

Avaliação da Probabilidade de Perda de Células em Redes ATM utilizando o Processo de Poisson Modulado por Markov*

Ricardo José Paiva de Britto Salgueiro
Universidade Federal de Pernambuco
Departamento de Eletrônica e Sistemas
rjpbs@di.ufpe.br

José Augusto Suruagy Monteiro
Universidade Federal de Pernambuco
Departamento de Informática
suruagy@di.ufpe.br

Sumário

Em Redes Digitais de Serviços Integrados de Faixa Larga (RDSI-FL) utilizando o Modo de Transferência Assíncrono (ATM), uma das questões principais é a determinação da probabilidade de perda de células (pacotes de comprimento fixo). Baseado neste parâmetro é que os mecanismos da rede devem decidir pela aceitação ou não de uma nova chamada numa determinada rota.

Esta probabilidade de perdas pode ser obtida através de métodos analíticos ou de simulação. Dentre os métodos analíticos propostos na literatura, encontra-se o Processo de Poisson modulado por Markov (MMPP — Markov Modulated Poisson Process). Enquanto que outros métodos produzem resultados apenas para a multiplexação de fontes homogêneas, o MMPP é capaz de obter resultados também para fontes heterogêneas.

Neste artigo apresentamos o MMPP e o algoritmo para obtenção da distribuição do tamanho do sistema para filas MMPP/G/1 e MMPP/D/1, a partir da qual obtemos a probabilidade de perdas de células em uma fila MMPP/D/1/k através de aproximações. Os resultados obtidos para fontes homogêneas são comparados com resultados obtidos através de outros métodos. Os resultados obtidos levam a crer que o MMPP seja um modelo pessimista, o que pode ser uma vantagem pois introduz uma margem de segurança na alocação de largura de faixa. Em trabalhos futuros, investigaremos se esta é realmente uma característica do modelo ou se foi devida a imprecisões dos parâmetros utilizados.

*Esse trabalho foi desenvolvido com o apoio da FACEPE, do CNPq e da CAPES.

Abstract

In Broadband Integrated Services Digital Networks (BISDN) implemented using the Asynchronous Transfer Mode (ATM), one of the main issues is the evaluation of cell loss probabilities. A new call will be accepted or rejected in a given path depending on this parameter.

Cell loss probabilities can be obtained either through analytic methods or simulation. Among the analytical methods proposed in the literature is the Markov Modulated Poisson Process (MMPP). While other methods produce results only for homogeneous sources, MMPP is able to obtain results also for heterogeneous ones.

In this article, we present MMPP and the algorithm which obtains system length distributions for MMPP/G/1 and MMPP/D/1 queues. From these, through approximations, we obtain the loss probabilities for a MMPP/D/1/k queue. We compare the results obtained for homogeneous sources with those obtained with other methods. From the results MMPP seems to be a pessimistic model, which can turn out to be an advantage since it introduces a safety margin in the bandwidth allocation. In future works, we plan to investigate whether this is a model characteristic or just due to imprecisions in the parameters used.

1 Introdução

O alto índice de crescimento da indústria eletrônica ocorrido neste fim de século, ocasionou uma abrupta evolução nos sistemas de telecomunicações. Equipamentos como telefones, computadores, telex, fac-símile, etc., popularizaram-se e problemas de custos, manutenção, supervisão e gerência de redes diferentes para suprir tais serviços tornaram-se intoleráveis. Neste contexto, surgiu a Rede Digital de Serviços Integrados (RDSI), que trouxe como proposta a integração dos serviços de voz, dados e imagens.

RDSI-FL é o nome dado para as RDSI de altas velocidades (faixa larga). Estas altas velocidades são conseguidas através do uso de cabos de fibra óptica que permitem comunicações em taxas bem maiores que os outros tipos de cabos. A denominação faixa larga se refere a redes com taxas de transmissão superiores a 2 Mbps.

O Comitê Consultivo Internacional de Telefonia e Telegrafia (CCITT) escolheu o Modo de Transferência Assíncrono (*Asynchronous Transfer Mode — ATM*) como solução para os problemas de comutação e multiplexação da RDSI-FL [cci88].

O ATM é estruturado em unidades de informações de tamanho fixo denominadas de células, compostas por 48 bytes de dados e por 5 bytes de cabeçalho que contém os bits de endereçamento e controle. No campo do cabeçalho estão contidas todas as informações necessárias para o transporte ao destino desejado, enquanto que no

campo de informação estão contidos os bits de mensagem do usuário e informações que possibilitam a sua reconstituição.

Em redes ATM serão transportadas desde dados de telemetria, que requer baixas taxas de transmissão e chamadas de curta duração, até vídeo de alta definição, que necessita de altas taxas de transmissão e pode ter a duração de várias horas.

Por outro lado, alguns destes tráfegos são periódicos, outros transmitem em rajadas, e outros ainda transmitem em taxas variadas. Assumindo tráfego em rajadas (*bursty*), podemos caracterizá-lo através dos seguintes parâmetros: taxa de pico (B_p), a taxa média (B_m), o nível de explosividade (*burstiness* — b), carga média oferecida e duração média da rajada.

A taxa de pico é a taxa máxima de transmissão de células. O nível de explosividade pode ser definido como a relação entre a taxa de pico e a taxa média, isto é $b = B_p/B_m$ [Mon90]. A carga média oferecida é definida como $A_o = N(B_{mux}/B_p)(1/b)$, onde B_{mux} é a taxa de multiplexação. E a duração do pico é o período em que a taxa de transmissão é igual a de pico.

A utilização do ATM permite uma alocação dinâmica da banda passante, e ganhos estatísticos na multiplexação [Min89]. Em contraste, porém, a multiplexação estatística possibilita o surgimento de problemas de congestionamento [CP90], [Boy90], [GRF89], [MGF91]. O congestionamento aparece se, em um determinado momento, um dado serviço necessita de uma certa largura de faixa, mas, devido à dinamicidade da alocação, não haja disponibilidade.

Mecanismos de controle de congestionamento devem ser implementados para que a qualidade do sistema seja garantida. Podemos dizer que o controle de congestionamento é o preço que deve ser pago para obtermos ganhos estatísticos na multiplexação.

As funções de controle de congestionamento podem ser classificadas em quatro grupos [CP90]: controle de admissão, gerenciamento da memória e da fila, controle de tráfego e controle reativo.

