

Caracterização de tráfego VoIP

Carlos Marcelo Pedroso
Departamento de Computação
PUCPR
Curitiba, Paraná, Brasil
Email: c.pedroso@pucpr.br

Jefferson P. Caldeira
Departamento de Computação
PUCPR
Curitiba, Paraná, Brasil
Email: caldeira@gmail.com

Keiko Fonseca
Pós Graduação em Engenharia
Elétrica e Informática Industrial UTFPR
Curitiba, Paraná, Brasil
Email: keiko@utfpr.edu.br

Abstract—This paper presents the traffic characterization of a VoIP system based on data collected from a Brazilian carrier with more than 10.000 active users. The resulting model updates existing models and can be applied to capacity planning and scientific research to workload generation for VoIP system simulation.

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 serviços baseados em VoIP têm experimentado um grande crescimento de demanda por estes serviços.

Um dos problemas a serem enfrentados na implementação de novas tecnologias é o correto dimensionamento do sistema. Já é conhecido que o tráfego gerado por sistemas VoIP não possui as mesmas características apresentadas pelo sistema de telefonia convencional. Modelos modernos [1] indicam que o tráfego de sistemas VoIP pode ser modelado através de um modelo ON-OFF onde os períodos ON e OFF são caracterizados por distribuições de cauda pesada.

Este trabalho apresenta a caracterização de tráfego VoIP amostrado em uma grande operadora de telecomunicações no Brasil contando com aproximadamente 10.000 usuários ativos. Esta caracterização do tráfego se fez a nível de pacotes, utilizando aproximadamente 90 segundos de tráfego.

O estudo apresentado analisa os dados obtidos do sistema real, identifica os principais *codecs* em uso e realiza a modelagem do tráfego no nível de pacotes. É importante citar que a maioria dos artigos encontrados baseia-se em análise de dados de sistemas VoIP operando em Universidades ou em empresas - diferentemente do nosso universo de usuários, o que torna os resultados aqui apresentados ainda mais importantes. Não é de conhecimento dos autores nenhum trabalho anterior caracterizando um sistema VoIP em uso no Brasil. As características aqui reveladas configuram-se como dados importantes para análise de desempenho e pesquisas científicas sobre sistemas VoIP.

Este artigo está estruturado como a seguir. A seção II introduz o conceitos básicos sobre sistemas VoIP. A seção III mostra um resumo com trabalhos anteriores sobre o tema. A IV descreve o conjunto de dados em estudo, e a seção V apresenta a modelagem de diversas métricas de desempenho

em estudo. Finalmente, as conclusões e trabalhos futuros são apresentados na seção VI.

II. SISTEMAS VOIP: CONCEITOS BÁSICOS

Um sistema Voip é caracterizado pelo transporte da voz através de uma infra-estrutura baseada em IP (rede IP). Usualmente se refere à conversão da voz tradicional (sinal analógico) em sinais digitais a serem transportados em pacotes IP. Esta conversão conhecida como codificação da voz pode ser realizada diretamente no próprio aparelho telefônico (telefonia IP), no PABX ou em gateways específicos antes de serem encaminhados para roteadores IP.

Vários algoritmos têm sido padronizados para codificação de voz com taxas de bit variando desde 5,3kb/s to 64kb/s [2] [3]. Um *codec* (de *Coder-Decoder*) é um dispositivo ou programa capaz de realizar a codificação e a decodificação de voz em fluxo de dados. Seus componentes usuais são amostragem da voz, conversão analógica para digital, codificação, encapsulamento, transmissão, decodificação, armazenamento e ajuste do atraso (*play out*). Exemplos de codecs são o G.711, G.729 e G.723 [4]. Os dispositivos que operam com Voz sobre IP devem suportar no mínimo o codec ITU G.711, que é a especificação da técnica PCM (*Pulse Code Modulation*). No entanto, os dispositivos podem implementar codecs mais complexos, com possível impacto no perfil de tráfego gerado. Os codecs que apresentam melhores taxas de compressão tendem a utilizar pacotes maiores transmitidos a intervalos de tempo maiores se comparados ao G.711. A escolha do codec é do cliente.

Um protocolo amplamente utilizado para troca de sinalização de telefonia é o SIP (*Session Initiation Protocol*) [5]. Ele estabelece, modifica e finaliza chamadas telefônicas através do protocolo IP e atuando na camada de aplicação. Criado originalmente por Henning Schulzrinne (Columbia University) e Mark Handlery (UCL) em 1996, 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 na chamada; uma chamada pode utilizar diferentes tipos de dados, incluindo áudio e vídeo. O protocolo caracteriza-se por ser simples, flexível, estável, de não depender do protocolo de transporte e de ser baseado em texto. Uma vez iniciada

a conferência, a voz será codificada/decodificada utilizando um algoritmo (codec) e transmitida através do protocolo RTP (*Real Time Protocol*[6]).

