

Módulo de Análise de Parâmetros de Qualidade de Serviço em Aplicações de Áudio na Internet

Mário Luiz Tronco¹, Júlio César Magro², Tânia Regina Tronco Fudoli², Júlio Cezar Estrella¹

1 Universidade Estadual Paulista - Instituto de Biociências, Letras e Ciências Exatas -
Departamento de Ciências de Computação e Estatística
Rua Cristóvão Colombo, 2265 - CEP 15054 - 000 - São José do Rio Preto - SP - Brasil

2 Diretoria de Redes de Telecomunicações - Fundação CPqD
Rodovia Campinas Mogi Mirim km 118,5 - Campinas -SP

E-mails: mariot@dce.ibilce.unesp.br, magro@cpqd.com.br, tania@cpqd.com.br, jcezar2@lycos.com

Abstract

Internet Multimedia applications have increased in the last years. However, the quality of these applications still depends on mechanisms that provide end-user interactivity. Voice over IP traffic (VoIP), a real time application, requires certain levels of Quality of Service (QoS). Because the Internet is still a best-effort network, QoS parameters (related to delay, jitter and packet loss) must be measured and monitored, in order to provide feedback to applications. Packet Loss, by routers discarding packets in intensive traffic period, represents gaps in streams of reproduced audio, modifying the speech quality levels. There are many techniques to minimize the effect of these factors. In audio applications, the main approach is recovery of lost packet, in end-to-end way. This paper describes the implementation of a Parameters's Analysis module of Quality of Service, based on Neural Networks, which provides a measurement of "degree of congestion" (conditions to be faced by packets during audio playout in Packet Networks). These network analysis measurement can be used in packet loss recovery and retransmission based tools.

1. Introdução

O acesso à Internet com baixo custo e alta qualidade e a disponibilidade de um telefone são itens essenciais na sociedade atual. Existe uma mudança de paradigma já que as diversas formas de comunicação vem sendo digitalizadas e transportadas em redes de pacotes tais como Redes IP, ATM e Frame Relay. Como o tráfego de dados vem crescendo muito mais rápido que o tráfego telefônico, existe um interesse considerável no transporte de voz sobre redes de dados, em oposição ao tradicional transporte de dados sobre redes de voz. A qualidade da telefonia sobre IP vem se tornando um dos fatores determinantes para a

convergência das tecnologias de voz, vídeo e comunicações de dados. A praticidade de transportar voz pela Internet já foi demonstrada mas a disponibilização de produtos comerciais de alta qualidade, estabelecendo serviços públicos e convencendo usuários a adquiri-los dentro deste contexto, está apenas começando. Voz sobre IP ou, simplesmente, VoIP pode ser definida como a capacidade de fazer chamadas telefônicas (ou seja, fazer tudo o que se pode fazer hoje com as redes telefônicas comutadas) sobre redes de dados baseadas no protocolo IP, com uma qualidade de serviço (QoS) aceitável e uma relação custo/benefício elevada. Embora existam atualmente limitações no uso de voz sobre redes de pacotes, existe um considerável interesse pelo serviço e algumas experiências vem sendo realizadas. Está claro que um mercado se estabeleceu e existem oportunidades para desenvolvedores nesta área.

Existem diversos parâmetros para se caracterizar a Qualidade de Serviço para aplicações VoIP, entre eles o atraso e a perda de pacotes. Neste trabalho, propõe-se um Módulo, baseado em Redes Neurais, o qual atua no sentido de analisar o tráfego de pacotes em uma determinada aplicação, envolvendo VoIP, monitorando a Rede (Internet) e fornecendo em sua saída uma medida do grau de congestionamento que os pacotes (de áudio) deverão enfrentar nos próximos segundos. Este módulo, em conjunto com uma ferramenta de controle e recuperação de perdas de pacotes, permite uma melhor qualidade das aplicações VoIP.