O controle de admissão verifica se uma nova chamada pode ser admitida ou não, isto é, se a alocação da banda passante solicitada por um determinado serviço, pode ser satisfeita, de modo a garantir o grau de serviço esperado. O gerenciamento da memória e da fila responsabiliza-se pela ordenação das células a serem processadas. O controle de tráfego que também é chamado de *policimento*, observa se as fontes estão obedecendo aos parâmetros estabelecidos durante a alocação. O policimento é um controle preventivo. Por fim, o controle reativo tem a função de aliviar o congestionamento porventura existente.

A eficiência do controle de admissão é o ponto de partida para um bom funcionamento do sistema de controle de congestionamento. Para que o controle de admissão funcione de forma eficiente, é imprescindível que as alocações das bandas passantes sejam feitas também de forma eficiente, pois se para um determinado serviço fizer-

mos uma alocação maior do que a necessária, correremos o risco de impossibilitar indevidamente a aceitação de outros.

A alocação de capacidade para cada serviço deve satisfazer a um determinado grau de serviço desejado (GDS). O GDS em uma rede com multiplexação estatística pode ser expresso, entre outros parâmetros pela probabilidade máxima de perdas de células. Para avaliarmos o desempenho de uma rede em função da probabilidade de perdas de células, número de fontes multiplexadas e comprimento da fila, utilizamos técnicas de modelagem.

Há duas maneiras de modelar um sistema: uma é através de simulação e a outra é através de modelos analíticos.

O uso de simulação para modelagem de sistemas de filas para RDSI-FL, onde a probabilidade de perdas de células é da ordem de 10^{-9} , é impraticável. O tempo de CPU requerido chega a valores exorbitantes. A implementação de modelos analíticos que apresentem uma boa precisão é portanto, imprescindível.

Dois modelos analíticos podem ser destacados: o UAS (*Uniform Arrival and Service Model*) [AMS82], [Tuc88] e o Processo de Poisson Modulado por Markov (*Markov Modulated Poisson Process* — MMPP).

No UAS, a taxa de chegada de células ao sistema depende do número de fontes ativas e a taxa de serviço é constante. O tempo de execução deste modelo depende apenas do número de fontes consideradas e não da probabilidade de perdas de células a ser obtida. Este modelo é válido apenas para fontes homogêneas. Porém, no caso de fontes heterogêneas, podemos utilizar os resultados obtidos para fontes homogêneas para cada uma das fontes envolvidas e a partir deles determinar, através de modelos aproximados, quais são as combinações de número de fontes de cada tipo que podem ser multiplexadas.

O UAS foi utilizado por alguns autores, [MGF91], [Mon91], e apresentou bons resultados quando comparado com simulação.

No MMPP [HL86], [Neu89], [FM91] as chegadas ocorrem segundo um processo de Poisson com taxa variável. As vantagens de sua utilização são as seguintes: o MMPP é analiticamente tratável; permite a caracterização de tráfegos agregados ou não de fontes homogêneas ou heterogêneas (caso típico em redes ATM); a superposição de dois ou mais MMPP's dá origem a um novo MMPP; mesmo que haja um grande aumento na quantidade de fontes o seu nível de complexidade pode ser mantido através de uma combinação das técnicas de superposição de tráfegos agregados para gerar um MMPP e superposição de MMPP's; é implementável computacionalmente; algoritmos recentes [FM91], [Luc91] reduziram bastante a carga computacional exigida (tempo de CPU *versus* memória).

Neste trabalho analisamos o grau de serviço em multiplexador estatístico em função da probabilidade de perdas de células, comprimento da fila e carga oferecida por um tráfego modelado através do MMPP. Mostramos resultados obtidos através

da implementações do algoritmo apresentado em [FM91] para uma fila MMPP/G/1. Estes resultados são aproximados para o caso de filas finitas, obtendo assim as probabilidades de perda de células que são comparadas com valores obtidos através de simulação e através do modelo UAS. Aqui neste artigo, nos restringiremos à multiplexação de fontes homogêneas representadas através de um MMPP de dois estados.

Na seção 2 descrevemos o Processo de Poisson Modulado por Markov. Na seção 3 apresentamos uma síntese da teoria para uma fila MMPP/G/1. Na seção 4 comentamos os resultados obtidos nas implementações dos algoritmos para obtenção da distribuição do tamanho do sistema representado por uma fila MMPP/D/1 e da probabilidade de perdas de células em uma fila MMPP/D/1/k. E finalmente, na seção 5 apresentamos nossas conclusões finais e sugestões para trabalhos futuros.

2 O Processo de Poisson Modulado por Markov

O Processo de Poisson Modulado por Markov (MMPP) é um Processo de Poisson Duplamente Estocástico em que a taxa de chegada é dada por $\lambda^*[J(t)]$, onde $J(t)$, $t \geq 0$, é um Processo de Markov irreduzível com m estados. Ou, em outras palavras, a taxa média de chegadas é uma variável aleatória cujo valor é função do estado de uma cadeia de Markov irreduzível e contínua no tempo. Esta cadeia de Markov modula a taxa média do processo de Poisson, daí o nome de processo de Poisson modulado por Markov. A figura 1 mostra uma cadeia de Markov representando um MMPP de dois estados.

O MMPP fica completamente parametrizado através do: Vetor de taxas de chegada; Gerador infinitesimal do processo de Markov; e, Vetor de probabilidades iniciais do processo de Markov.

Do vetor de taxas de chegadas origina-se a matriz de taxas de chegadas.