III. MODELAGEM 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 [7] e ainda utilizada para modelagem de tráfego na rede de telefonia fixa. Neste tipo de rede, usualmente, os períodos ON e OFF são modelados por distribuições de decaimento exponencial. Em redes de pacotes, a modelagem de tráfego usualmente utiliza distribuições de probabilidade para prever a chegada de pacotes e o tamanho dos mesmos.

Trabalhos recentes [8] [9] relatam uma nova alternativa para caracterização do tráfego gerado pelos *codecs* G.729 e G.723 não baseada no modelo ON-OFF. Outros trabalhos realizados sobre redes de serviço diferenciado também mostram que a caracterização de tráfego diverge daquela utilizada para rede de telefonia convencional; em [10] é reportado o comportamento de cauda pesada para os períodos ON/OFF e em [11] é reportado a aderência à distribuição de Weibull, que possui decaimento semi-exponencial. [12] considera o uso da distribuição de Pareto e modelos multi-fractais para reproduzir e caracterizar o tráfego VoIP. Investigações em [13] de uma rede GPRS (General Packet Radio) utilizando G.729 sugere que o período OFF pode ser modelado por uma distribuição de Weibull, enquanto [14] argumenta que a distribuição Lognormal pode ser utilizada para ambos os períodos ON e OFF. [15] analisa o tráfego VoIP no nível de pacotes, utilizando os *codecs* G.723 e G.729 e reporta a distribuição de Pareto para os períodos ON-OFF e também para o intervalo entre pacotes.

IV. 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. A rede analisada possuía aproximadamente 10.000 usuários na época da coleta do tráfego.

Em termos gerais, a rede de acesso é formada de enlaces ADSL (Asymmetric Digital Subscriber Line). O tráfego VoIP gerado pelos usuários é direcionado para uma rede Ethernet não congestionada. Os usuários enviam solicitações ao *Call Server* conectado à rede Ethernet. O tráfego da porta do *switch* Ethernet onde está conectado o *Call Server* foi espelhado em uma porta onde foi configurado o analisador de protocolos *Open Source* Wireshark [16] e os dados coletados.

Os dados foram coletados em 4 de maio de 2006, durante um período de 92.8 segundos no momento de maior tráfego. A Tabela I mostra um sumário dos dados coletados.

V. MODELAGEM DO TRÁFEGO

A partir da amostra coletada, foram analisados:

- Protocolos utilizados;

TABELA I
SUMÁRIO COLETA DE DADOS

Tempo total	Quantidade de pacotes	Tamanho do arquivo
92,8 segundos	2.420.654	315.073.656 bytes

- Intervalo entre pacotes dos fluxos RTP de voz;
- Tamanho do pacote dos fluxos RTP de voz;
- Codec utilizado pelos fluxos RTP.

O item V-A a seguir introduz conceitos fundamentais, que serão utilizados na avaliação das métricas citadas acima.

A. Conceitos Fundamentais

Para entendimento dos resultados da análise dos dados obtidos fazem se necessários os conceitos relacionados a *Processo com dependência de longa duração* ou *Long Range Dependence (LRD)*.

Um *Processo com dependência de longa duração* ou *Long Range Dependence (LRD)* é o fenômeno que ocorre quando observações de um evento são significativamente correlacionadas com observações muito distantes no tempo. Este fenômeno é de particular interesse na modelagem de tráfego, uma vez que pesquisas comprovaram que o tráfego em redes modernas de comunicação frequentemente apresenta *LRD* [17].

Em um processo *LRD* a função de auto-correlação ρ_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, se existe uma constante c tal que $\gamma_k \sim c/k^\alpha, 0 < \alpha < 1$, então o processo possui *LRD*.

O termo *auto-similar* descreve a geometria de objetos em que uma pequena parte quando expandida se parece com o todo [18]. Muitos objetos encontrados na natureza apresentam características auto-similares (ou fractais aproximados). Pode-se citar como exemplo a estrutura geométrica de nuvens, montanhas e árvores.

Um processo será *auto-similar* do ponto de vista estocástico se ele mantém características estocásticas sobre uma certa faixa de escalas, de modo a satisfazer a relação dada pela Equação 1. Processos auto-similares apresentam *LRD* [1].

