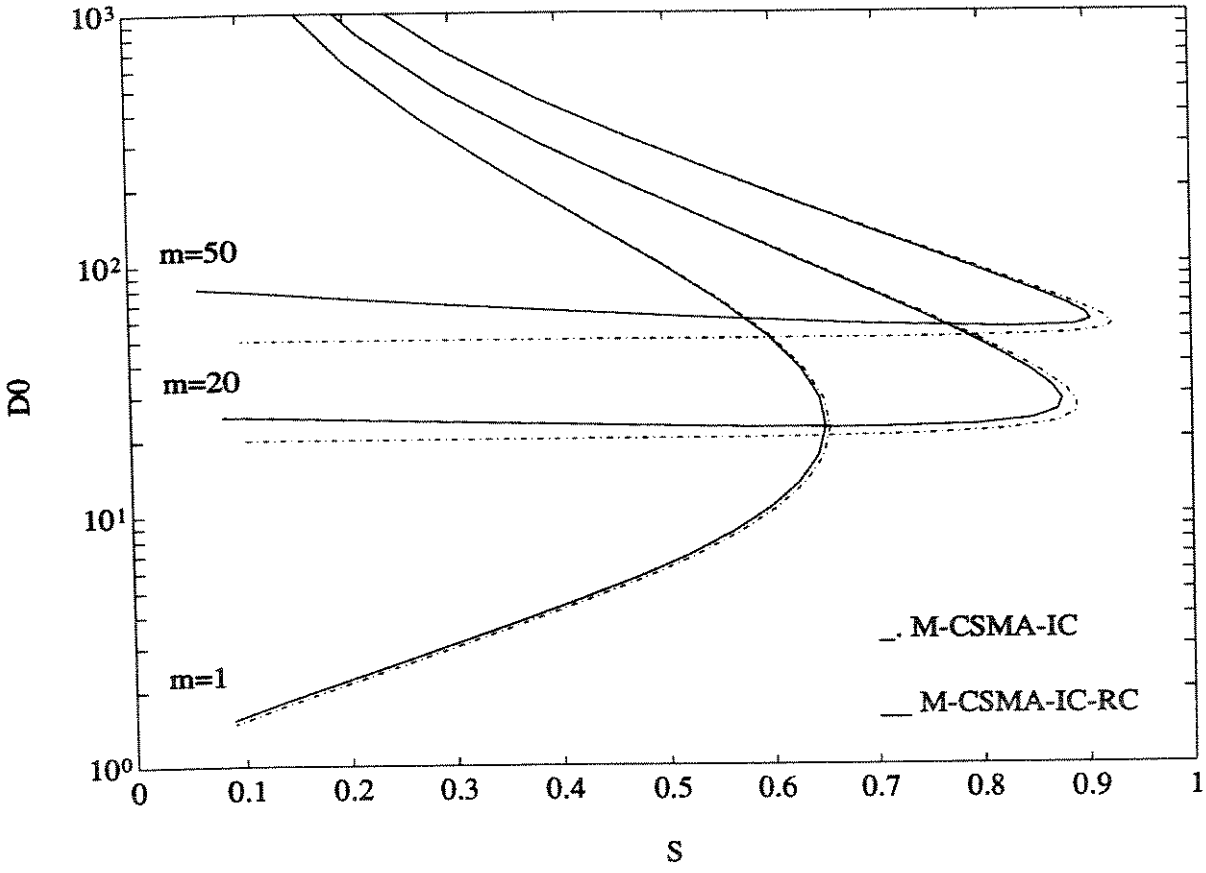

Figura B.1: Atraso x Vazão: M-CSMA-IC, eq. (3.24) e considerando colisões no receptor, eq. (3.25), com $M = 1, 20, 50$ e $N = 100$.

