

# Método de Análise de Tráfegos VoIP Sobrepostos

Marcos Eduardo Coelho Garcia, Marcus Pantoja da Silva e Leonardo de Sousa Mendes

**Resumo**— Neste trabalho é feita uma análise de um modelo aproximado para representar a sobreposição de tráfegos VoIP com características de períodos de fala e silêncio distintos. A análise é realizada tanto analiticamente como por simulação, utilizando parâmetros sistêmicos do ITU.

Para a análise do comportamento da fila utilizou-se o modelo de MMPP/M/1 que descreve de forma simples o comportamento do sistema. Os resultados dessa análise mostram que o modelo é adequado para o estudo desse tipo de rede, o que leva a um melhor dimensionamento da rede.

**Palavras-Chave**— VoIP, sobreposição de tráfegos, MMPP/M/1.

**Abstract**— This work is an analysis of an approximate model to represent the superposition of VoIP traffic with the features of different periods of silence and speak. The analysis is performed both analytically and by simulation, using systemic parameters of the ITU.

For the analysis of the performance of the queue is used the model of MMPP/M/1 that describes the behavior of a simple system. The results of this analysis show that the model is suitable for the study of this type of system, which leads to a better scaling of the network.

**Keywords**— VoIP, traffic superposition, MMPP/M/1.

## I. INTRODUÇÃO

A tecnologia de Voz sobre IP (VoIP) se tornou uma das aplicações mais comuns da Internet. Aplicações de voz exigem uma grande preocupação do gerente da rede, pois possuem requisitos de Qualidade de Serviço (QoS) mais rígidos. Para ajudar no dimensionamento confiável da rede, faz-se necessário um modelo de tráfego que consiga representar as características do comportamento do tráfego de voz [5]. À medida que o número de usuários simultâneos de VoIP vai aumentando, o número necessário de operações para encontrar os parâmetros de desempenho de um enlace aumenta. Os trabalhos [1] [7] tratam somente de sobreposição de tráfegos de voz que possuem todos os mesmos períodos de fala e silêncio, que podem ser chamados de pertencentes a mesma categoria de conversa. Neste trabalho é apresentado um método para aproximar o tráfego sobreposto de múltiplas fontes pertencentes a categorias de conversa diferentes e também apresenta o método de obter a probabilidade escalar do tamanho da fila e o tamanho médio de uma fila MMPP/M/1.

## II. CODIFICAÇÃO DE VOIP

Para serem processados, armazenados ou transmitidos digitalmente, os sinais de voz precisam ser primeiro digitalizados, isto é, precisam passar por uma conversão do tipo (A/D). Esta

conversão compreende a amostragem do sinal, a quantização das amostras e a representação (ou codificação) das amostras quantizadas por palavras binárias. O formato digital obtido com esse procedimento é denominado PCM (*pulse-code modulation*) [3]. Para sinais de voz de banda estreita (largura de banda menor que 4 kHz), normalmente utiliza-se uma taxa de amostragem de 8 kHz, o que implica na necessidade de um canal digital de 64 Kbps para a transmissão de cada voz [4].

De modo a reduzir o consumo de banda na transmissão dos sinais de voz, são empregadas técnicas de codificação (ou compressão) de voz. A compressão do sinal de voz é conseguida explorando e reduzindo as suas redundâncias temporais. As técnicas de codificação de voz são divididas em três classes: codificação de forma de onda, codificação paramétrica e codificação híbrida. A codificação de forma de onda de sinais de voz é baseada principalmente na predição linear e o esquema mais utilizado é o ADPCM (*Adaptive Differential Pulse-Code Modulation*). A classe de codificadores paramétricos mais utilizada é a dos *vocoders* LPC (*Linear Predictive Coding*). Os codificadores híbridos são aqueles que combinam características dos codificadores de forma de onda e dos *vocoders*. Atualmente a maioria dos codificadores híbridos utilizam o modelo de codificação CELP (*Code Excited Linear Predictive*) [4].

Os algoritmos de compressão de voz utilizam técnicas de supressão de silêncio. Esta técnica atua nas “pausas entre as falas”, provendo uma compressão adicional. Durante uma conversação normal, as falas ocorrem tipicamente durante 40% do tempo, o restante é silêncio. A supressão de silêncio tira proveito destas pausas (silêncio) e somente transmite durante as rajadas de voz. Durante os períodos de silêncio, outros pacotes de voz ou dados podem usar a banda passante; o total de redução prático dependerá do tipo de codificação adotada, chegando a valores 25% menores, no caso do G.729 [4].