Para um Processo de Markov com m estados, a matriz e o vetor de chegadas são definidos como:

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m) \quad (1)$$

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m) \quad (2)$$

Também chamado de matriz de taxas de transições, o gerador infinitesimal de um Processo de Markov com m estados é dado por:

$$Q = \begin{bmatrix} -\sigma_1 & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & -\sigma_2 & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2} & \cdots & \sigma_m \end{bmatrix}, \quad (3)$$

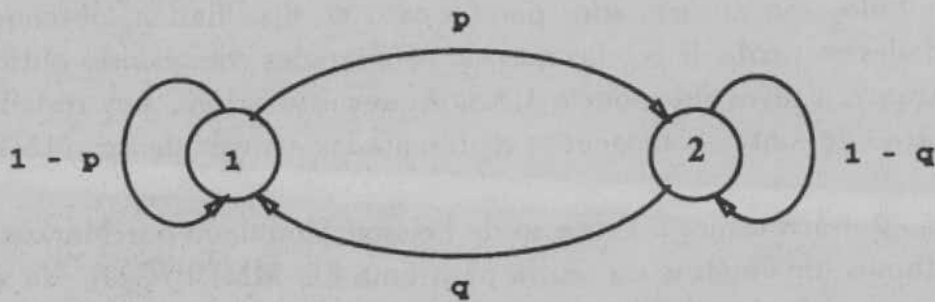


Figura 1: Um MMPP de 2 estados

onde

$$\sigma_i = \sum_{j=1; j \neq i}^m \sigma_{ij}.$$

O gerador infinitesimal de um processo de Markov contínuo no tempo é definido e bem discutido em [Ç75]. Neste trabalho estamos assumindo o prévio conhecimento por parte do leitor.

O vetor estacionário da cadeia de Markov (π) é definido, então, como:

$$\pi \mathbf{Q} = 0, \quad \pi \mathbf{e} = 1, \quad (4)$$

onde $\mathbf{e} = (1, 1, \dots, 1)^T$ é um vetor coluna de comprimento m .

O vetor de probabilidade inicial pode ser escolhido como o vetor estacionário da cadeia de Markov, π . Em [FM91] e [Neu89] são apresentadas outras formas de escolhas.

Baiocchi et al, em [BMAW91], estudaram a probabilidade de perdas de células em um multiplexador estatístico com entrada composta por N fontes independentes, homogêneas, do tipo rajadas, modeladas por um MMPP. Como em fontes do tipo rajadas cada uma das fontes se alternam em períodos de silêncio e períodos de transmissão em taxa de pico (ativo), eles construíram uma cadeia de Markov de dois estados onde um dos estados representa a situação em que o número de fontes ativas supera a capacidade do sistema assumindo a alocação de pico, enquanto que o outro representa o caso em que o número de fontes ativas está abaixo. A partir da caracterização de uma única fonte pela sua taxa de pico, taxa média de transmissão

e comprimento médio do surto eles desenvolveram uma técnica para superposição de fontes homogêneas e obtenção dos parâmetros do MMPP (matriz de taxas de chegadas e matriz de taxas de transições). Estes parâmetros foram utilizados para a obtenção da probabilidade de perdas de células em uma fila mmpp/D/1/k.

O primeiro grande passo para a utilização do MMPP como modelo analítico capaz de caracterizar tráfego de voz e dados foi dado por Heffes e Lucantoni [HL86]. Após haverem obtido os parâmetros do MMPP para fontes de voz superpostas, eles incorporaram o fluxo de dados ao modelo. Para isto, primeiro foi considerado que o fluxo de dados seguia a distribuição de Poisson. Depois, eles notaram que a superposição de um processo de Poisson de taxa λ_d a um MMPP de dois estados com parâmetros $\lambda_1, \lambda_2, r_1$ e r_2 é novamente um MMPP de dois estados com parâmetros $\lambda_1 + \lambda_d, \lambda_2 + \lambda_d, r_1, r_2$.

Um outro modo de modelarmos fontes heterogêneas é utilizarmos a técnica de superposição dos parâmetros do MMPP mostrada em [FM91] e [Neu89].

A superposição de MMPP's resulta em um novo MMPP. Os parâmetros deste novo MMPP, i.e., a matriz de taxas de chegadas Λ e o gerador infinitesimal Q , são determinados pela soma de Kronecker dos MMPP's superpostos.

A soma de Kronecker, simbolizada por \oplus , é definida como sendo

$$A \oplus B = (A \otimes I_B) + (I_A \otimes B),$$

onde \otimes simboliza o produto de Kronecker e I_A e I_B são matrizes identidade de mesma ordem que as matrizes A e B respectivamente.

O produto de Kronecker, por sua vez, é definido como sendo

$$C \otimes D = \begin{bmatrix} c_{11}D & c_{12}D & \dots & c_{1m}D \\ \vdots & \vdots & \vdots & \\ c_{n1}D & c_{n2}D & \dots & c_{nm}D \end{bmatrix}.$$

Após o exposto, podemos, então, definir a equação do gerador Q e da matriz de taxas Λ do MMPP resultante da superposição de n MMPP's que apresentam geradores individuais Q_i e matrizes de taxa Λ_i como sendo

$$Q = Q_1 \oplus Q_2 \oplus \dots \oplus Q_n, \quad (5)$$

$$\Lambda = \Lambda_1 \oplus \Lambda_2 \oplus \dots \oplus \Lambda_n. \quad (6)$$

Como já deve ter sido notado, há um acréscimo na dimensão da matriz resultante. Isto reflete-se num cuidado que devemos tomar ao caracterizarmos o tráfego de um sistema através de MMPP's superpostos, pois é possível chegarmos a matrizes de tamanhos inviáveis. Para atenuar este problema podemos optar por combinação de técnicas de superposição de fontes para gerar um MMPP e a superposição de MMPP's propriamente dita.

Se Q_i e Λ_i são matrizes de dimensão k_i , o acréscimo na dimensão da matriz é expresso pela seguinte relação:

$$k = \prod_{i=1}^n k_i,$$

onde k é a dimensão das matrizes superpostas Q e Λ .

Quando a superposição envolve apenas MMPPs de dois estados que possuem os mesmos parâmetros, a complexidade é bastante reduzida.