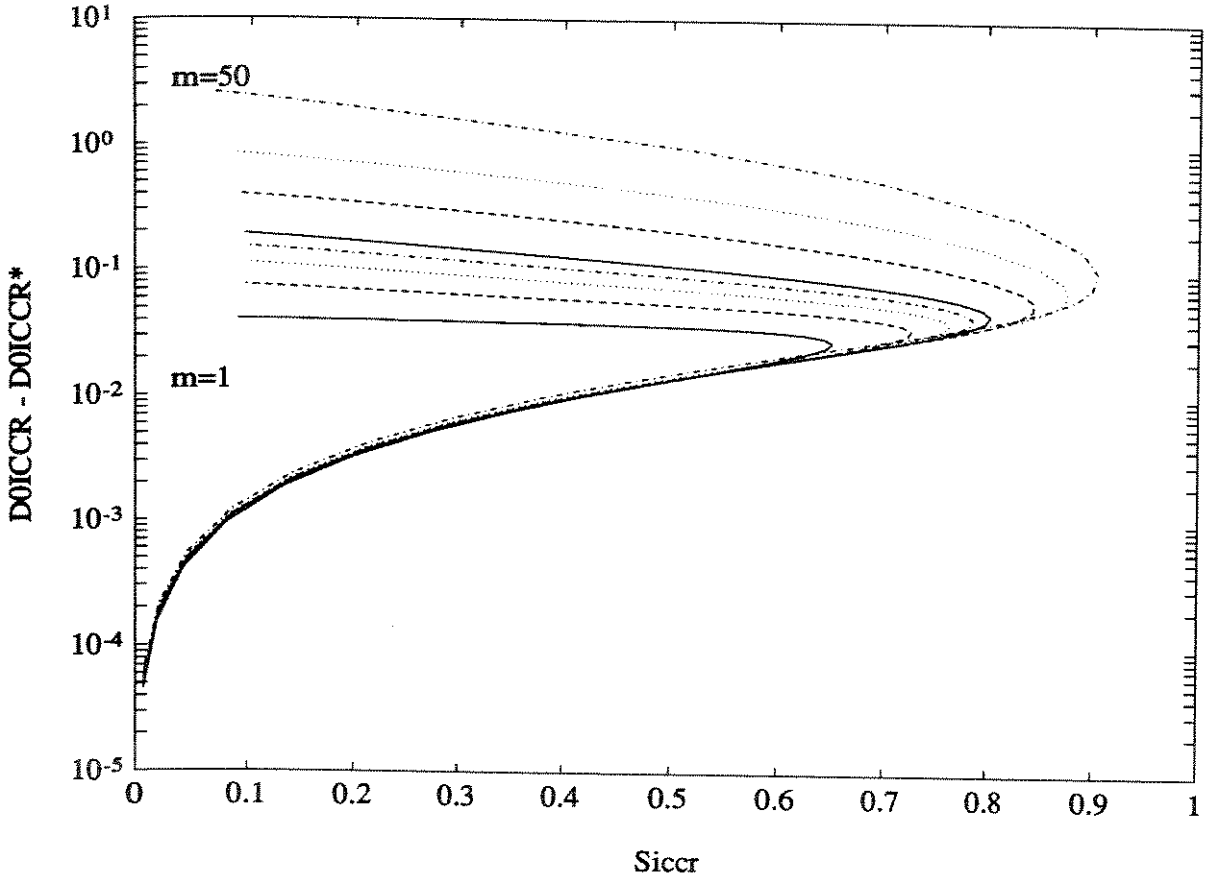

Figura B.2: M-CSMA-IC-RC: Diferença entre eq. (3.25) e eq. (3.26) x Vazão com $M = 1, 2, 3, 4, 5, 10, 20, 50$ e $N = 100$.