Estes algoritmos de codificação de voz tem uma taxa de transmissão de bits fixa e um tempo fixo de empacotamento. A fonte é caracterizada por um fluxo constante de pacotes durante o período ativo (período *ON*) e nenhuma chegada durante o silêncio (período *OFF*). Durante o período *ON* a fonte envia pacotes em intervalos regulares de tamanho  $T$  (*Packetization Time*). A duração dos períodos ativos e inativos é geralmente estimada por distribuições exponenciais independentes com parâmetros  $\alpha$  e  $\beta$ . A média de duração do período *ON* é dada por  $T_{on} = \frac{1}{\alpha}$  e a média de duração do período *OFF* por  $T_{off} = \frac{1}{\beta}$ . A média de taxa de bits do modelo *ON-OFF* é igual à:

$$\bar{\lambda} = \frac{T_{on}}{(T_{on} + T_{off}) \cdot T} = \frac{\beta}{(\beta + \alpha) \cdot T}$$

Os pacotes de voz codificados são transmitidos pela internet através do protocolo UDP enquanto que os pacotes com

informações sobre controle de erros (*RTP Control Protocol*) são transportados pelo protocolo TCP.

Este trabalho concentra o estudo no codec, G.729 com três diferentes categorias de conversa: *Classical*, *Conversation* e *Reading*. Na Tabela I são mostrados os valores numéricos dos modelos utilizados.

TABELA I  
MODELOS DE APLICAÇÃO VOIP

Aplicação	$T_{on}$ SEG	$T_{off}$ SEG	Tamanho Pacotes BYTES	$\bar{\lambda}$ KBPS
G729C	0.352	0.65	70	6.3
G729V	7.24	5.69	70	10.2
G729R	3.23	0.41	70	16.2

### III. Markov Modulated Poisson Process - MMPP

O MMPP, como já diz o nome, é um processo de Poisson duplamente estocástico, em que a transição entre os estados segue uma cadeia de Markov. Cada estado possui uma taxa de chegadas que obedece a uma distribuição de Poisson. Tem sido muito utilizado para modelar os processos de chegada de pacotes de voz [1], pois ele trata efetivamente a variação das taxas de chegadas no tempo e captura a correlação das chegadas. Como será visto, é bastante simples tratar o MMPP matematicamente [6].

O tráfego unitário de VoIP pode ser modelado com uma MMPP de dois estados, estado *ON* e estado *OFF*. No estado *ON*, temos um fluxo de pacotes de intensidade  $\lambda$  e, no estado *OFF*, temos uma ausência de pacotes. Assume-se que todos os pacotes possuem o mesmo tamanho. A passagem dos estados são mostrados na cadeia de Markov na Figura 1. A taxa de nascimento  $\alpha$  e a taxa de morte  $\beta$  são dadas respectivamente por

$$\alpha = \frac{1}{T_{on}} \quad \text{e} \quad \beta = \frac{1}{T_{off}} \quad (1)$$

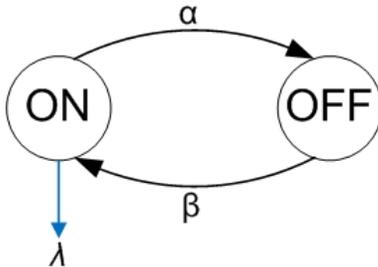


Fig. 1. Modelo de Tráfego ON-OFF.