2. Parâmetros de Rede e Avaliação de QoS para aplicações VoIP

Um requisito básico para o serviço de VoIP é fornecer um nível de qualidade de serviço (QoS) pelo menos igual ao do Sistema Telefônico Fixo Comutado (STFC). Embora QoS esteja relacionada à fidelidade da voz transmitida, também pode ser aplicada à disponibilidade da rede, disponibilidade característica de telefone (conferência, display de número de

chamada, etc.) e escalabilidade (qualquer um para qualquer um, universal, expansível). A qualidade de reprodução de áudio sobre a rede telefônica é fundamentalmente subjetiva. No entanto, medidas padronizadas tem sido desenvolvidas pela ITU (*International Telecommunication Union*). As medidas mais comuns para se avaliar a qualidade de voz são:

MOS (*Mean Opinion Score*), PSQM+ (*Perceptual Speech Quality Measurement Enhanced*), PAMS (*Perceptual Analysis Measurement System*) e PESQ (*Perceptual Evaluation of Speech Quality*). As recomendações P.800 e P.830 do ITU-T fornecem os guias para avaliação dos codecs de voz. Os parâmetros que mais afetam a qualidade da voz numa rede de pacotes são o atraso, a variação do atraso (*jitter*), a perda de pacotes e a codificação de voz [1], [2], [3].

Atraso

A latência e o atraso são parâmetros importantes para a qualidade de serviço das aplicações. Ambos os termos podem ser utilizados na especificação de QoS, embora o termo "latência" seja mais utilizado para equipamentos e o termo "atraso" seja mais utilizado nas transmissões de dados (por ex.: atrasos de transmissão, atrasos de propagação, etc). De maneira geral, a latência da rede pode ser entendida como o somatório dos atrasos impostos pela rede e equipamentos utilizados na comunicação. Do ponto de vista da aplicação, a latência (atraso) resulta em um tempo de resposta (tempo de entrega da informação - pacotes) para a aplicação. Atraso é o maior desafio da qualidade de serviço para voz em redes de pacotes e corresponde ao tempo necessário para transmitir os pacotes de dados da origem ao destino.

O atraso que ocorre nas redes IP é consequência do compartilhamento da largura de banda e do processamento nos roteadores e terminais. As aplicações de dados, para as quais as redes de pacotes foram desenvolvidas, são mais tolerantes ao atraso que as redes de voz, onde se considera que o máximo atraso tolerável, em um sentido de transmissão, é de 250ms. O atraso da rede consiste de atraso de propagação, atraso de processamento, atraso de transmissão e atraso nas filas que é causado por congestionamento. O atraso de propagação é constante e o atraso de transmissão é proporcional ao tamanho do pacote. O atraso pode ser classificado como fixo ou variável. O atraso fixo corresponde ao atraso fim-a-fim para qualquer pacote de voz, independentemente de pontos de congestionamento na rede. O atraso variável é incremental, causado por congestionamentos na rede ou nos *gateways*.

O atraso fixo pode ocorrer, para as aplicações de VoIP, nas seguintes áreas: compressão (codificação) processamento nas extremidades da rede, transmissão de dados internos, transmissão na rede, buffer (configurável) e decodificação. Em aplicações de VoIP, as amostras de voz analógicas são transmitidas em uma rede de dados, devendo primeiramente serem comprimidas e, em seguida, digitalizadas. Os atrasos de

compressão são estimados como sendo da ordem de 20 a 45 ms, dependendo do algoritmo de compressão utilizado. Ao chegar no destino, a descompressão é estimada como sendo da ordem de 10ms. Os atrasos de processamento nos gateways ou roteadores são de cerca de 10 ms em cada ponta (origem e destino). Atrasos de transmissão são altamente dependentes da velocidade do enlace, sendo quase desprezíveis (0,25ms) em cada ponta em enlaces T1/E1 e da ordem de 7ms em cada ponta, para enlaces de 56Kbps. Todos estes componentes formam o atraso fim-a-fim. Para uma conversação ser considerada interativa, este atraso não deve exceder certos limites: 150ms para a maioria das aplicações de tempo real, entre 150 e 400ms como potencialmente intolerável e acima de 400ms como inaceitável. A Tabela 1, a seguir, mostra o atraso fixo em VoIP.

