

# Análise da Evolução de Características de Tráfego VoIP

Carlos Marcelo Pedroso, Jefferson Caldeira, Keiko Fonseca e Mateus Cruz

**Resumo**—A demanda por sistemas de voz sobre IP (*Voice over IP, VoIP*) têm crescido em comparação aos sistemas de telefonia convencionais e atrai a atenção das grandes operadoras de telecomunicações. A caracterização e modelagem do tráfego de sistemas VoIP é essencial para o adequado dimensionamento da estrutura de redes bem como para a condução de trabalhos de pesquisa relacionados ao tema VoIP. No entanto, as características do tráfego VoIP não são necessariamente similares ao tráfego de voz convencional, além de que o crescimento do número de usuários e a adoção de determinadas tecnologias afetam estas características. Este trabalho apresenta a caracterização de tráfego de um sistema VoIP, analisando características do tráfego a partir de amostras obtidas nos anos de 2006 e 2007. As alterações nestas características são discutidas e suas possíveis causas apontadas.

**Palavras-Chave**—VoIP, modelos de tráfego.

**Abstract**—The growing number of users of Voice over IP (VoIP) systems compared to the traditional telephony systems attracts the attention of large telecommunication carriers. In order to provide VoIP services, system designers and implementors should characterize and model VoIP traffic, which may not be the same of the conventional telephony. This step is essential for the correct sizing of network facilities and also for scientific research. Based on the analysis of data collected over two years in a telecommunication carrier operator, this paper presents a VoIP traffic characterization. The analysis points out changes in traffic characteristics over two consecutive years, discusses their possible causes and trends.

**Keywords**—VoIP, traffic models.

## I. INTRODUÇÃO

A utilização de sistemas transmitindo Voz sobre IP (*Voice over IP, VoIP*) tende a alterar o perfil de uso das redes de telecomunicações. Atualmente o uso de sistemas VoIP está em seu estado inicial - no entanto, as operadoras que oferecem o serviço têm experimentado um grande crescimento na demanda por serviços VoIP.

Em sistemas VoIP, a voz humana é digitalizada, codificada e enviada na forma de fluxo de pacotes através de uma rede IP. Dispositivos ou programas denominados *codecs* (*coder-decoder*) realizam a codificação e decodificação de um fluxo de dados de voz. Os equipamentos que operam com VoIP devem suportar no mínimo o *codec* ITU G.711 [1] baseado na técnica PCM (*Pulse Code Modulation*). No entanto, os dispositivos podem implementar *codecs* mais complexos, por

exemplo G.723 [2] e G.729 [3] com possível impacto no perfil de tráfego gerado. Os *codecs* que obtêm melhores taxas de compressão tendem a utilizar pacotes maiores transmitidos a intervalos de tempo maiores se comparados ao G.711.

O serviço VoIP pode ser implementado utilizando-se a sinalização definida pela recomendação H.323 da ITU-T [4] ou pelo SIP (*Session Initiation Protocol* [5]). Neste artigo, a análise dos dados de tráfego refere-se à um sistema que utiliza o protocolo de sinalização SIP, com fluxos de voz transportados pelo protocolo RTP (*Real Time Protocol* [6]).

Um dos problemas a serem enfrentados na implementação da tecnologia é a realização do correto dimensionamento de recursos na rede da operadora. Já é conhecido que o tráfego gerado por sistemas VoIP não possui as mesmas características apresentadas pelo sistema de telefonia convencional [7].

Em um trabalho anterior [8] apresentamos uma análise do tráfego VoIP coletado em uma operadora de telecomunicações baseada em dados de 2006. Desde então, a rede física desta operadora não sofreu grandes alterações exceto pelo incremento no número de usuários. Mesmo assim, no momento de maior tráfego, a rede em estudo tinha uma ocupação menor que 10% de sua capacidade máxima.

Este trabalho apresenta a análise de uma amostra de tráfego coletada no ano de 2007, comparando as características de tráfego observadas neste ano com as características observadas na amostra coletada em 2006 na mesma rede. Não é de conhecimento dos autores nenhum trabalho anterior semelhante acompanhando as características de tráfego de sistemas VoIP em uso no Brasil. As características aqui apresentadas são importantes para análise de desempenho e dimensionamento de sistemas VoIP, além de possibilitar geração de tráfego sintético com perfil próximo do real para simulação de tais sistemas.

Este artigo está estruturado da seguinte forma. A seção II mostra os principais modelos para caracterização de tráfego VoIP em uso atualmente. A seção III descreve o conjunto de dados em estudo, e a seção IV apresenta a modelagem de diversas características de tráfego do sistema. Finalmente, os resultados são discutidos na seção V e as conclusões e trabalhos futuros são apresentados na seção VI.