A sobreposição de  $N$  tráfegos de VoIP com supressão de silêncio pode ser representada por um processo MMPP de  $N + 1$  estados, supondo que todos os tráfegos tenham as mesmas características de fala e silêncio [1], [6]. Cada estado representaria um número de vozes ativas, como pode ser visto na Figura 2. Quando o processo markoviano está em um estado  $i$ , as chegadas de pacotes ocorrem de acordo com um processo poissoniano de taxa  $\lambda_i$ . O MMPP é parametrizado por uma

matriz de transição  $Q$  e uma matriz diagonal  $\mathcal{L}$  contendo as taxas de chegadas para cada estado do processo markoviano. Estas duas matrizes são mostradas a seguir:

$$Q = \begin{bmatrix} -N\beta & N\beta & 0 & \dots \\ \alpha & -[\alpha + (N-1)\beta] & (N-1)\beta & \dots \\ 0 & 2\alpha & -[2\alpha + (N-2)\beta] & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \quad (2)$$

$$\mathcal{L} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N). \quad (3)$$

O vetor de probabilidade dos estados estacionários da cadeia de Markov é  $\pi$ , de forma que

$$\pi Q = 0, \quad \pi \varepsilon' = 1, \quad (4)$$

onde  $\varepsilon' = (1, 1, \dots, 1)^T$  é um vetor coluna de largura  $N$ .

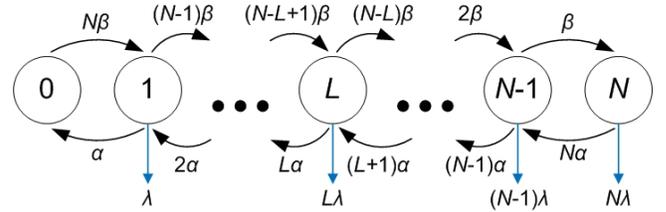


Fig. 2. Modelo MMPP.

### IV. REDUÇÃO DE $N + 1$ ESTADOS PARA SOMENTE 2

Com o intuito de simplificar o modelo, existem técnicas para reduzir os  $N + 1$  estados da MMPP em 2 estados, os estados *underload* (UL) e *overload* (OL). O estado UL representa os estados  $\{0, \dots, M\}$  e o estado OL representa os estados restantes  $\{M + 1, \dots, N\}$ , sendo  $M = \lfloor NP_{on} \rfloor$ . Juntos, estes dois estados são usados para aproximar o modelo original. Para isto, precisamos calcular 4 parâmetros,  $\lambda_{ul}$ ,  $\lambda_{ol}$ ,  $r_{ul}$  e  $r_{ol}$ , como pode ser visto na Figura

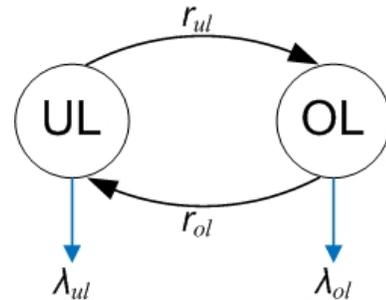


Fig. 3. Modelo MMPP de 2 Estados.

As equações que calculam os valores de  $\lambda_{ul}$  e  $\lambda_{ol}$  são mostrados a seguir:

$$\lambda_{ul} = \lambda \sum_{i=0}^M i \frac{\pi_i}{\pi_{ul}} \quad (5)$$

e

$$\lambda_{ol} = \lambda \sum_{i=M+1}^N i \frac{\pi_i}{\pi_{ol}} \quad (6)$$

sendo

$$\pi_{ul} = \sum_{i=0}^M \pi_i, \quad \pi_{ol} = \sum_{i=M+1}^N \pi_i$$

e

$$\pi = \frac{N!}{i!(N-1)!} P_{on}^i (1 - P_{on})^{N-1}$$

Para calcular o  $r_{ul}$  e  $r_{ol}$  foi usado o mesmo método usado na referência [7]. Considera-se primeiro o estado OL. Seja  $v$  uma variável aleatória definida como a duração do período em que o processo de nascimento e morte permanece nos estados *overload*. E seja  $v^*$  uma variável aleatória que representa o período de duração do estado OL no MMPP de 2 estados. Podemos encontrar o valor de  $r_{ol}$  aproximando as distribuições de  $v$  e  $v^*$ , que é uma exponencial com média  $\frac{1}{r_{ol}}$ .

A variável aleatória  $v$  pode ser identificada como sendo o tempo de absorção de um processo markoviano  $\mathcal{M}$ , obtido através de um processo de nascimento e morte considerando somente os estados  $\{M, M+1, \dots, N\}$ , sendo o estado  $M$  absorvente. A modelagem deste processo é ilustrada na Figura 4.