Tabela 1. Atraso Fixo em VoIP.

| Processamento | Típico (ms) |
|--------------------------------------|-------------|
| Compressão | 20-45 |
| Inter-Processos (Origem) | 10 |
| Transmissão (Origem) | 0,25-7 |
| Transmissão na Rede (um sentido) | 10-200 |
| Transmissão (Destino) | 0,25-7 |
| Inter-Processos (Destino) | 10 |
| Controle de Buffer (Variável) | 10-80 |
| Decodificação | 20-45 |
| Total dos Atrasos Fixos (sem Buffer) | 91-403 |

O atraso variável muda, em tempo real, em função do congestionamento do tráfego na rede. Corresponde essencialmente, à soma dos atrasos nas filas, em cada comutador ou roteador intermediário na rede. Redes IP podem ser projetadas para minimizar o atraso, acrescentando-se largura de banda e reduzindo-se as aplicações que concorram entre si. Aumentar a largura de banda da rede, entretanto, irá reduzir as vantagens econômicas da rede IP. A especificação de requisitos de dimensionamento para uma rede VoIP do *European Telecommunication Standardization Institute* (ETSI), TR 101 329 v2.1.1 (1999-06) – “*Telecommunications and Internet Protocol Harmonization over Networks* (TIPHON); *General Aspects of Quality of Service* (QoS)” tem por escopo o estabelecimento de requisitos mínimos de QoS para uma rede de pacotes com serviços em tempo real, como o de voz. A Tabela 2 a seguir mostra os requisitos de atraso para um dos cenários considerados.

Tabela 2. Atraso para chamada originada no STFC, destinada ao STFC, cursada via rede IP [4].

| Classe de QoS | Requisitos de Atraso para Redes |
|---------------|---------------------------------|
| Ótima | < 150ms |
| Alta | < 250ms |
| Média | < 350ms |
| Baixa | < 450ms |

O *jitter* é outro parâmetro importante para a qualidade de serviço. Do ponto de vista de uma rede de

computadores, o *jitter* pode ser entendido como a variação de atraso na chegada de pacotes em relação ao atraso com o qual os mesmos foram transmitidos na rede pela origem devido à variação na latência (atrasos) da rede. Remover o *jitter* requer coletar pacotes e segurá-los o tempo suficiente para permitir que os pacotes mais lentos cheguem, de tal maneira que todos sejam entregues na seqüência correta, o que causa atraso adicional. A Tabela 3 relaciona os níveis de degradação da rede baseada no *jitter* usando valores do ETSI TIPHON.

Tabela 3 – Níveis de degradação da rede baseada no *jitter*.

| Categoria de Degradação da Rede | Pico de <i>jitter</i> (ms) | MOS |
|---------------------------------|----------------------------|-----|
| Perfeita | 8 | 4.5 |
| Boa | 75 | 4.0 |
| Média | 125 | 3.5 |
| Pobre | 225 | 3.0 |

As perdas de pacotes em redes IP ocorrem principalmente em função de fatores tais como: o descarte de pacotes nos roteadores e computadores por congestionamento (*buffer overflow*), instabilidade do roteamento (tal como mudanças de rota, falhas de links e links “instáveis” e a perda de pacotes devido a erros ocorridos na camada 2 (PPP - Point-to-Point Protocol, Ethernet, Frame Relay, ATM, ...) durante o transporte dos mesmos. Para se detectar perdas, todo pacote pesquisado deve ter um número de seqüência único. É difícil de saber, no entanto, se um pacote foi perdido se ele não chegou. Na prática, um valor de *timeout* é estabelecido. Para atrasos *one-way*, torna-se difícil usar *timeouts* se os relógios não estão sincronizados. Mas, se chegaram vários pacotes com um número de seqüência alto e o pacote esperado não chegou, é quase certeza que este pacote foi perdido. Do ponto de vista da qualidade de serviço da rede (QoS) a preocupação é normalmente no sentido de especificar e garantir limites razoáveis (taxas de perdas) que permitam uma operação adequada da aplicação. A maioria dos protocolos de aplicação de dados, tais como o TCP, automaticamente retransmitem os pacotes perdidos. Devido a sua característica de tempo real, as aplicações de VoIP utilizam os protocolos UDP e RTP, que não efetuam a retransmissão dos pacotes perdidos. A Tabela 4 relaciona a perda de pacote e o MOS usando valores do ETSI TIPHON.