## II. MODELOS DE TRÁFEGO VOIP

A voz humana pode ser modelada como uma seqüência de períodos de atividade e silêncio, conhecida como padrão ON-OFF. Primeiramente descrito por [9] é ainda utilizado para modelagem de tráfego de voz na rede pública de telefonia. Nestas redes, os períodos ON e OFF são modelados por distribuições de decaimento exponencial [10]. A Figura 1

Carlos Marcelo Pedroso é Professor na Universidade Federal do Paraná e na Pontifícia Universidade Católica do Paraná.

Keiko Fonseca é Professora na Universidade Tecnológica Federal do Paraná. Jefferson Caldeira é estudante do curso de Engenharia de Computação na Pontifícia Universidade Católica do Paraná.

Mateus Cruz é Engenheiro da Operadora de Telecomunicações GVT, Paraná. E-mails: pedroso@eletrica.ufpr.br, keiko@utfpr.edu.br, jefferson@puccpr.br, mateus.cruz@gvt.br

mostra um dos modelos mais populares para geração de tráfego de voz, o modelo *ON-OFF* [10]. Neste modelo, os pacotes são gerados somente durante o período onde o estado do sistema é *ON*. Durante este período, os pacotes são gerados com intervalos fixos de tempo entre chegadas; a sucessão entre estados *ON* e *OFF* é modelada pela distribuição exponencial.

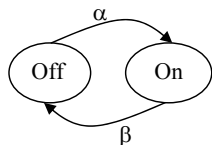


Fig. 1. Modelo ON-OFF e IPP.

O modelo *IPP* (*Interrupted Poisson Process*) modifica o modelo *ON-OFF* de tal maneira que, quando o processo está no estado *ON*, a chegada dos pacotes será realizada de acordo com o processo de Poisson com taxa  $\lambda$  [10].

Em redes de pacotes, a modelagem de tráfego usualmente utiliza distribuições de probabilidade para antecipar a chegada de pacotes e o tamanho dos mesmos. É de particular interesse na modelagem de tráfego de redes, os *Processos com Dependência de Longa Duração* ou *Long Range Dependence* (*LRD*) uma vez que as pesquisas indicam que o tráfego em redes modernas de comunicação frequentemente apresenta *LRD* [11]. Um processo *LRD* descreve sistemas em que observações de um evento são significativamente correlacionadas com observações muito distantes no tempo.

Em um processo *LRD* a função de auto-correlação  $\rho_k$  (ou *ACF*, *Auto-Correlation Function*) decai para zero lentamente em função do deslocamento (ou *Lag*)  $k$ .

Seja  $X_t$ ,  $t = 1, 2, \dots$ , um processo estacionário de segunda ordem. A função de auto-covariância com defasagem  $k$  será denotada por  $E[(X_t - \mu)(X_{t-k} - \mu)] = \gamma_k$  e a função de auto-correlação com defasagem  $k$  será denotada por  $\rho_k = \gamma_k/\gamma_0$ .

Se existe uma constante  $c$  tal que a função de auto-covariância seja dada por  $\gamma_k \sim c/k^\alpha$ ,  $0 < \alpha < 1$ , então o processo possui *LRD*. Desta forma, é importante realizar o estudo da função de auto-correlação para identificação da dependência de longa ou curta duração, o que será realizado posteriormente neste artigo utilizando-se o conjunto de dados de tráfego em estudo.

Um processo é dito *auto-similar* do ponto de vista estocástico se mantém suas características estocásticas sobre uma certa faixa de escalas, de modo a satisfazer a relação dada pela Equação 1. Esta relação indica que um processo dado por  $X_{mt}$  apresenta as mesmas características em termos de distribuição de probabilidade (indicada pelo símbolo  $\cong_d$ ) de um processo  $X_t$  re-escalado por um fator  $m$  e multiplicado pelo fator  $m^H$ .

$$X_{mt} \cong_d m^H X_t, m > 0 \quad (1)$$

Um *Processo com Memória de Curta Duração* ou *Short Range Dependence* (*SRD*) representa o comportamento onde a observação corrente não é correlacionada com observações

muito antigas. Para um processo *SRD*, a função de auto-correlação decai rapidamente para zero.

Processos auto-similares apresentam *LRD* [12]. O parâmetro de Hurst  $H = 1 - \alpha/2$  é utilizado para representar a auto-similaridade. O valor  $H = \frac{1}{2}$  é o limite entre a *LRD* e a *SRD*. Processos com  $\frac{1}{2} < H < 1$  são *LRD*, enquanto processos com  $0 < H < \frac{1}{2}$  são *SRD*. O efeito da *LRD* será maior quando  $H$  estiver próximo a 1.