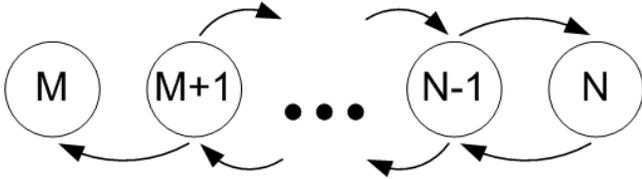


Fig. 4. Modelo dos estados de M a N, sendo M absorvente.

De acordo com [7], para encontrar o valor de  $r_{ol}$  basta calcular autovalor com parte real máxima de  $Q_M$ , matriz de transição de estados de  $\mathcal{M}$ . O resultado será um valor negativo  $-\eta$ , então  $r_{ol} = \eta$ .

Uma vez tendo os auto-valores  $\lambda_{ul}$ ,  $\lambda_{ol}$  e  $r_{ol}$ , é preciso calcular  $r_{ul}$  de modo a obter uma taxa de chegadas média igual ao modelo original. Para isto, usa-se a seguinte equação:

$$r_{ul} = r_{ol} \frac{N\lambda P_{on} - \lambda_{ul}}{\lambda_{ol} - N\lambda P_{on}} \quad (7)$$

## V. TRÁFEGOS VOIP COM PERÍODOS DE FALA DIFERENTES

Até o presente momento só foram mostrados modelos que sobrepõem tráfegos de VoIP pertencentes à mesma categoria de conversa. Em muitos casos, vários tráfegos sobrepostos seriam representados de forma mais próxima à realidade se considerar que nem todas as fontes pertencem à mesma categoria de conversa. Uma forma de fazer isso é sobrepor MMPPs de categorias diferentes.

Um aspecto interessante de tráfegos MMPP é que a sobreposição de dois MMPP resulta em uma novo MMPP [6]. Então a matriz de transição de estados  $Q$  e a matriz de taxas  $\mathcal{L}$  podem ser calculadas por

$$Q = Q_1 \oplus Q_2 \oplus \dots \oplus Q_n \quad (8)$$

e

$$\mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \dots \oplus \mathcal{L}_n, \quad (9)$$

onde  $\oplus$  representa a soma de Kronecker. Se  $MMPP_1$  e  $MMPP_2$  são da ordem de  $k_1$  e  $k_2$ , respectivamente, o processo resultante será da ordem de  $k_1 \times k_2$ . Esta característica é conhecida como explosão de estados, que, em muitos casos, impossibilita a análise matemática.

A solução neste trabalho para evitar a explosão de estados é a soma dos MMPPs após a redução para dois estados. A cadeia de Markov de transição de estados resultante da soma de Kronecker é mostrada na Figura 5.

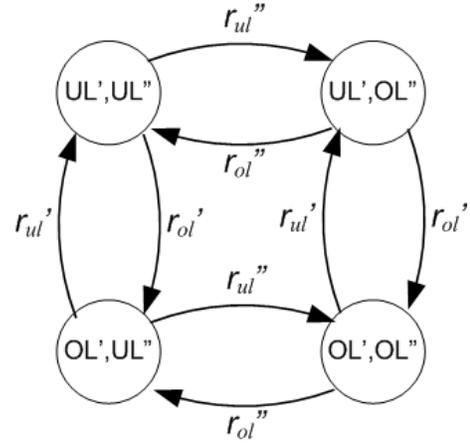


Fig. 5. Modelo de dois MMPPs reduzidos sobrepostos.

## VI. FILA MMPP/M/1

Para analisar o comportamento do tráfego em um enlace, ele precisa ser injetado em um sistema de filas. Como os pacotes de VoIP possuem tamanhos constantes, a melhor representação de um enlace seria um uma fila do tipo MMPP/D/1. Entretanto, por ser matematicamente mais simples, usa-se uma fila do tipo MMPP/M/1. Como este tipo de fila considera que o tamanho dos pacotes obedece uma distribuição exponencial, os resultados tenderão a ser mais conservadores.