Tabela 4 – Relação da Perda de pacote e MOS.

| Porcentagem de Perda | Qualidade de Voz | MOS |
|----------------------|------------------|-----|
| 3% | Boa | 4.2 |
| 15% | Média | 3.8 |
| 25% | Pobre | 3.0 |

2.1 Mecanismos de Recuperação e Controle de Perdas

Abordagens para evitar e recuperar perdas de pacotes, para uso na aplicação VoIP incluem métodos *fim-a-fim*, métodos *hop-by-hop* locais ou distribuídos e métodos utilizando uma combinação dos dois métodos anteriores. O método de recuperação de perdas *fim-a-fim* consiste em implementar mecanismos apenas na origem e no destino para fazer o controle de perda, sem fazer modificações na rede por onde os pacotes irão trafegar até o destino. Módulos de análise funcionando antes ou depois do codificador (*encoder*) extraem informação do sinal, trabalhando em conjunto com o processo convencional de codificação e “empacotamento” (*packetization*). A informação gerada pode então ser usada para influenciar a forma como os pacotes são gerados - *Interleaving* [2] ou pode ser adicionada aos dados a serem transmitidos - FEC [5], [6]. No receptor, essa informação adicional transmitida é extraída e as perdas de pacotes são detectadas. Então, logo que possível, a *stream* codificada é reconstruída e entregue ao decodificador. Tanto o processo de codificação quanto o de decodificação podem ser influenciados, minimizando o impacto das perdas. Finalmente, após o decodificador, o sinal pode ser processado para aumentar a qualidade do sinal. Métodos *hop-by-hop* utilizando a abordagem local de gerenciamento de fila [7], [2] tentam melhorar o mecanismo FIFO (First-In-First-Out) de descarte de pacotes em caso de *overflow*, levando em conta aspectos de *throughput* alcançável e atraso mínimo para todos os fluxos. Um exemplo é a técnica RED (Random Early Detection) a qual influencia a probabilidade de um descarte de pacote antes que a fila esteja cheia. Embora melhore o desempenho geral das redes de *best effort*, esta abordagem é limitada e possui baixo desempenho para tratar fluxos de comportamento variável. Abordagens distribuídas incluem, *IntServ* - Arquitetura de Serviços Integrados e *DiffServ* - Arquitetura de Serviços Diferenciados. As abordagens *fim-a-fim* e *hop-by-hop* combinadas utilizam mecanismos onde um fluxo usa alguma forma de QoS inter-fluxo (reserva por fluxo / priorização por pacote) embora não para a largura de banda do fluxo necessária. Então mecanismos de FEC/cancelamento podem adaptá-lo e explorar as condições de perda modificadas em um contexto *fim-a-fim* [8], [9].

3. Implementação do Módulo Análise de Parâmetros de QoS

A figura 1 mostra esquematicamente o Módulo de Análise de Parâmetros de Qualidade de Serviço proposto atuando em conjunto com uma ferramenta de VoIP em uma conversação via Internet.