Neste trabalho, a estimação do parâmetro de Hurst das séries em estudo será realizada através do método da transformada *Wavelet* utilizando-se as ferramentas disponibilizadas por [13]. *Wavelets* são funções matemáticas que seccionam os dados em diferentes componentes de frequência e então estudam cada componente com uma resolução ajustada à sua escala. *Wavelets* foram desenvolvidas independentemente nos campos da matemática, física quântica, engenharia elétrica e geologia sísmica [14]. Em [15] é estabelecido um método para determinação do parâmetro de Hurst em uma série baseada na transformada *Wavelet*. O método é relativamente imune a ruído e a não estacionariedade dos dados.

A auto-similaridade do tráfego pode ser obtida através de uma soma de fontes de tráfego *ON-OFF*, onde os períodos de tempo *ON* e *OFF* têm uma distribuição de cauda pesada, como a distribuição de Pareto [11].

Os modelos de tráfego do sistema VoIP baseados na modelagem dos períodos *ON* e *OFF* são frequentes na literatura: por exemplo, [16] analisa o tráfego VoIP no nível de pacotes, utilizando os *codecs* G.723 e G.729 e a distribuição de Pareto para caracterizar os períodos *ON-OFF* e também para o intervalo de tempo entre chegada de pacotes; em [17] é reportado o comportamento de cauda pesada para os períodos *ON/OFF*. Outras investigações, por exemplo em redes GPRS (*General Packet Radio Service*) utilizando G.729, sugere que o período *OFF* pode ser modelado por uma distribuição de Weibull [18], enquanto [19] argumenta que a distribuição Lognormal pode ser utilizada para caracterizar ambos, os períodos *ON* e *OFF*.

Os modelos baseados em distribuições de cauda pesada são, em geral, de difícil solução por apresentarem dificuldades na amostragem, parametrização e análise dos dados, pois este tipo de distribuição pode apresentar características de variância ou até mesmo média não convergente [12]. Trabalhos recentes [20] [21] relatam novas alternativas para caracterização do tráfego gerado por *codecs* G.729 e G.723 e levam em conta a influência de quadros *SID* (*Silence Insertion Descriptor*), gerados durante os períodos de inatividade (silêncio), e que descrevem as características principais do ruído de fundo. Este modelo, conhecido como *ON-SID*, ainda é objeto de estudos e não será abordado neste trabalho.

O estudo aqui apresentado analisa os dados obtidos de uma operadora de telefonia, identifica os principais *codecs* em uso e realiza a modelagem do tráfego no nível de pacotes. Diferentemente da literatura pesquisada cujos resultados baseiam-se na análise de dados de sistemas VoIP operando em Universidades ou em empresas (redes corporativas), nossos dados referem-se a dados com um nível de agregação significativamente maior se comparado com os trabalhos publicados anteriormente sobre a caracterização de sistemas VoIP.

TABELA I

SUMÁRIO DE ESTATÍSTICAS BÁSICAS DOS DADOS EM ESTUDO

Amostra	Data	Tempo total	Quantidade de pacotes
2006	4/05/2006 10:15h	92,8 segundos	2.420.654
2007-1	17/09/2007 10:15h	18,1 segundos	481.921
2007-2	18/09/2007 10:15h	54,6 segundos	1.528.895
2007-3	19/09/2007 10:15h	60,9 segundos	1.816.864
2007-4	20/09/2007 10:15h	60,4 segundos	1.690.489
2007-5	21/09/2007 10:15h	60,9 segundos	1.696.016

TABELA II

SUMÁRIO DE ESTATÍSTICAS SOBRE O INTERVALO DE TEMPO ENTRE PACOTES RTP, COM VALORES EM *microsegundos* ( $10^{-6}$  SEGUNDOS)

Amostra	Mediana	Média	Variância	Erro (h) para 99% de Confiança
2006	23	41,87	0,004768	0,04415
2007-1	2	39,37	0,003976	0,09030
2007-2	4	36,33	0,003694	0,04091
2007-3	4	33,60	0,003069	0,04089
2007-4	4	36,55	0,003419	0,04475
2007-5	4	36,74	0,003471	0,04501

### III. CONJUNTO DE DADOS UTILIZADO

O conjunto de dados em estudo refere-se ao tráfego coletado em uma grande operadora de telecomunicações brasileira. A operadora fornece um serviço de Voz sobre IP utilizando o protocolo SIP para sinalização. O SIP é utilizado para estabelecer chamadas através do protocolo IP, atua na camada de aplicação e é um padrão da *Internet Engineering Task Force* (IETF).