O espaço de estados de uma fila MMPP/M/1 é o produto do espaço de estados do processo de chegadas com os possíveis tamanhos de fila [5]. A matriz de transição de taxas bloco diagonal  $\hat{Q}$  para fila MMPP/M/1 é a matriz infinita

$$\hat{Q} = \begin{bmatrix} \bar{A}_1 & A_0 & 0 & 0 & \dots \\ A_2 & A_1 & A_0 & 0 & \dots \\ 0 & A_2 & A_1 & A_0 & \dots \\ 0 & 0 & A_2 & A_1 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad (10)$$

sendo

$$A_0 = \mathcal{L} \quad (11)$$

$$\bar{A}_1 = Q - \mathcal{L} \quad (12)$$

$$A_1 = Q - \mathcal{L} - vI \quad (13)$$

$$\mathcal{A}_2 = v\mathcal{I} \quad (14)$$

O vetor de distribuição estacionária do tamanho de fila  $\hat{\pi}$  pode ser particionado em estados com o mesmo tamanho de fila:

$$\hat{\pi} = [\hat{\pi}(0), \hat{\pi}(1), \hat{\pi}(2), \dots]. \quad (15)$$

Então, a equação  $\hat{\pi}\hat{\mathcal{Q}} = 0$  possui soluções do tipo:

$$\hat{\pi}(k) = \hat{\pi}(0)\mathcal{R}^k, \quad (16)$$

onde a matriz de taxas  $\mathcal{R}$  é a solução da seguinte equação quadrática matricial:

$$\mathcal{A}_0 + \mathcal{R}\mathcal{A}_1 + \mathcal{R}^2\mathcal{A}_2 = 0 \quad (17)$$

onde a condição de contorno dada por:

$$\hat{\pi}(0) = \pi(\mathcal{I} - \mathcal{R}). \quad (18)$$

Na referência [5] explica várias formas numéricas de resolver a equação quadrática matricial (17).

Finalmente, a probabilidade escalar do tamanho da fila e o tamanho médio da fila são dados, respectivamente, pelas equações a seguir:

$$r(k) = \hat{\pi}(k)\varepsilon' = \pi(\mathcal{I} - \mathcal{R})\mathcal{R}^k\varepsilon', \quad k = 0, 1, \dots \quad (19)$$

e

$$E\{Q\} = \sum_{k=0}^{\infty} k\hat{\pi}(k)\varepsilon' = \pi\mathcal{R}(\mathcal{I} - \mathcal{R})^{-1}\varepsilon'. \quad (20)$$

## VII. RESULTADOS

Para comparar os métodos apresentados neste trabalho com resultados de simulações, foram estudados 2 cenários.

No primeiro cenário, encontram-se 20 fontes pertencentes a categoria de conversa *Classical*, apresentada na Tabela I. Este cenário é ilustrado na Figura 6.

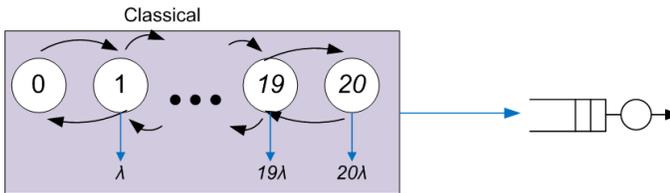


Fig. 6. Cenário 1.

Os resultados foram obtidos de 5 formas diferentes: MMPP/D/1 simulado, MMPP/M/1 simulado, MMPP/M/1 analítico sem redução de estados, MMPP/M/1 analítico com redução e M/M/1 analítico.

A Tabela II apresenta os tamanhos médios de acordo com a utilização, obtidos no primeiro cenários. Nota-se que o MMPP/M/1 analítico sem redução de estados foi a forma que melhor conseguiu acompanhar os resultados simulados. O MMPP/M/1 analítico com redução de estados não se mostrou adequado para utilizações  $\rho < 0.58$ .

A Figura 7 apresenta a probabilidade escalar do tamanho da fila para  $\rho = 0.86$  no primeiro cenário. Mais uma vez, o MMPP/M/1 analítico sem redução de estados foi o que melhor conseguiu acompanhar o comportamento simulado da fila.