Sejam Q_n e Λ_n , respectivamente, o gerador infinitesimal e a matriz de taxas de chegadas do MMPP resultante da superposição de n processos idênticos que apresentam geradores Q e matrizes de taxas de chegadas Λ , tal que

$$Q = \begin{bmatrix} -\sigma_1 & \sigma_1 \\ \sigma_2 & -\sigma_2 \end{bmatrix} \quad e \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

então

$$\begin{aligned} (Q_n)_{i,i} &= -i\sigma_1 - (n-i)\sigma_2, & 0 \leq i \leq n \\ (Q_n)_{i,i-1} &= i\sigma_1, & 1 \leq i \leq n \\ (Q_n)_{i,i+1} &= (n-i)\sigma_2, & 0 \leq i \leq n-1 \\ 0, & & \text{caso contrário,} \end{aligned}$$

e

$$\Lambda_n = \text{diag}(i\lambda_1 + (n-i)\lambda_2, 0 \leq i \leq n). \quad (7)$$

Pelo que foi exposto, pode ser visto que o acréscimo na dimensão da matriz passa, neste caso, a ser de apenas uma unidade a cada superposição. Se compararmos com o caso geral, onde devemos aplicar a soma de Kronecker, notamos que esta particularidade viabiliza a superposição para um número consideravelmente maior de MMPP's. A dimensão das matrizes Q_n e Λ_n podem ser representadas por

$$\dim(Q_n) = \dim(\Lambda_n) = n + 1.$$

3 A fila MMPP/G/1

Utilizamos a denominação "fila MMPP/G/1" para uma fila que represente um sistema que possui apenas 1 servidor com distribuição de tempo de serviço geral e com o processo de chegada descrito por um MMPP. Se a fila possui um comprimento k , dizemos que se trata de uma fila MMPP/G/1/k.

Na seção anterior comentamos o trabalho publicado por Heffes e Lucantoni [HL86] que, além de tratar de superposições de tráfegos agregados de voz e dados, faz uma análise da desempenho de uma fila MMPP/G/1. Anteriormente, Lucantoni e

Ramaswami [LR85] desenvolveram um algoritmo para resolver equações matriciais não lineares aplicadas a uma fila PH/G/1, onde PH simboliza uma distribuição de probabilidade do tipo fase (*phase* - PH). Este algoritmo também poderia ser aplicado a uma fila MMPP/G/1. Depois, novos resultados para a fila MMPP/G/1 foram conseguidos por Ramaswami [Ram88], Neuts [Neu89], Lucantoni, Meier e Neuts [LMHN90] e, finalmente, Lucantoni [Luc91]. Fischer e Meier [FM91] fizeram uma coletânea destes últimos resultados e aqui, neste nosso trabalho, nos limitaremos a definições e breves exposições de valores de interesses essenciais ao entendimento deste algoritmo para a obtenção da distribuição do comprimento de uma fila MMPP/G/1, ou ao cálculo das probabilidades de perdas (de células) para uma fila MMPP/D/1/k. O leitor interessado em mais detalhes deve consultar [FM91], ou examinar diretamente as demais referências supracitadas.

3.1 A matriz de probabilidades de transição

O processo de renovação de Markov embutido, obtido considerando o comprimento da fila e o estado do MMPP, tem uma matriz de probabilidades de transição do tipo M/G/1. Em [Neu89] é feito um estudo detalhado da estrutura estocástica de tais matrizes.

A matriz de probabilidade de transição é dada por

$$\tilde{Q}(x) = \begin{bmatrix} \tilde{B}_0(x) & \tilde{B}_1(x) & \tilde{B}_2(x) & \cdots \\ \tilde{A}_0(x) & \tilde{A}_1(x) & \tilde{A}_2(x) & \cdots \\ 0 & \tilde{A}_0(x) & \tilde{A}_1(x) & \cdots \\ 0 & 0 & \tilde{A}_0(x) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad x \geq 0, \quad (8)$$

onde para $n \geq 0$; $\tilde{A}_n(x)$ e $\tilde{B}_n(x)$ são matrizes $m \times m$.

Os elementos de $\tilde{A}_n(x)$ são definidos como a probabilidade de que dada uma partida no instante $t = 0$, que deixou o processo de chegadas no estado i e pelo menos um usuário no sistema, a próxima partida ocorra em um instante $t \leq x$ com o processo de chegada no estado j , e que durante este serviço tenham havido n chegadas.

Pela sua definição, verificamos que no intervalo entre a primeira partida e a próxima há a ocorrência de n chegadas. Se denotarmos por $P(n, t)$ a função de contagem para estas n chegadas e por $\tilde{H}(t)$ a distribuição de probabilidade do tempo de serviço para tal sistema, a matriz $\tilde{A}_n(x)$ pode ser expressa como

$$\tilde{A}_n(x) = \int_0^x P(n, t) d\tilde{H}(t), \quad n \geq 0, x \geq 0. \quad (9)$$

Por sua vez, os elementos de $\tilde{\mathbf{B}}_n(x)$ são definidos como a probabilidade de que dada uma partida no instante $t = 0$, que deixou o processo de chegadas no estado i e o sistema vazio, a próxima partida ocorra em um instante $t \leq x$ com o processo de chegadas no estado j , deixando n usuários no sistema.

Utilizando as transformadas de Laplace-Stieltjes¹ e transformadas z podemos definir as matrizes

$$\mathbf{A}_n(s) = \int_0^\infty e^{-sx} d\tilde{\mathbf{A}}_n(x), \quad \mathbf{B}_n(s) = \int_0^\infty e^{-sx} d\tilde{\mathbf{B}}_n(x),$$

$$\mathbf{A}(z, s) = \sum_{n=0}^\infty \mathbf{A}_n(s) z^n, \quad \mathbf{B}(z, s) = \sum_{n=0}^\infty \mathbf{B}_n(s) z^n,$$

e as matrizes

$$\mathbf{A}_n = \mathbf{A}_n(0) = \tilde{\mathbf{A}}_n(\infty), \quad \mathbf{B}_n = \mathbf{B}_n(0) = \tilde{\mathbf{B}}_n(\infty),$$

$$\mathbf{A} = \mathbf{A}(1, 0), \quad \mathbf{B} = \mathbf{B}(1, 0).$$

Conforme mostrado em [FM91] a seguinte relação é obedecida

$$\mathbf{B}_n = (\Lambda - \mathbf{Q})^{-1} \Lambda \mathbf{A}_n. \tag{10}$$

3.2 A distribuição do comprimento da fila no instante de partida

Através da cadeia de Markov embutida nos instantes de partida podemos obter o comprimento da fila nestes mesmos instantes. Se denotarmos como $\tilde{\mathbf{Q}}(\infty)$ a matriz de probabilidades de transição para esta cadeia, podemos definir o seu vetor estacionário como $\mathbf{x} = (x_0, x_1, \dots)$ e o seguinte sistema de equações pode ser obtido

$$\mathbf{x}_i = \mathbf{x}_0 \mathbf{B}_i + \sum_{v=1}^{i+1} \mathbf{x}_v \mathbf{A}_{i+1-v}, \quad i \geq 0 \tag{11}$$

onde os elementos $x_{i,j}$ são a probabilidade que uma partida ocorra deixando o sistema anteriormente com i usuários e o MMPP no estado j .