O módulo, uma vez atuando em tempo real, em um aplicativo de VoIP, analisa o comportamento da rede a cada instante de tempo pré-determinado, emitindo uma informação que poderá ser utilizada pelo mecanismo de recuperação de perdas (FEC). Essa informação será interpretada pela FEC como sendo um padrão do

comportamento da rede naquele instante e pode ser utilizada para controlar a quantidade de informação redundante enviada pelo mecanismo de FEC ou para gerar uma ordem de retransmissão de pacotes. O detalhamento do Módulo é mostrado a seguir.

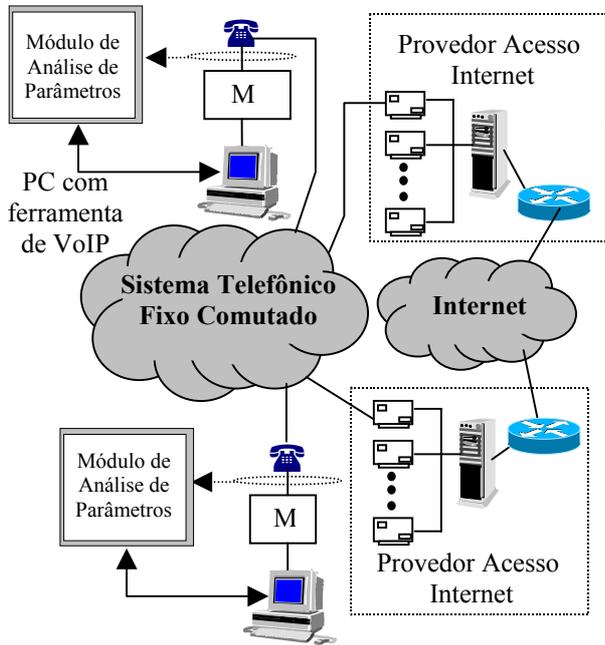


Figura 1. Módulo de Análise de Parâmetros de Qualidade de Serviço em aplicação VoIP.

3.1 Detalhamento do Módulo

O Módulo, mostrado na figura 2 a seguir, foi projetado para, através da coleta de informações provenientes da Rede (Internet), analisar seu comportamento e gerar em suas saídas uma medida de seu grau de congestionamento. Sua estrutura contempla sub-módulos (um para cada parâmetro a ser analisado) e uma Rede Neural. A Rede Neural utiliza as saídas dos sub-módulos como entradas e gera a saída desejada (medida do grau de congestionamento da Rede) em função de exemplos aprendidos através do treinamento com padrões de congestionamento conhecidos.

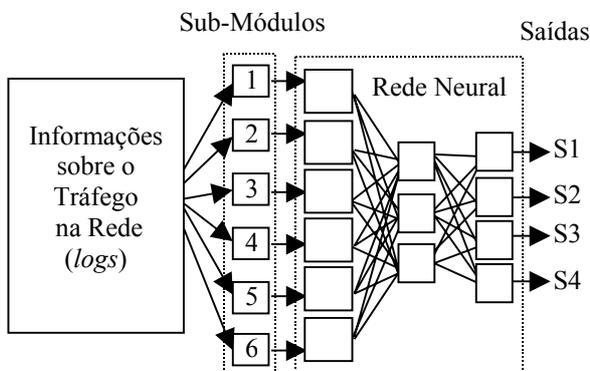


Figura 2. Detalhamento do Módulo. O sistema é composto por seis sub-módulos:

- Sub-Módulo de Análise de Perdas de Pacotes;
- Sub-Módulo de Análise de Atraso dos Pacotes;
- Sub-Módulo de Análise de Rajadas de Perda de Pacotes;
- Sub-Módulo de Análise de Comprimento de Rajada de Perda;
- Sub-Módulo de Análise de *Jitter*;
- Sub-Módulo de Análise de Probabilidade de Perda;

As entradas para estes módulos são obtidas, do tráfego real na Rede, através de uma ferramenta de captura de pacotes, a qual gera os parâmetros necessários para cada módulo. A figura a seguir mostra o detalhamento dos sub-módulos.