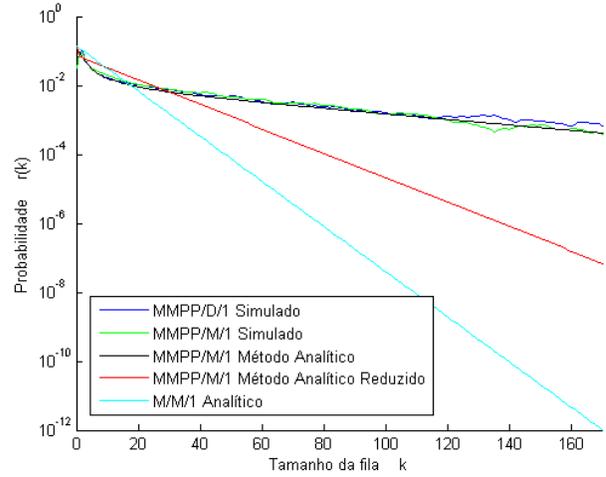


Fig. 7. Probabilidade de largura de fila.

Observa-se que a redução de estados, deixa o comportamento da fila mais próxima a de uma fila M/M/1.

No segundo cenário, encontram-se 60 fontes que foram divididas em 3 grupos de 20. Todas as fontes de um grupo possuem características de períodos de fala e silêncio de categorias de aplicações de voz semelhantes. Cada um dos grupos representa uma das categorias apresentadas na Tabela I. Este cenário é ilustrado na Figura 8.

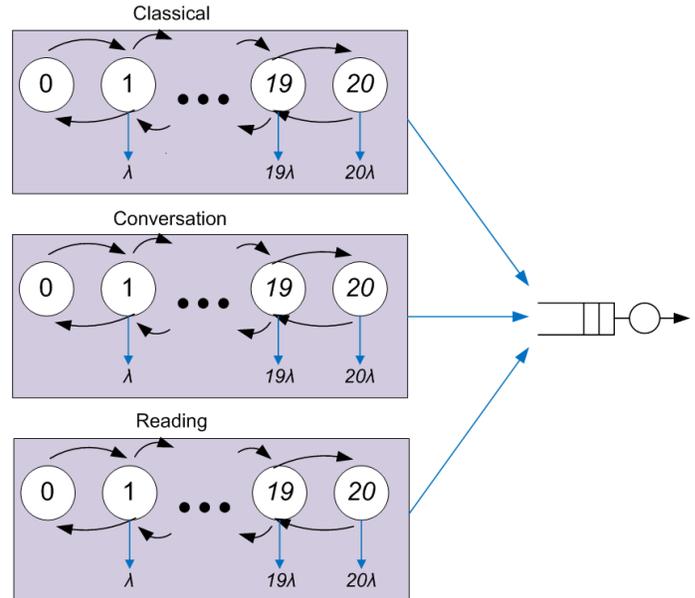


Fig. 8. Probabilidade de largura de fila.

Os resultados foram obtidos de 4 formas diferentes: 3MMPP/D/1 simulado, 3MMPP/M/1 simulado, 3MMPP/M/1 analítico com redução e M/M/1 analítico. Na aproximação analítica, após cada grupo ser modelado por MMPPs de 2 estados, estes MMPPs foram somados pela soma de Kronecker para resultarem em somente um MMPP de 8 estados. O método analítico só é computacionalmente possível com

redução de estados, já que sem redução a soma de Kronecker resultaria em uma explosão de estados que impossibilitariam os cálculos.

Na Tabela III, são apresentados os resultados dos tamanhos médios das filas de acordo com a utilização. Neste cenário, o 3MMPP/M/1 analítico com redução de estados e usando a soma de Kronecker se mostrou satisfatório de acordo com os resultados simulados.

A Figura 9 apresenta a probabilidade escalar do tamanho da fila para  $\rho = 0.86$  no segundo cenário. O que se nota é que com o aumento do número de fontes, as curvas dos resultados simulados seguem de forma mais próxima a curva do método analítico com redução de estados.

TABELA II

COMPARAÇÃO DOS TAMANHOS MÉDIOS DE FILA (PRIMEIRO CENÁRIO)

Classical					
$\rho$	MMPP/D/1 Sim.	MMPP/M/1 Sim	Mét. Aná. c/ Red. Est.	Mét. Aná. s/ Red Est.	M/M/1 Aná.
0,86	34,3877	41,1752	10,9246	30,4534	6,1428
0,8	14,8053	22,2749	6,2836	15,1547	4,0000
0,7	6,0087	7,2759	3,1508	5,5291	2,3333
0,58	1,7153	2,5083	1,6656	2,0627	1,3809
0,46	0,7722	1,2356	0,9581	1,0071	0,8518
0,3	0,3817	0,4603	0,4547	0,4543	0,4285
0,12	0,1293	0,1384	0,1389	0,1383	0,1363