O vetor \mathbf{x}_i é representado por

$$\mathbf{x}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,m})$$

Apresentamos no apêndice A o algoritmo, desenvolvido por Ramaswami [Ram88], utilizado para determinar os vetores \mathbf{x}_i . Uma condição inicial deste algoritmo é o cálculo do vetor \mathbf{x}_0 . Antes, porém, faz-se necessário o cálculo da matriz \mathbf{G} que é definida de tal modo que os elementos $G_{i,j}$ sejam a probabilidade que um período

¹A transformada de Laplace-Stieltjes é a transformada de Laplace onde, no lugar de se usar a integral de Riemann, utiliza-se a integral de Stieltjes.

ocupado comece com o MMPP no estado i e termine com o MMPP no estado j . Pode ser mostrado que a matriz G é a raiz da equação

$$G = \int_0^{\infty} \exp(Q - \Lambda + \Lambda G) x \, d\tilde{H}(x). \quad (12)$$

O vetor estacionário da matriz G satisfaz às seguintes equações

$$g \cdot G = g, \quad g \cdot e = 1.$$

O vetor x_0 pode agora ser definido [LMHN90] por

$$x_0 = \frac{1 - \rho}{\lambda_{tot}} \cdot g(\Lambda - Q), \quad (13)$$

onde

$$\begin{aligned} \lambda_{tot} &= \pi \lambda, \\ \rho &= h \lambda_{tot} \text{ é a intensidade de tráfego, e,} \\ h &\text{ é o tempo médio de serviço.} \end{aligned}$$

3.3 A distribuição do tamanho do sistema em um instante arbitrário

A distribuição do tamanho do sistema em um instante arbitrário é obtida através do comprimento da fila nos instantes de partida. Definindo $y_{i,j}$ como a probabilidade de que, em um instante qualquer, haja i usuários no sistema e o MMPP esteja no estado j ; e o vetor

$$y_i = (y_{i,1}, y_{i,2}, \dots, y_{i,m}),$$

podemos mostrar que

$$y_i = (y_{i-1} \Lambda - \lambda_{tot}(x_{i-1} - x_i)) (\Lambda - Q)^{-1}. \quad (14)$$

O vetor y_0 é dado por

$$y_0 = (1 - \rho)g \quad (15)$$

e a distribuição de probabilidade incondicional do tamanho do sistema em um instante arbitrário pode, finalmente, ser expressa por

$$p_i = y_i \cdot e. \quad (16)$$

Observe que, como p_i representa a probabilidade incondicional de existir i usuários no sistema, a probabilidade de perdas de células (P) em uma fila MMPP/D/1/k pode ser aproximada por

$$P = 1 - \sum_{n=0}^i p_n \quad (17)$$

Esta aproximação nos leva a maiores imprecisões à medida que o número de usuários aumenta, pois p_i decresce com o aumento de i e, como pode ser facilmente visto, quanto menor for $\sum_{n=0}^i p_n$ maior será o erro introduzido pela aproximação.

Como consequência, podemos esperar que a implementação desta aproximação dê bons resultados apenas para a obtenção de probabilidade de perdas pequenas. Como em redes ATM estamos interessados em baixas probabilidade de perdas esta aproximação pode ser considerada satisfatória.

4 Resultados Numéricos

O algoritmo apresentado no apêndice A considera que os parâmetros do MMPP foram, de algum modo, previamente obtidos. Em [BMAW91], comentado na seção 2, é apresentado um método que a partir dos parâmetros característicos do tráfego tais como carga oferecida, taxa de pico, nível de explosividade e taxa de multiplexação, chega-se à determinação dos parâmetros do MMPP que, por sua vez, são utilizados para obtenção da probabilidade de perdas de células. Como foi visto, este método aplica-se apenas ao caso de fontes homogêneas e cadeia de Markov com dois estados.

Aqui neste artigo, fizemos uma concatenação do método para obtenção dos parâmetros do MMPP, descrito em [BMAW91], com o algoritmo contido no apêndice A, aproximando-o para o caso de filas finitas. Assim sendo, a validação do algoritmo, ou de nossas implementações, deve envolver uma avaliação na procedência das possíveis imprecisões, ou, de outra forma, deve envolver testes e comparações com resultados que possam ser tomados como referência, de modo que não só os possíveis erros resultantes do próprio algoritmo possam ser identificados, mas também, os erros derivados da técnica de obtenção dos parâmetros.

Nesta seção mostramos os resultados obtidos com o MMPP comparando-o, inicialmente, com os resultados provenientes de simulações do modelo do MMPP de dois estados para diversos parâmetros, depois comparando-o com resultados obtidos através do modelo UAS, e, por fim, comparando-o com os resultados apresentados em [BMAW91].

Na figura 2 mostramos um gráfico das probabilidades de perdas de células (P) em função da carga de trabalho oferecida (A_o), obtidas através do MMPP, de simulação e do UAS. Utilizamos fontes homogêneas com taxa de pico de 10 Mbits/s, taxa de multiplexação de 150 Mbits/s, nível de explosividade igual a 10, comprimento médio do surto de 100 células e comprimento da célula igual a 53 Bytes.

Comparando a curva do MMPP com a curva da simulação verificamos que elas se interceptam nas proximidades de $A_o = 0.3$. Se prolongássemos a curva da simulação para pontos referentes a valores de $A_o < 0.3$ notaríamos destacadas duas regiões.