O SIP foi projetado para interagir com outros protocolos da internet funcionando como um mecanismo de estabelecimento de sessão, iniciando, modificando e terminando a sessão independente do tipo de mídia ou aplicação que será usada. O protocolo caracteriza-se como simples, flexível, estável, baseado em texto, e não depende do protocolo de transporte. Uma vez iniciada a conferência, a voz é codificada/decodificada por um *codec* e transmitida através do protocolo RTP.

A rede analisada possuía aproximadamente 10.000 usuários na época da coleta do tráfego em 2006. Em termos gerais, a rede de acesso é formada por enlaces ADSL (*Asymmetric Digital Subscriber Line*). O tráfego VoIP gerado pelos usuários é direcionado para uma rede Ethernet, não congestionada. Para a coleta de dados, diversas portas do *switch* Ethernet que serve como *backbone* da rede foram espelhadas, de modo que o tráfego total do sistema VoIP foi capturado utilizando-se o analisador de protocolos *Open Source* Wireshark [22].

Os dados em estudo foram coletados em maio de 2006 (amostra 2006) e entre 17 e 21 de outubro 2007 (amostras 2007-1, 2007-2, 2007-3, 2007-4 e 2007-5), no momento de maior tráfego. A Tabela I mostra um sumário dos dados coletados. A precisão de tempo na captura foi de 1 microsegundo.

A confiança estatística da amostra pode ser calculada para as variáveis de interesse. O intervalo de confiança compreende um intervalo numérico que possui uma probabilidade de incluir o verdadeiro valor da variável ou medida de desempenho sob análise [23]. As Tabelas II e III mostram respectivamente um sumário do intervalo de tempo entre chegada de pacotes (também chamado de intervalo entre pacotes) e do tamanho

TABELA III

SUMÁRIO DE ESTATÍSTICAS SOBRE O TAMANHO DOS PACOTES RTP (BYTES)

Amostra	Mediana	Média	Variância	Erro (h) para 99% de Confiança
2006	74	107,60	3364,7	0,037
2007-1	74	86,03	1588,3	0,057
2007-2	74	84,80	1402,8	0,030
2007-3	74	85,44	1572,0	0,029
2007-4	74	84,75	1381,8	0,028
2007-5	74	85,25	1450,0	0,029

TABELA IV

PROTOCOLOS ENCONTRADOS

Protocolo	Quantidade de pacotes	%
RTP	2.221.441	91,77
SIP	39.998	1,65
Outros	159.215	6,57

dos pacotes do fluxos RTP. Nestas tabelas são mostrados os intervalos de confiança para média das respectivas variáveis em estudo. Desta forma,  $h$  será o erro admitido ao se concluir sobre a presença do verdadeiro valor da variável com o nível de confiança desejado.

O intervalo de tempo entre chegada de pacotes RTP da amostra 2006 foi de 41.87 microsegundos e o erro para um nível de confiança de 99% é dado por  $h = 44,1$  nanosegundos (Tabela II). Para as amostras coletadas em 2007, a análise do tempo entre chegada de pacotes RTP indica que estas também possuem erros pequenos em relação à média. Isto se deve à grande quantidade de pacotes amostrados e confirma a representatividade de ambas amostras (2006 e 2007).

Os erros para um intervalo de confiança de 99% foram calculados para o tamanho do pacote e o resultado é apresentado na Tabela III. Da mesma forma que o intervalo de tempo entre chegada de pacotes, o desvio em relação à média é pequeno.

### IV. MODELAGEM DO TRÁFEGO

Para a modelagem do tráfego VoIP apresentada neste trabalho foram analisados:

- Protocolos utilizados;
- *Codec* utilizado pelos fluxos RTP;
- intervalo de tempo entre chegada de pacotes dos fluxos RTP de voz;
- Tamanho do pacote dos fluxos RTP de voz;
- Tráfego agregado total.

#### A. Protocolos Utilizados

O maior número de pacotes amostrados carregava o protocolo RTP, transportando voz codificada (91,77% dos pacotes). Pacotes carregando sinalização SIP corresponderam a 1,65% do total e protocolos diversos completaram a amostra (principalmente ARP, ICMP, VRRP e HTCP). A Tabela IV sumariza os principais protocolos encontrados na amostra 2006.

#### B. Codec utilizado pelos fluxos RTP

A Tabela V mostra os principais *codecs* identificados nas amostras de 2006 e 2007. Percebe-se que em 2006 os principais *codecs* em uso foram o G.711 e o G.729. As amostras

TABELA V  
Codecs IDENTIFICADOS

Amostra	G.711	G.729	iLBC	Outros
2006	22,52%	70,37%	6,67%	0,33%
2007-1	7,60%	93,39%	0,00%	0,00%
2007-2	6,88%	93,11%	0,00%	0,00%
2007-3	7,22%	92,77 %	0,00%	0,00%
2007-4	7,17%	92,82%	0,00%	0,00%
2007-5	7,37%	92,62 %	0,00%	0,00%

de 2007 indicam que o *codec* G.711 está sendo preterido em relação ao G.729, que passou a ter uma participação maior, em torno de 92% do total. O G.711 passou de 22,5% em 2006 para em torno de 7% em 2007.