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

Hurst foi um famoso hidrólogo interessado no Egito e no rio Nilo que mostrou em 1951 que o nível do rio Nilo é uma seqüência *LRD*. Por razões históricas, o parâmetro de Hurst $H = 1 - \alpha/2$ é utilizado para representar a auto-similaridade. Um processo com memória longa apresenta auto-similaridade assintótica de segunda ordem.

Processo com memória de curta duração ou *Short Range Dependence (SRD)* indica o fenômeno onde a observação corrente não é correlacionada com observações muito antigas. Para um processo *SRD*, a função de auto-correlação decai rapidamente para zero.

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 $0 < H < \frac{1}{2}$

TABELA II
PROTÓCOLOS ENCONTRADOS

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

TABELA III
SUMÁRIO DE ESTATÍSTICAS SOBRE O INTERVALO ENTRE PACOTES RTP
(SEGUNDOS)

Min.	1° Q	Mediana	Média	3° Q.	Max.	Var.
0.00	8.0e-06	2.30e-05	4.187e-05	5.0e-05	6.649e-03	4.768e-09

$\frac{1}{2}$ são *SRD*. O efeito da *LRD* será maior quando *H* estiver próximo a 1.

B. 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 (principalmente ARP, ICMP, VRRP e HTCP) foram identificados no restante dos pacotes. A Tabela II sumariza o percentual de carga (quantidade de pacotes) oferecido pelos principais protocolos encontrados.

C. Intervalo entre pacotes dos fluxos RTP

Da análise dos protocolos verificou-se que a maior parte dos pacotes observados pertence à transmissão de voz codificada utilizando protocolo RTP. Assim sendo o estudo concentrou-se em analisar os dados para modelagem do tráfego RTP. A Tabela III mostra um sumário com estatísticas básicas sobre os intervalos entre chegada dos pacotes RTP da amostra.

A Figura 1 mostra a função de auto correlação para diversos níveis de sub-amostragem. A Figura 1(a) a (d) mostra, respectivamente, a função de autocorrelação para os níveis de sub-amostragem de 1.000.000 pontos, 100.000 pontos, 10.000 pontos e 1.000 pontos. O lento decaimento da função de auto-correlação (*ACF*, *auto-correlation function*) é um indicativo da auto-similaridade da série. O alto nível de auto-correlação observado impede o uso de distribuições de probabilidade, como Weibull ou Pareto, para a modelagem desta métrica de desempenho.

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. Percebe-se que mesmo a sub-amostragem de 1.000 pontos apresenta um alto grau de auto-correlação. A observação da série temporal, apresentada na Figura 2 indica que a série é estacionária (séries não estacionárias apresentam lento decaimento da função de auto-correlação).

A estimação do parâmetro de Hurst do tráfego agregado se fez através do método da transformada Wavelet utilizando-se as ferramentas disponibilizadas por Roughan et al. em [19]. Wavelets são funções matemáticas que seccionam os dados

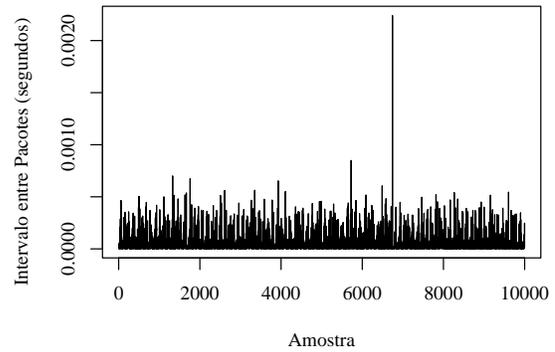


Fig. 2. Série temporal representando os intervalos entre pacotes RTP

TABELA IV
PARÂMETRO DE HURST PARA OS INTERVALOS ENTRE PACOTES

Tamanho da amostra	Parâmetro de Hurst	Intervalo de confiança para 95%
1.000.000	0.5766	[0.575, 0.578]
100.000	0.5807	[0.576, 0.585]
10.000	0.5935	[0.579, 0.609]
1.000	0.6207	[0.564, 0.678]

em diferentes componentes de frequência e permite o estudo de cada componente com uma resolução ajustada à sua escala. O método apresenta vantagens sobre os métodos tradicionais de Fourier para a análise de situações onde o sinal contém descontinuidades e picos. Wavelets foram desenvolvidas independentemente nos campos da matemática, física quântica, engenharia elétrica e geologia sísmica [20].

Os testes sobre os dados revelaram que a amostra apresenta auto-similaridade leve, com o parâmetro de Hurst de aproximadamente 0,58. Os resultados para diversos níveis de sub-amostragem são mostrados na Tabela IV.