Na primeira região, os valores para simulação não foram obtidos, pois, como já

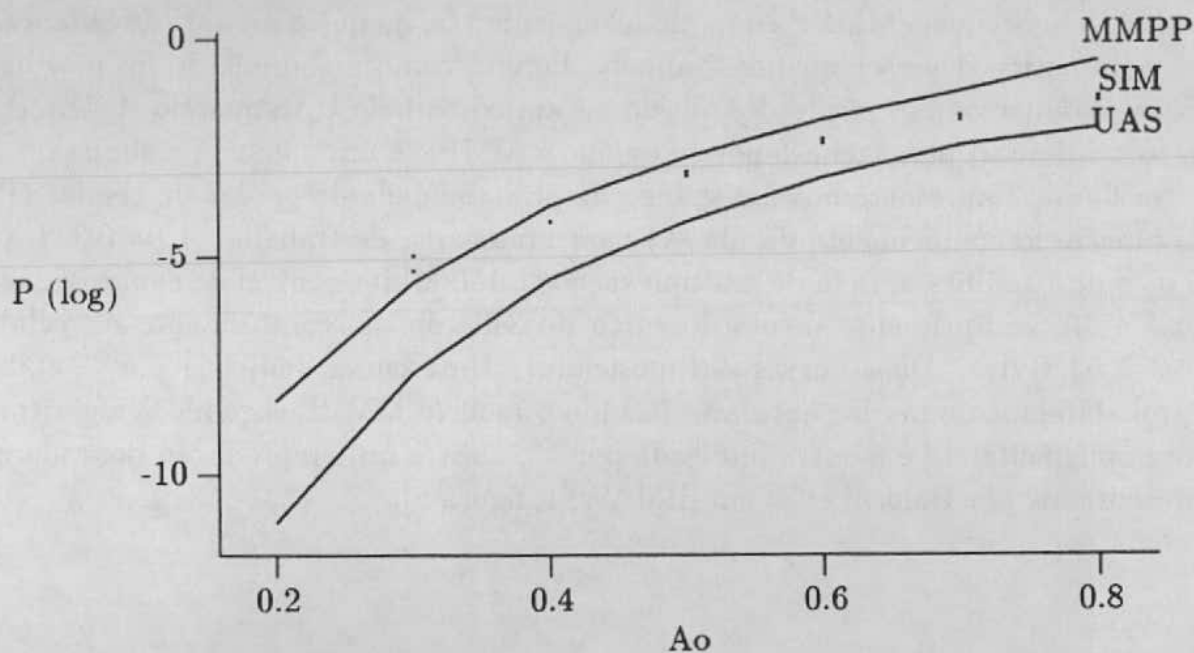


Figura 2: Comparação das probabilidades de perdas de células (P) em função da carga de trabalho (A_o) obtidas através do MMPP, simulação e UAS.

foi mencionado na seção 1, para obtenção de pequenas probabilidades de perdas de células é desaconselhável o uso de simulação. Note que, por exemplo, para $A_o = 0.2$, temos aproximadamente 27 fontes transmitindo em uma taxa média de 1 MBps ($B_m = B_p/b$), multiplexadas em 150 Mbps. Isto, sem dúvida, resulta numa baixa probabilidade de perdas. Para valores como estes torna-se impraticável o uso de simulação.

Na segunda região, observamos o processo inverso. As curvas partem de um ponto de igualdade distanciando-se com o incremento da carga de trabalho. Isto pode ser reflexo do aumento da imprecisão da aproximação – discutida na seção 3.3 – adotada para o cálculo da probabilidade de perdas em uma fila MMPP/D/1/k. Perceba que à medida que a carga de trabalho aumenta os resultados da simulação se aproximam do UAS.

Comparando, desta vez, as curvas do MMPP e do UAS notamos que elas apresentam praticamente o mesmo desenho, mas com o MMPP apresentando sempre como resposta probabilidades de perdas maiores. Alteramos, respectivamente, os valores do comprimento médio da rajada para 340 células e 1000 células e observamos a mesma relação.

Considerando que o UAS é um modelo que apresentou bons resultados quando comparado com simulação [MGF91, Mon91], portanto já validado, poderíamos che-

gar à conclusão que o MMPP é um modelo pessimista, ou que o método de obtenção dos parâmetros deve ser melhor avaliado. Porém, como na simulação fizemos uso destes parâmetros e os resultados obtidos se aproximam do UAS quando A_o aumenta, torna-se mais provável a hipótese de que o MMPP é um modelo pessimista.

Na figura 3 apresentamos um gráfico da probabilidade de perdas de células (P) em função do comprimento da fila (k) para uma carga de trabalho $A_o = 0.2$, taxa de pico de 10 Mbits/s, taxa de multiplexação de 150 Mbits/s, nível de explosividade igual a 10, comprimento médio do surto de 340 células, comprimento da célula igual a 53 Bytes. Duas curvas são mostradas. Uma curva, indicada por *, exibe a probabilidade de perdas obtida utilizando o modelo MMPP segundo o algoritmo que implementamos e a outra, indicada por **, mostra uma reprodução dos valores apresentados por Baiocchi et al em [BMAW91, figura 8].

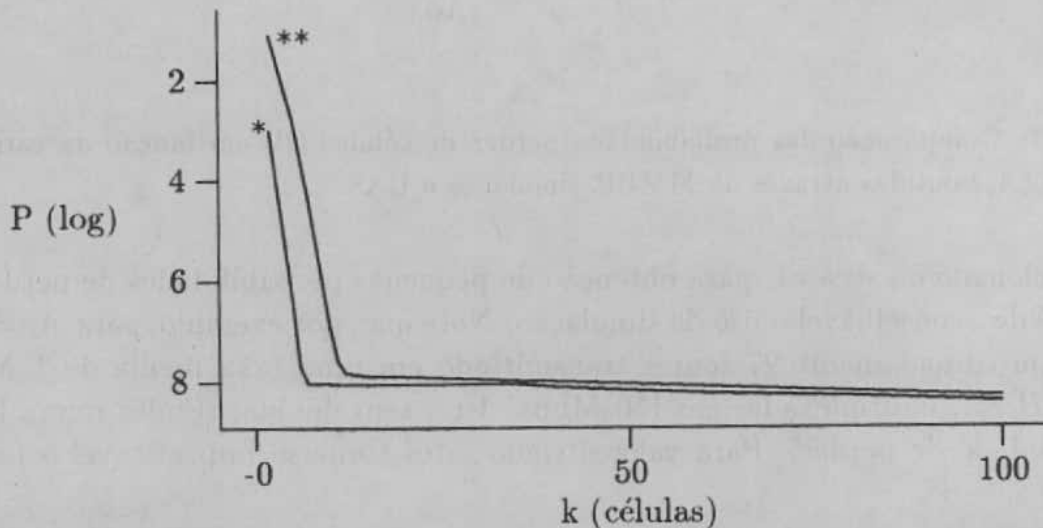


Figura 3: Comparação das probabilidades de perdas de células (P) em função do comprimento da fila (k) obtidas através de nossas implementações (*) e extraídas de [BMAW91][figura 8] (**).