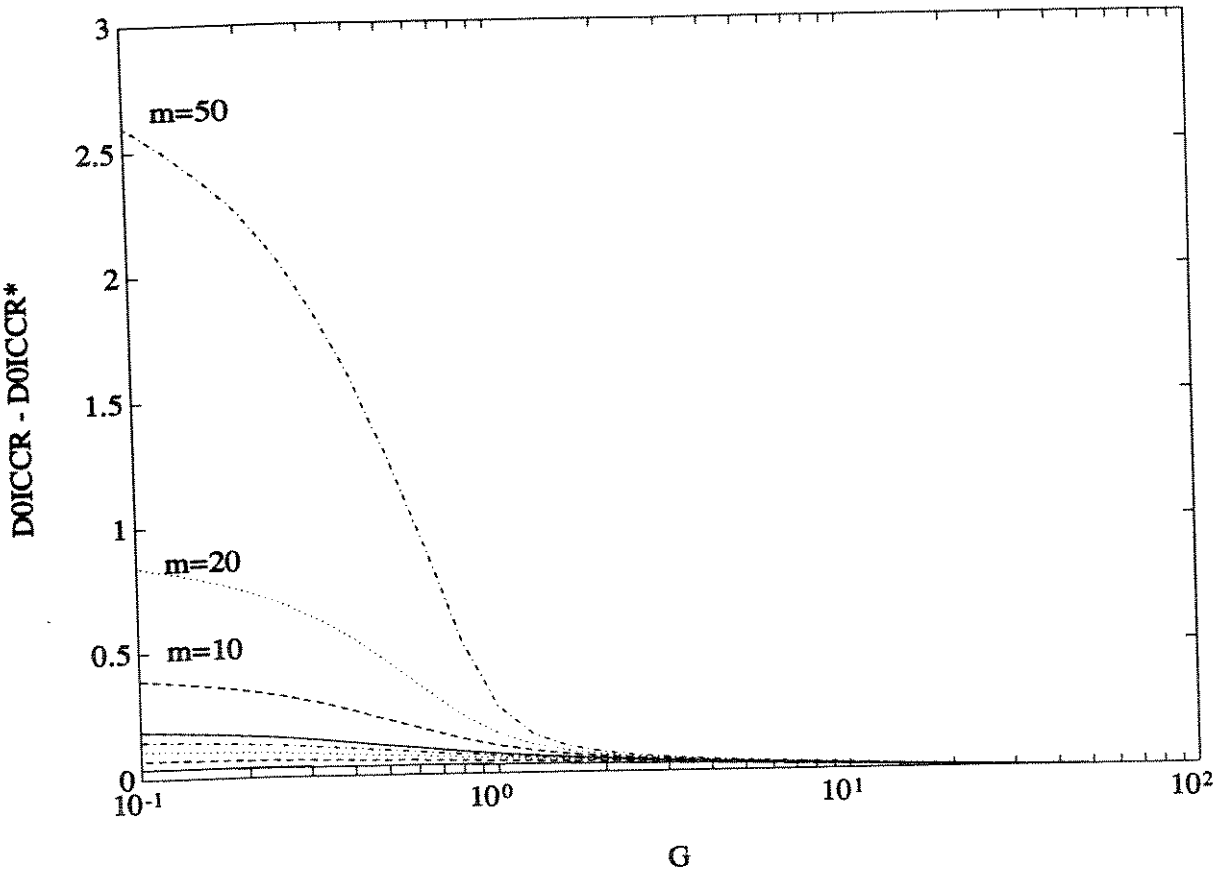

Figura B.3: Diferença entre as eqs. (3.25) e (3.26) x Tráfego:
M-CSMA-IC-RC, com $M = 1, 2, 3, 4, 5, 10, 20, 50$ e $N = 100$.

Bibliografia

[1] M. ABRAMOWITZ and I. A. STEGUN, *Handbook of Mathematical Functions - with formulas, graphs, and mathematical tables*, Dover Publications, Inc., New York.

[2] L. KLEINROCK, *Queueing Systems, vol. I: Theory*. John Wiley & Sons, 1975.

[3] L. KLEINROCK, *Queueing Systems, vol. II: Computer Applications*. John Wiley & Sons, 1976.

[4] L. KLEINROCK and F. A. TOBAGI, "Packet Switching in Radio Channels: Part I - Carrier Sense Multiple-Access Modes and Their Throughput-Delay Characteristics", *IEEE Transactions on Communications*, vol. COM-23, no. 12, pp 1400-1416, December 1975.

[5] L. LELLIS, *Redes Locais de Computadores*. Caderno de Informática, Série Estado da Arte 2, Compucenter.

[6] M. A. MARSAN e D. ROFFINELLA, "Multichannel Local Area Network Protocols", *IEEE Journal on Selected Areas in Communications*, vol. SAC-1, No. 5, pp.885-897, November 1983.

[7] Notas de aula do mini-curso sobre redes locais ministrado por L.F.G. SOARES durante a "I Jornada de Informática", realizada pela PUC-CAMP e IBM, Campinas, setembro, 1989.

[8] R. P. GRIMALDI, *Discrete and Combinatorial Mathematics - An Applied Introduction*, Addison-Wesley Publishing Company.

[9] S. R. SACHS, "Alternative Local Area Network Access Protocols", *IEEE Commun. Magazine*, vol. 26, no. 3, pp. 25-45, March, 1988.

[10] L. F. G. SOARES, *Redes Locais*, Editora Campus Ltda., Rio de Janeiro, 1986.

[11] F. A. TOBAGI, "Multiaccess Protocols in Packet Communication Systems", *IEEE Transactions on Communications*, vol. COM-28, no. 4, pp.468-488, April, 1980.

[12] L. S. PUPO e R. PALAZZO Jr., "Revisão das Técnicas de Múltiplo Acesso para Utilização em Redes Integradas com Multi-Canais", Relatório Técnico No. RT-189, Convênio Unicamp-Telebrás, contrato No. 308/90, Transmissão Digital, Abril/90.

[13] L. S. PUPO e R. PALAZZO Jr., "Modelamento Baseado em Teoria de Filas para Obtenção de Parâmetros para a Análise de Desempenho", Relatório Técnico No. RT-199, Convênio Unicamp-Telebrás, contrato No. 387/90, Transmissão Digital PT 01, Julho/90.