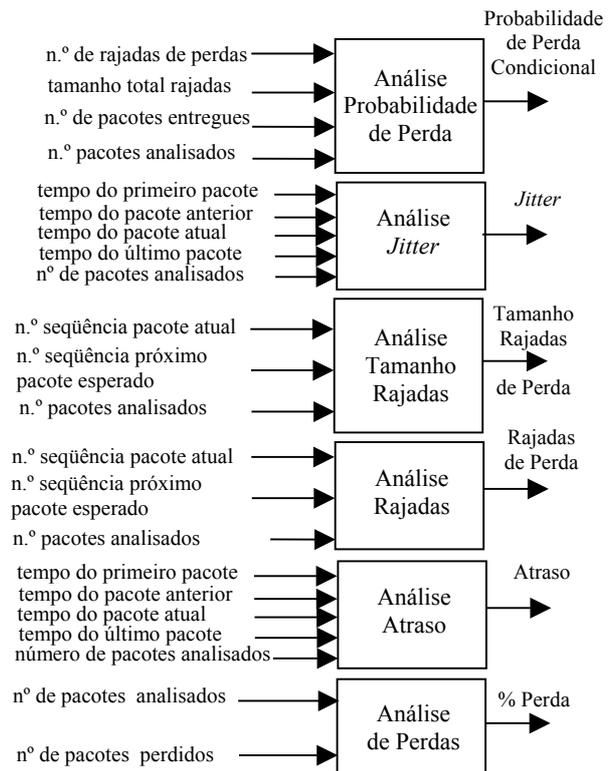


Figura 3. Detalhamento do Módulo.

Através de padrões conhecidos de congestionamento da Rede (Internet), é feito o treinamento da Rede Neural, forçando suas saídas a se comportarem como um código representativo do mesmo, como mostrado a seguir.

Tabela 5. Relação entre as saídas da Rede Neural e a condição do tráfego na Rede (Internet).

| S1 | S2 | S3 | S4 | Condição |
|----|----|----|----|--------------------------|
| 0 | 0 | 0 | 0 | Sem congestionamento |
| . | . | . | . | . |
| 0 | 1 | 1 | 1 | Congestionamento Crítico |
| . | . | . | . | . |
| 1 | 1 | 1 | 1 | Congestionamento Máximo |

Os padrões utilizados para treinamento da Rede Neural foram obtidos utilizando-se o aplicativo *ott_mon* [10] para capturar pacotes que trafegavam na Rede, durante uma conversação envolvendo VoIP. A utilização deste aplicativo permite salvar as informações referentes à conversação em um arquivo de log (ASCII), de registros de pacotes UDP. A figura 4 mostra um exemplo dos dados obtidos pela ferramenta *ott_mon* plotados em função do tempo.

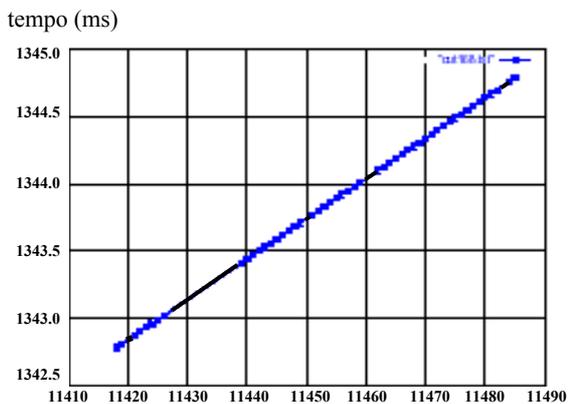


Figura 4. Pacotes recebidos em função do tempo.