Nossos resultados mostram probabilidades de perdas maiores para valores pequenos de k , porém para comprimentos maiores da fila, nossas implementações mostram resultados bem próximos. Poderíamos relacionar este acontecimento, com o erro que introduzimos na aproximação para o cálculo da probabilidade de perdas. Como para comprimentos de filas pequenos a probabilidade de perdas é grande, era de se esperar que em nossa implementação ocorressem imprecisões quando estivéssemos tratando de filas pequenas.

5 Conclusões

Em Redes ATM, uma das questões principais é a determinação da probabilidade de perda de células. Baseado neste parâmetro é que os mecanismos da rede devem decidir pela aceitação ou não de uma nova chamada numa determinada rota.

Esta probabilidade de perdas pode ser obtida através de métodos analíticos ou de simulação. Dentre os métodos analíticos propostos na literatura, encontra-se o Processo de Poisson modulado por Markov (MMPP — *Markov Modulated Poisson Process*).

Uma das principais vantagens na utilização do MMPP é a capacidade de obtenção de resultados para fontes heterogêneas (caso típico em redes ATM). Outros modelos, tais como o UAS, são apropriados apenas para fontes homogêneas. Porém, com o uso de aproximações, podemos obter resultados também para tráfegos de fontes heterogêneas.

Nossas implementações ficaram restritas a tráfegos de fontes homogêneas, porém mostramos o algoritmo para superposição de MMPP's e deixamos sua implementação como trabalho futuro. Devemos, então, na oportunidade, verificar o comportamento do MMPP em relação às aproximações do UAS.

Na análise de nossos resultados procuramos diferenciar duas possibilidades de erros. Uma devido ao método de obtenção dos parâmetros e outra devido ao próprio algoritmo implementado.

Algumas questões foram relacionadas aos resultados, tais como a impossibilidade de avaliação dos resultados através de simulação para baixas cargas de trabalho oferecida, imprecisão da aproximação que utilizamos para obtenção das perdas em filas finitas para altas cargas de trabalho e a hipótese do MMPP ser um modelo pessimista.

A comparação feita na figura 3, de certa forma, valida nossas implementações, pois partimos de um mesmo método de obtenção dos parâmetros e chegamos a valores próximos à medida que tratamos de menores probabilidades de perdas.

Já a comparação com o UAS nos mostra que deve existir alguma relação entre seus valores, trazendo a idéia de um algum pessimismo existente ou de algum erro introduzido pelos parâmetros utilizados. Mas o fato dos resultados obtidos através de simulação tender aos resultados obtidos por intermédio do UAS, à medida que aumentamos a carga oferecida, nos indica que devemos ter cautela antes de assumirmos que os valores dos parâmetros utilizados é o causador destas imprecisões.

Pelo exposto, podemos concluir que o MMPP deve ser um modelo pessimista, o que pode ser uma vantagem pois introduz uma margem de segurança na alocação da largura de faixa. Porém deixamos uma melhor investigação para trabalhos futuros.

A Algoritmo MMPP/G/1

Nesta Apêndice reproduziremos o algoritmo apresentado em [FM91] e que foi implementado para obter os resultados da seção 4.

1. Compute a matriz G ;

Defina

$$\begin{aligned} G_0 &= 0, \quad H_{0,k} = I, \quad k = 0, 1, 2, \dots, \\ \Theta &= \max_i [(\Lambda - Q)_{ii}], \\ \gamma_n &= \int_0^\infty e^{-\Theta x} \frac{(\Theta x)^n}{n!} d\hat{H}(x), \quad n = 0, 1, \dots, n^*, \end{aligned}$$

onde n^* é escolhido tal que $\sum_{k=0}^{n^*} \gamma_k > 1 - \varepsilon_1$, $\varepsilon_1 \ll 1$.

Caso o serviço seja determinístico a computação de γ_n é obtida pela equação recursiva:

$$\begin{aligned} \gamma_n &= \int_0^\infty e^{-\Theta x} \frac{(\Theta x)^n}{n!} \delta(h-x) dx \\ &= e^{-\Theta h} \frac{(\Theta h)^n}{n!} \end{aligned}$$

$$\begin{aligned} \gamma_0 &= e^{-\Theta h}, \\ \gamma_n &= \gamma_{n-1} \frac{\Theta h}{n}, \quad n \geq 1. \end{aligned}$$

Após estas definições segue a equação recursiva:

$$H_{n+1,k} = \left[I + \frac{1}{\Theta} (Q - \Lambda + \Lambda G_k) \right] H_{n,k}, \quad n = 0, 1, \dots, n^*,$$

$$G_{k+1} = \sum_{n=0}^{n^*} \gamma_n H_{n,k}.$$

O critério para o término da recursão é

$$\| G_{k-1} - G_k \| < \varepsilon_2 \ll 1.$$

Faça $G = G_{k+1}$.

2. Compute o vetor estacionário g que satisfaça

$$g G = g, \quad g e = 1.$$

3. Compute o vetor x_0

$$x_0 = \frac{1 - \rho}{\lambda_{tot}} \cdot g(\Lambda - Q).$$

4. Compute a distribuição do tamanho do sistema nos instantes de partidas

Neste passo é que se situam as mais recentes evoluções que tornaram mais viável a implementação computacional do MMPP. Ramaswami em [Ram88] desenvolveu o algoritmo para o cálculo do vetor x_i , Neuts em [Neu89] resumiu estes procedimentos, e Lucantoni em [Luc91] simplificou os cálculos das matrizes A_n .

A partir da equação 11 é possível chegar, como pode ser visto em [Ram88] e em [Neu89, teorema 3.2.5] na fórmula

$$x_i = [x_0 \bar{B}_i + \sum_{\nu=1}^{i-1} x_\nu \bar{A}_{i+1-\nu}] (I - A_1)^{-1}, \quad i \geq 1, \quad (18)$$

onde as matrizes \bar{A}_k e \bar{B}_k são definidas como

$$\bar{A}_k = \sum_{i=k}^{\infty} \bar{A}_i G^{i-k}, \quad \bar{B}_k = \sum_{i=k}^{\infty} \bar{A}_i G^{i-k}, \quad k \geq 0. \quad (19)$$

Uma observação importante é que quando $i \rightarrow \infty$, $\bar{A}_i \rightarrow 0$ e $\bar{B}_i \rightarrow 0$.