D. Tamanho dos pacotes dos fluxos RTP de voz

A análise do tamanho dos pacotes dos fluxos RTP de voz mostra que o tamanho possui um comportamento praticamente discreto: foram detectados pacotes de tamanho igual a 64, 74, 84, 104 e 214 bytes. O histograma apresentado na Figura 3 mostra a frequência observada em cada um dos tamanhos citados. A Tabela V mostra um sumário de estatísticas básicas sobre o tamanho dos pacotes.

A análise da função de auto-correlação do tamanho de pacotes sucessivos (definir), apresentada na Figura 4, mostra um decaimento lento, indicativo da auto-similaridade da série.

A Tabela VI mostra o cálculo do parâmetro de Hurst para diversos níveis de sub-amostragem dos dados em estudo

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

Min.	1° Q	Mediana	Média	3° Q.	Max.	Var.
60.0	74.0	74.0	107.6	104.0	214.0	3364.72

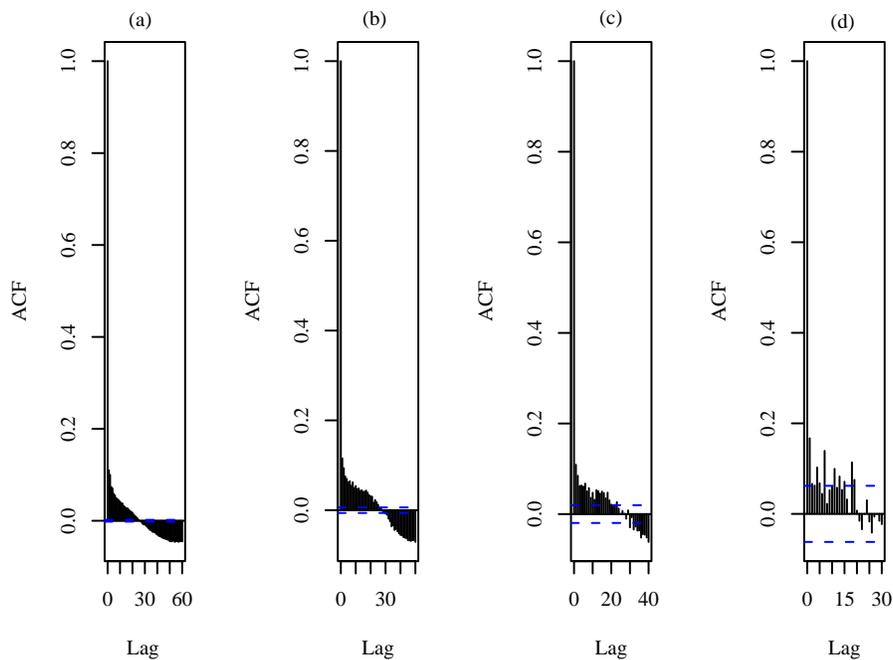


Fig. 1. Função de auto correlação

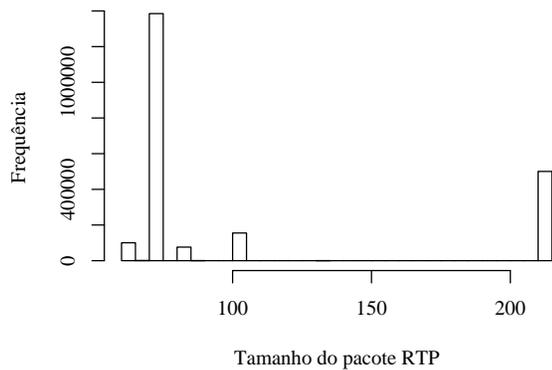


Fig. 3. Histograma do tamanho do pacote dos fluxos RTP

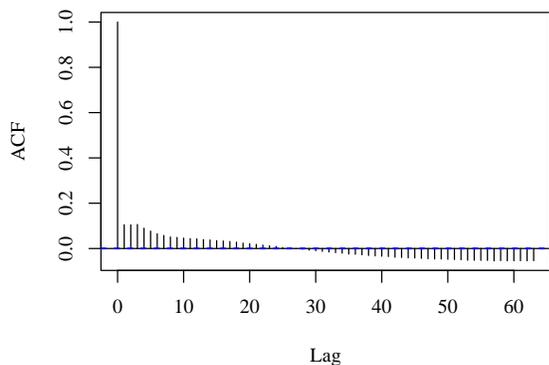


Fig. 4. Função de auto-correlação da série representada pelos tamanhos de pacotes sucessivos

TABELA VI
PARÂMETRO DE HURST PARA O TAMANHO DOS PACOTES

Tamanho da amostra	Parâmetro de Hurst	Intervalo de confiança para 95%
1.000.000	0.5661	[0.565, 0.568]
100.000	0.5777	[0.573, 0.582]
10.000	0.5922	[0.577, 0.607]
1.000	0.5906	[0.533, 0.648]

TABELA VII
CODECS IDENTIFICADOS

Codec	Número de pacotes	%
G.729	1.563.402	70.37%
G.711	500.423	22.52%
iLBC	148.262	6.67%
Outros	7.504	0.33%

através do método da transformada de Wavelet. O resultado indica que a série apresenta características auto-similares, com $H \approx 0.56$.