### C. intervalo de tempo entre chegada de pacotes RTP

A principal contribuição deste artigo é apresentar a caracterização dos fluxos RTP. Como mostrado no item anterior, a maior parte dos protocolos observados pertence à transmissão de voz codificada utilizando protocolo RTP e a sua modelagem é de fundamental importância para o estudo de tais sistemas. A Tabela II mostra um sumário com estatísticas básicas sobre os intervalo de tempo entre chegada de pacotes RTP das amostras. Percebe-se que o intervalo médio entre pacotes apresentou redução entre os anos de 2006 e 2007, causado principalmente pelo aumento do número de usuários e conseqüente incremento no número de conferências simultâneas. Observa-se também uma pequena redução na variância das amostras de 2007 quando comparadas com 2006.

A Figura 2 mostra as funções de auto-correlação do intervalo de tempo entre chegada de pacotes RTP das amostras de 2006 e 2007. Na amostra 2006 pode ser observado o lento decaimento da função de auto-correlação, o que é um indicativo da auto-similaridade da série, o que ocorre em menor intensidade nas amostras de 2007.

Para que os valores da função de auto-correlação fossem considerados não correlacionados para um nível de confiança de 95%, os valores correspondentes para uma determinada defasagem deveriam estar abaixo da linha pontilhada na Figura 2. Percebe-se que a amostra de 2006 apresenta maior grau de auto-correlação se comparado às amostras de 2007. A comparação da ACF das amostras de 2006 com 2007 evidencia uma alteração de características do tráfego VoIP.

Os parâmetros de Hurst encontrados para os intervalo de tempo entre chegada de pacotes RTP, bem como os respectivos intervalos de confiança são apresentados na Tabela VI. Os testes mostraram que a amostra de 2006 apresenta auto-similaridade leve com o parâmetro de Hurst de aproximadamente 0,58. No entanto, as amostras de 2007 apresentam  $H < 0,5$ , o que é uma característica de variáveis com memória de curta duração.

A Figura 3 mostra a função distribuição acumulada empírica da amostra de 2006 comparada com a amostra 2007-2<sup>1</sup>. Em relação a 2006, percebe-se uma alteração no comportamento

<sup>1</sup>A distribuição acumulada de probabilidade da amostra 2007-2 é semelhante às demais amostras coletadas em 2007 - desta forma, não é necessário apresentar gráficos com a distribuição de probabilidade de tais amostras.

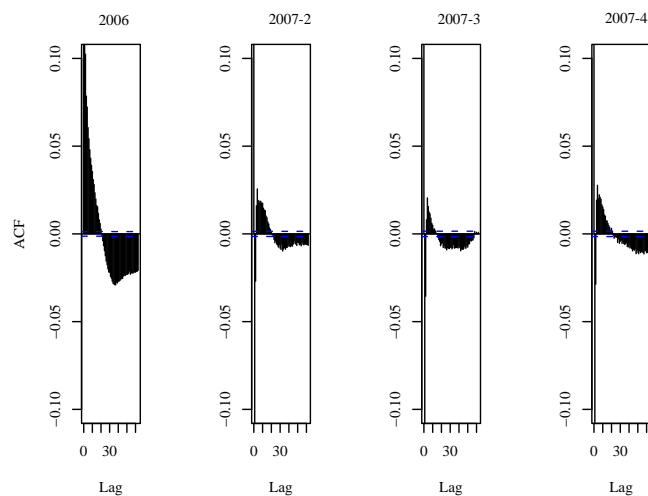


Fig. 2. Função de auto correlação (ACF) para o tempo entre chegada de pacotes RTP

TABELA VI  
PARÂMETRO DE HURST PARA OS INTERVALO DE TEMPO ENTRE CHEGADA DE PACOTES RTP

Amostra	Parâmetro de Hurst	Intervalo de confiança para 95%
2006	0,5766	[0,575, 0,578]
2007-1	0,4242	[0,422, 0,426]
2007-2	0,4417	[0,441, 0,443]
2007-3	0,4174	[0,416, 0,418]
2007-4	0,3971	[0,396, 0,398]
2007-5	0,4106	[0,410, 0,412]

que em 2007 passou a apresentar 3 regimes distintos, dependendo do intervalo de tempo considerado.