TABELA III

COMPARAÇÃO DOS TAMANHOS MÉDIOS DE FILA (SEGUNDO CENÁRIO)

Classical + Conversation + Reading				
$\rho$	MMPP/D/1 Sim.	MMPP/M/1 Sim	Mét. Aná. c/ Red. Est.	M/M/1 Aná.
0,86	8,175808241	10,21864617	8,2377	6,142857143
0,8	3,32186286	5,319185668	4,7758	4
0,7	1,660628158	2,657550708	2,5493	2,333333333
0,58	1,007622811	1,4674321	1,443	1,380952381
0,46	0,661343097	0,863355675	0,8719	0,851851852
0,3	0,366294125	0,435042361	0,4328	0,428571429
0,12	0,128347692	0,136927213	0,1367	0,136363636

VIII. CONCLUSÕES

O objetivo deste trabalho foi obter um modelo aproximado para representar várias fontes de pacotes de voz. O modelo apresentado pode ser usado para fontes com características de períodos de fala e silêncio distintos.

Para estudar o comportamento de uma fila quando várias fontes são multiplexadas em um enlace, foi apresentada uma forma matemática de calcular a probabilidade escalar do tamanho da fila e o tamanho médio da fila. O modelo de fila usado é o MMPP/M/1; o serviço é markoviano por ser matematicamente mais simples, o que faz com que os resultados sejam conservadores em relação a um serviço determinístico.

Foram feitas simulações com MMPPs sem aproximações e com serviço determinístico para comparar o comportamento da fila com o método do artigo. Os resultados foram bastante satisfatórios e mostraram que o método é adequado para constantemente obter resultados conservadores, sem correr o risco de planejar um enlace de forma subestimada.

REFERÊNCIAS

- [1] H. Hassan, J.M. Garcia, C. Bockstal, *Aggregate Traffic Models for VoIP Applications*.: IEEE.
- [2] N. T. Bogdan , P. D. Elias , *Robustez na Comunicação de Voz Sobre o Protocolo IP Através do Uso de Redundâncias*.
- [3] ITU-T Recommendation G.729, *Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction*. CS-ACELP, Genève, March 1996.
- [4] D. B. O. Martins , *Dissertação de Mestrado*, Departamento de Engenharia Elétrica, Universidade de Brasília, Brazil. 2005 <http://www.gtel.ufc.br/martin/publicacoes.html>.
- [5] S. Hans-Peter, Tese de Doutorado, *Performance Analysis of Intermediate Systems Serving Aggregated ON/OFF Traffic with Long-Range Dependent Properties*. Institute für Informatik der Technischen Universität München Lehrstuhl für Informatik VIII.
- [6] W. Fischer, K. Meier-Hellestern, *The MMPP cookbook*, Performance Evaluation, 18, pp. 149- 171, 1992.
- [7] Anders Andersson. *Capacity study of statistical multiplexing for ip telephony* Technical Report T2000:03, SICS Swedish Institute of Computer Science, January 2000.

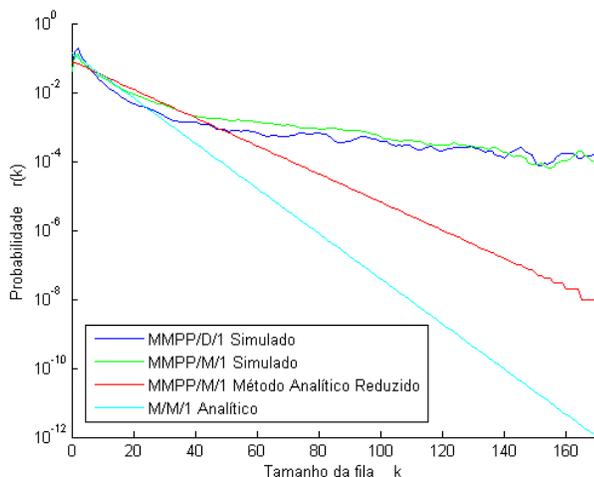


Fig. 9. Probabilidade de largura de fila da simulação e análise aproximada.