As matrizes \bar{A}_k e \bar{B}_k podem ser computadas escolhendo-se um índice máximo k suficientemente grande tal que $\bar{A}_k = 0$ e $\bar{B}_k = 0$ e implementando as fórmula inversas recursivas

$$\bar{A}_k = A_k + \bar{A}_{k+1} \quad \bar{B}_k = B_k + \bar{B}_{k+1}.$$

A equação $G = A_0 + \bar{A}_1 G$ pode ser usada para verificar se o índice máximo foi bem escolhido.

O algoritmo abaixo descreve o procedimento utilizado para obtenção das matrizes A_ν

$$A_\nu = \sum_{n=\nu}^{\infty} \gamma_n K_\nu^{(n)}, \quad \nu \geq 0, \quad (20)$$

$$\begin{aligned} K_0^{(0)} &= I, \\ K_\nu^{(0)} &= 0 && \nu \geq 1, \\ K_0^{(n)} &= K_0^{(n-1)} [\Theta^{-1}(Q - \Lambda) + I], && \nu \geq 0 \\ K_\nu^{(n)} &= K_\nu^{(n-1)} [\Theta^{-1}(Q - \Lambda) + I] + K_{\nu-1}^{(n-1)} \Theta^{-1} \Lambda, && n \geq \nu \geq 1, \\ K_\nu^{(n)} &= 0 && n < 0, \quad n < \nu. \end{aligned}$$

Na implementação da matriz \mathbf{A}_ν , a equação 20 deve ser truncada em um índice N que seja suficientemente grande. Para uma boa escolha pode-se optar por uma das duas formas apresentadas a seguir que resultar no menor valor para N :

(a) N é escolhido tal que

$$\sum_{n=0}^N \gamma_n \geq 1 - \varepsilon, \quad \text{com } \varepsilon < 10^{-8};$$

(b) ou N é escolhido tal que

$$\max_j \left[\sum_{\nu=0}^N (A_\nu e)_j - 1 \right] < \varepsilon.$$

Note que em (a) o N corresponde ao n^* já utilizado para a obtenção da matriz \mathbf{G} . Enquanto que em (b) estamos nos baseando no princípio de que a matriz $\mathbf{A} = \sum_{\nu=0}^{\infty} \mathbf{A}_\nu$ é estocástica.

5. Compute o vetor \mathbf{y}_0

$$\mathbf{y}_0 = (1 - \rho)\mathbf{g}.$$

6. Compute a distribuição do comprimento da fila em um tempo arbitrário de acordo com a equação 14

$$\mathbf{y}_i = (\mathbf{y}_{i-1}\Lambda - \lambda_{tot}(\mathbf{x}_{i-1} - \mathbf{x}_i))(\Lambda - \mathbf{Q})^{-1}.$$

Referências

- [AMS82] D. Anick, D. Mitra, e M. M. Sondhi. Stochastic theory of data-handling system with multiple sources. *Bell System tech.*, 68(8):1871 - 1894, 1982.
- [BMAW91] A. Baiocchi, N. B. Melazzi, A. Roveri, e R. Winkler. Modeling issues on an ATM multiplexer within a bursty traffic environment. *IEEE - INFOCOM 91*, 1991.
- [Boy90] P. Boyer. A congestion control for the ATM. Em *Proc. International Teletraffic Symposium*, Norristown, NJ, outubro 1990.
- [Ç75] E. Çinlar. *Introduction to Stochastic Process*. Englewood Cliffs, NJ, 1975.
- [cci88] *CCITT Recomendação I.121 Broadband aspect of ISDN*, Melbourne, 1988.
- [CP90] C. A. Cooper e K. I. Park. Toward a broadband congestion control strategy. *IEEE Network Magazine*, 4(3):18 - 23, maio 1990.
- [FM91] Wolfgang Fischer e K. S. Meier. The MMPP cookbook. INRS-Telecommunications, Verdun Quebec, and AT&T Bell Laboratories, Holmdel, NJ, 1991.
- [GRF89] G. Gallassi, G. Rigolio, e L. Fratta. ATM: Bandwith assignment and bandwidth enforcement polices. Em *Anais do GLOBECOM'89*, volume 3, páginas 1788 - 1793, Dalas, Texas, novembro 1989.
- [HL86] Harry Heffes e David M. Lucantoni. A Markov modulated characterization of packetized voice and data traffic related statistical multiplexer performance. *IEEE Journal on Selected Areas in Communications*, SAC - 4:856 - 868, 1986.
- [LMHN90] David M. Lucantoni, Kathleen S. Meier-Hellstern, e Marcel F. Neuts. A single server queue with server vacation and a class of non-renewal arrival processes. *Journal of Applied Probability*, páginas 676 - 705, setembro 1990.
- [LR85] David M. Lucantoni e V. Ramaswami. Efficient algorithms for solving the non-linear matrix equations arising in phase type queues. *Stochastic Models*, 1:29 - 51, 1985.

- [Luc91] David M. Lucantoni. New result on the single server queue with a batch markovian arrival process. *Stochastic Models*, 7(1):1 – 46, janeiro 1991.
- [MGF91] José Augusto Suruagy Monteiro, Mario Gerla, e Luigi Fratta. Statistical multiplexing in ATM networks. *Performance Evaluation*, 12:157 – 167, 1991.
- [Min89] Steven E. Minzer. Broadband ISND and assynchronous transfer mode (ATM). *IEEE Communication Magazine*, 27(9):17 – 24, setembro 1989.
- [Mon90] José Augusto Suruagy Monteiro. *Bandwidth Allocation in Broadband Integrated Services Digital Networks*. Tese de PhD, Universty of California, Los Angeles, 1990.
- [Mon91] José Augusto Suruagy Monteiro. Alocação de Capacidades e Controle de Tráfego em Redes MTA. Em 9ª SBRC, páginas 446 – 447, 1991.
- [Neu89] Marcel F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and their Application*. New York, 1989.
- [Ram88] V. Ramaswami. Stable recursion for the steady vectorfor markov chains of M/G/1 type. *Stochastic Models*, 4:183 – 188, 1988.
- [Tuc88] Roger C. F. Tucker. Accurate method for analysis of a packet-speech multiplexing with limited delay. *IEEE Transations on Comunnications*, 36(4):479 – 483, abril 1988.