### D. Tamanho dos pacotes dos fluxos RTP de voz

A Tabela III mostra um sumário com estatísticas básicas sobre os tamanhos dos pacotes coletados nas diversas amostras. Percebe-se que de 2006 para 2007 ocorreu uma alteração no tamanho médio do pacote.

A análise do tamanho dos pacotes dos fluxos RTP de voz mostra que o tamanho possui um comportamento discreto: foram detectados pacotes de 64, 74, 84, 104, 214 e 294 bytes. O histograma apresentado na Figura 4 mostra a frequência observada em cada tamanho, comparando a amostra 2006 com a amostra 2007-2, tendo esta última comportamento semelhante entre todas as amostras de 2007. Esta alteração está relacionada com o uso do *codec* G.729 em detrimento ao G.711.

A análise da função de auto-correlação do tamanho de pacotes sucessivos, apresentada na Figura 5, mostra um decaimento lento para a amostra de 2006, indicativo da auto-similaridade da série. Já nas amostras de 2007 o decaimento da função de auto-correlação foi rápido em função do deslocamento (*lag*), à semelhança do que foi observado com os intervalo de tempo entre chegada de pacotes.

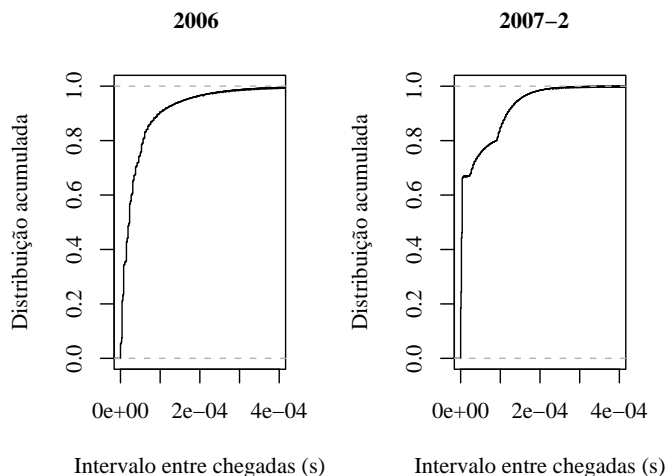


Fig. 3. Distribuição de probabilidade empírica acumulada para amostras 2006 e 2007-2

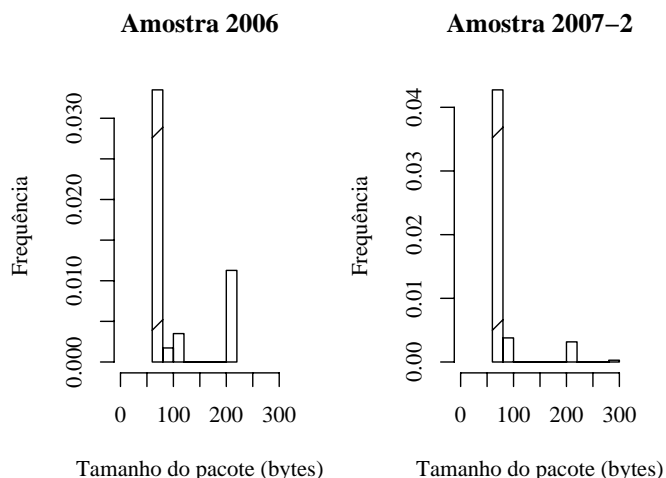


Fig. 4. Histograma do tamanho do pacote dos fluxos RTP

A Tabela VII mostra o cálculo do parâmetro de Hurst para as diversas amostras em estudo através do método da transformada de Wavelet. O resultado indica que a amostra de 2006 apresenta características auto-similares, com  $H \simeq 0,56$ . No entanto, as amostras de 2007 não apresentaram características auto-similares, comportando-se como séries SRD.

E. Tráfego Agregado Total

O estudo do tráfego agregado total foi realizado utilizando-se dados coletados ao longo dos dias 21-09-2007 e 22-09-2007, com uma totalização do tráfego agregado total gerado pelo sistema VoIP amostrado a cada 1 milissegundo. A função de auto-correlação apresenta um decaimento lento, como pode ser observado na Figura 6. Através do método da transformada Wavelet foi calculado o parâmetro de Hurst  $H = 0,582$ , no intervalo  $[0,581 \ 0,583]$  para um nível de confiança de