E. Codec utilizado pelos fluxos RTP

A Tabela VII mostra os principais codecs identificados no fluxo de dados amostrado. O codec mais utilizado foi o G.729, seguido pelo G.711.

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 correspondente a aproximadamente 90 segundos de tráfego, capturado diretamente no *call server* da operadora.

A principal contribuição deste artigo é a caracterização dos fluxos RTP, de fundamental importância para o estudo de sistemas VOIP. Os resultados mostram que os intervalos entre pacotes dos fluxos RTP podem ser caracterizados por um modelo auto similar. O tamanho dos pacotes observados apresentam uma característica bastante particular em termos dos valores discretos de tamanho observados, possivelmente devido aos *codecs* utilizados. O tamanho dos pacotes também possuem uma estrutura auto-similar.

Observou-se também que o *codec* mais utilizado foi o G.729, seguido pelo G.711. Esta característica está relacionada às capacidades dos aplicativos de usuários. O suporte ao G.711 é obrigatório em dispositivos VoIP e a utilização de outros *codecs* implica maior compactação do sinal de voz e conseqüente alteração nas características de tráfego geradas.

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 sintético com o perfil encontrado de fato em redes reais.

A seqüência deste trabalho está sendo executada com a análise de períodos maiores de tráfego, de modo a confirmar a análise aqui apresentada e incluir a caracterização do comportamento da sessão: tempo de duração da conversação e intervalo entre chamadas.

VII. AGRADECIMENTOS

Os autores agradecem a GVT pela disponibilização dos dados utilizados no estudo realizado.

REFERÊNCIAS

- [1] W. Willinger and K. Park. *Self-similar network traffic and performance evaluation*. John Wiley & Sons, New York, 1st edition, 2000.
- [2] A.Tyagi, H. De Meer, and J.K.Muppala. VoIP support on differentiated services using expedited forwarding. In *ICON 2000. IEEE International Conference on Networks 2000*, pages 574–580, February.
- [3] R.V.Cox. Three new speech coders from the itu cover a range of applications. *Communications Magazine, IEEE*, 35(9):40–47, September 1997.
- [4] ITU-T. Recommendation G.711. pulse code modulation (PCM) of voice frequencies, 1988.
- [5] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler. Sip: Session initiation protocol, 2002.
- [6] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. STD 64: RTP: A transport protocol for real-time applications, July 2003. See also RFC3550. Obsoletes RFC1889.
- [7] P.Brady. A technique for investigating on/off patterns of speech. *Bell Labs Tech Journal.*, January 1965.
- [8] 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.
- [9] A.Estepa, R.Estepa, and J.Vozmediano. A new approach for voip traffic characterization. *IEEE Communications Letters*, 8(10), October 2004.
- [10] 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.
- [11] Chen-Nee Chuah and R. H. Katz. Characterizing packet audio streams from internet multimedia applications. In *IEEE International Conference on Communications (ICC 2002)*, volume 2, pages 1199–1203, Apr-May 2002.
- [12] X.Liu and J.S.Baras. Aggregation of heavy-tailed on-off flows is multifractal. In *Technical Report CSHCN TR 2002-5 ISR TR2002*, 2003.
- [13] 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.
- [14] 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.
- [15] 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.
- [16] Angela Orebaugh, Gilbert Ramirez, Josh Burke, and Larry Pesce. *Wireshark & Ethereal Network Protocol Analyzer Toolkit (Jay Beale's Open Source Security)*. Syngress Publishing, 2006.
- [17] W.E. Leland, M.S. Qaqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of Ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, February 1994.
- [18] D. Kaplan and L. Glass. *Understanding non Linear Dynamics*. Springer, New York, 1st edition, 1998.
- [19] Matthew Roughan, Darryl Veitch, and Patrice Abry. On-line estimation of the parameters of long-range dependence. In *Proceedings Globecom '98*, volume 6, pages 3716–3721, Sydney, 1998.
- [20] Amara Graps. An introduction to Wavelets. *IEEE Comput. Sci. Eng.*, 2(2):50–61, 1995.