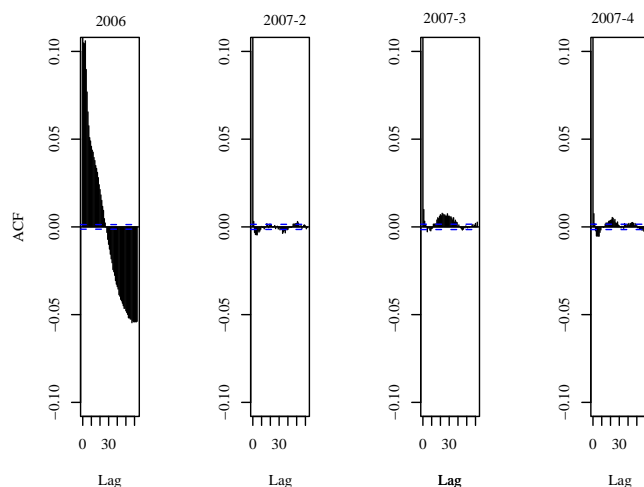


Fig. 5. Função de auto correlação para o tamanho dos pacotes RTP

TABELA VII  
PARÂMETRO DE HURST PARA O TAMANHO DOS PACOTES DOS FLUXOS RTP

Amostra	Parâmetro de Hurst	Intervalo de confiança para 95%
2006	0,5766	[0,575, 0,578]
2007-1	0,4865	[0,484, 0,489]
2007-2	0,5031	[0,502, 0,504]
2007-3	0,4882	[0,487, 0,489]
2007-4	0,4966	[0,495, 0,498]
2007-5	0,4893	[0,488, 0,490]

95%. Desta forma, observa-se auto-similaridade no tráfego agregado total - no entanto, esta auto-similaridade não é muito pronunciada (parâmetro de Hurst próximo de 0,5).

V. DISCUSSÃO DE RESULTADOS

Os resultados apresentados evidenciam a mudança de comportamento do tráfego VoIP na operadora, onde observam-se alterações nas características de tráfego do ano de 2007 em relação à amostra do ano anterior. Os autores acreditam que a mudança é influenciada por:

- (i) uma alteração na relação de *codecs* utilizados, com o *codec* G.729 sendo utilizado pela maior parte das sessões e sensível redução no uso do G.711.
- (ii) uma alteração observada nas características do intervalo de tempo entre chegada de pacotes RTP. Tudo indica que esta variável reduziu suas características auto-similares. Isto pode ter sido estimulado por dois fatores: (a) o aumento no número de usuários, com conseqüente incremento no nível de agregação de tráfego e (b) o menor uso do *codec* G.711.

Finalmente, o tamanho dos pacotes RTP também apresentou alteração de características em relação à 2006. Da mesma forma como no caso do tempo entre chegada de pacotes, o tamanho dos pacotes observado reduziu a sua característica

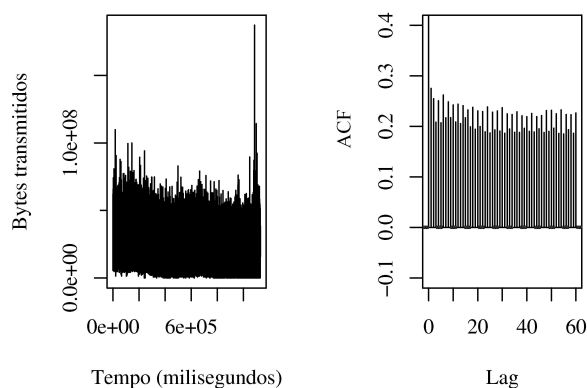


Fig. 6. Tráfego agregado total na escala de 1ms. Direita: função de auto-correlação. Esquerda: série temporal

auto-similar. As causas desta mudança estão sendo presente-mente investigadas.

## VI. CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho realizou a caracterização de tráfego VoIP de uma grande operadora de telecomunicações. O conjunto de dados disponibilizado foi capturado nos anos de 2006 e 2007, diretamente no *call server* da operadora.

O resultado mais marcante deste trabalho é a comparação sistemática de diversas características de tráfego observadas em amostras coletadas com 1 ano de intervalo e a constatação que estas características apresentaram alteração, principalmente em relação à sua estrutura de auto-correlação temporal. Duas importantes métricas de desempenho, o tempo entre chegada de pacotes e o tamanho dos pacotes dos fluxos RTP, reduziram as características auto-similares em comparação com àquelas coletadas no ano anterior.

Os resultados aqui apresentados são de grande valor para o estudo de desempenho das redes que operam sistema VoIP, permitindo o uso de ferramentas analíticas e de simulação apropriadas para realizar o dimensionamento e análises em redes existentes ou em pesquisas científicas. Com os dados aqui apresentados é possível alimentar geradores de tráfego com o perfil encontrado de fato em redes reais, semelhantes à rede estudada.

As principais implicações dos resultados refletem diretamente em:

- (i) Alteração em parâmetros de projeto e dimensionamento de equipamentos de rede e enlaces para operação de sistemas VoIP;
- (ii) Alteração de parâmetros nos modelos de geração de tráfego sintético para estudo de desempenho de sistemas VoIP.

A seqüência deste trabalho está sendo executada com a coleta de novas amostras com precisão de tempo de 1 nanosegundo, principalmente para confirmar as análises a respeito do intervalo entre pacotes. Outro estudo sendo conduzido é a caracterização do comportamento da sessão: o tempo de duração da conversação e intervalo entre chamadas será tratado em detalhes em um trabalho futuro.

## VII. AGRADECIMENTOS

Os autores agradecem à operadora de telecomunicações GVT pela disponibilização dos dados utilizados no estudo realizado.

## REFERÊNCIAS

- [1] ITU, "Recommendation G.711. pulse code modulation (PCM) of voice frequencies," 1988, International Telecommunication Union.
- [2] —, "G.723 : Extensions of recommendation G.721 adaptive differential pulse code modulation to 24 and 40 kbit/s for digital circuit multiplication equipment application," 1988, International Telecommunication Union.
- [3] —, "G.729 : Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (cs-acelp)," 1996, International Telecommunication Union.
- [4] D. Minoli, E. Minoli, and L. Sookchand, "ITU-T H.320/H.323," pp. 4018–4030, 2000.
- [5] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "RFC3261: Session initiation protocol," 2002.
- [6] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "STD 64: RTP: A transport protocol for real-time applications," Jul. 2003.
- [7] M. Fiedler, P. Carlsson, and A. Nilsson, "Voice and multi-fractal data in the internet," in *LCN '01: Proceedings of the 26th Annual IEEE Conference on Local Computer Networks*. Washington, DC, USA: IEEE Computer Society, 2001, p. 426.
- [8] C. M. Pedroso, J. Caldeira, and K. Fonseca, "Caracterização de tráfego VoIP," in *I2TS 2006. 6th International Information and Telecommunication Technologies Symposium*, 2007.
- [9] P.Brady, "A technique for investigating on/off patterns of speech." *Bell Labs Tech Journal*, January 1965.
- [10] A. Adas, "Traffic models in broadband networks," *IEEE Communications Magazine*, vol. 37, no. 7, pp. 82–89, July 1997.
- [11] W. Leland, M. Qaqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, February 1994.
- [12] W. Willinger and K. Park, *Self-similar network traffic and performance evaluation*, 1st ed. New York: John Wiley & Sons, 2000.
- [13] M. Roughan, D. Veitch, and P. Abry, "On-line estimation of the parameters of long-range dependence," in *Proceedings Globecom '98*, vol. 6, Sydney, 1998, pp. 3716–3721. [Online]. Available: [citeseer.ist.psu.edu/roughan98line.html](http://citeseer.ist.psu.edu/roughan98line.html)
- [14] A. Graps, "An introduction to Wavelets," *IEEE Comput. Sci. Eng.*, vol. 2, no. 2, pp. 50–61, 1995.
- [15] P. Abry and D. Veitch, "Wavelet analysis of long-range dependent traffic," *IEEE Trans. on Info. Theory*, vol. 44, no. 1, pp. 2–15, January 1998.
- [16] T.D.Dang, B.Sonkoly, and S.Molnr, "Fractal analysis and modeling of VoIP traffic," in *Report to Ministry of Education IKTA-0092/2002*, 2003.
- [17] W.Jiang and H.Schulzrinne, "Analysis of on-off patterns in voip and their effect on voice traffic aggregation," in *Proceedings of 9th IEEE International Conference on Computer Communication Networks*, 2000.
- [18] B.Bellalta, M.Oliver, and D.Rincn, "Capacity and traffic analysis of voip services over gprs mobile networks," in *report TIC2001-0956-C04. Technical University of Catalonia, University Pompeu Fabra, Spain.*, 2001.
- [19] E.Casilari, H.Montes, and F.Sandoval, "Modelling of voice traffic over ip networks," in *Proc. of Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, 2002.
- [20] A.Estepa, R.Estepa, and J.Vozmediano, "Packetization and silence influence on voip traffic profile," in *Lecture Notes on Computer Science, Multimedia Interactive Protocols and Systems MIPS 2003*, 2003.
- [21] A. Estepa, R. Estepa, and J. Vozmediano, "A new approach for VoIP traffic characterization." *IEEE Communications Letters*, vol. 8, no. 10, October 2004.
- [22] A. Orebaugh, G. Ramirez, J. Burke, and L. Pesce, *Wireshark & Ethereal Network Protocol Analyzer Toolkit (Jay Beale's Open Source Security)*. Syngress Publishing, 2006.
- [23] J. Banks, J. Carson, B. Nelson, and D. Nicol, *Discrete-event system simulation*, 3rd ed. New Jersey: Prentice Hall, 2001.