Na figura 4 temos a seguinte situação:

- Perda do pacote 11420
- Perda no intervalo de [11427, 11438]
- Perda do pacote 11450
- Perda no intervalo de [11460, 11461]
- Perda do pacote 11483

Estatísticas:

| | |
|---|----------|
| Número de pacotes analisados..... | 66 |
| Número de pacotes entregues..... | 49 |
| Número de pacotes perdidos..... | 17 |
| Taxa de pacotes perdidos..... | 0.257576 |
| Atraso médio entre pacotes..... | 0.014844 |
| Número de rajadas de perda..... | 5.0 |
| Comprimento total das rajadas de perda..... | 17 |
| Média do jitter..... | 0.006711 |
| Probabilidade de Perda Condicional..... | 0.705882 |

Saídas Normalizadas:

| Atraso | Taxa | Rajadas | Compr. Rajadas | Média Jitter | Prob. Perda |
|----------|----------|----------|----------------|--------------|-------------|
| 0.037111 | 0.257576 | 0.500000 | 0.548387 | 0.009278 | 0.705882 |

Os valores obtidos (Estatísticas) são normalizados para que os sub-módulos possam gerar saídas entre 0 e 1.

Foram capturados, utilizando a ferramenta *ott_mon* em uma seção de conversação via VoIP, 294 gráficos como o mostrado anteriormente, representando

situações de tráfego real na Internet, em período de grande utilização. Destes gráficos, 53 foram selecionados, por representarem situações particulares de congestionamento, para serem utilizados no treinamento da Rede Neural. Tais padrões representam as possíveis situações em que a ferramenta de Controle e Recuperação de Perdas terá que trabalhar. A Rede Neural foi implementada utilizando *Multilayer Perceptron* com *Backpropagation*, com um total de seis unidades de entrada (ligadas às saídas dos sub-módulos) e quatro unidades de saída (que geram o código representativo de congestionamento da Rede). Após treinada a Rede Neural, foram testados diversos padrões, obtidos da análise dos pacotes recebidos em uma aplicação real de VoIP, sob condições normais de utilização da Internet. Os resultados experimentais obtidos são mostrados a seguir.

4. Resultados Experimentais

Os testes realizados com dados reais obtidos da análise de pacotes recebidos em uma aplicação envolvendo a conversação via aplicativo de VoIP mostraram uma taxa alta de acerto (erro inferior a 10%, em média), o que representa a geração de códigos de saída compatíveis com as situações reais de tráfego presentes na Rede.

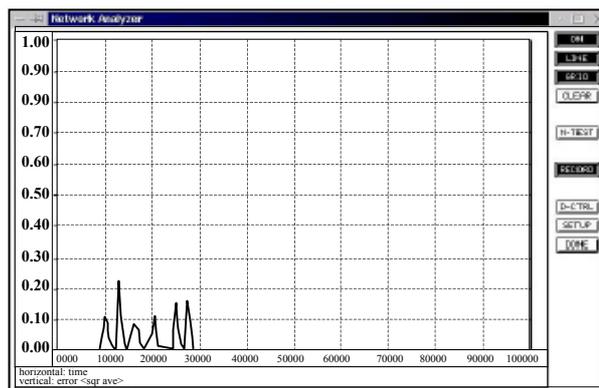


Figura 5. Taxa de Erro da Rede Neural na análise dos padrões de teste.

Os picos de erro representam três padrões que foram classificados de forma incorreta, porém gerando na saída do Módulo um código de saída muito próximo daquele que seria considerado o mais correto para a situação correspondente.

5. Conclusões

Para a implementação do Módulo proposto, foi realizada uma seção de conversação interativa utilizando um aplicativo de VoIP, através do uso de modem discado (via Sistema Telefônico Fixo Comutado). Juntamente com esse aplicativo, foi executada uma ferramenta para capturar os pacotes durante a conversação. A implementação dos sub-

módulos permitiu a extração de informações representativas de perdas e atrasos entre pacotes. Essas informações forneceram uma medida do grau de congestionamento da rede, considerando um intervalo de tempo pré-determinado, durante a reprodução do discurso. A implementação da Rede Neural para o reconhecimento dos padrões de congestionamento tornou representativas as informações fornecidas pelo Módulo, atuando como um indicativo de uma determinada situação de perda e atraso na rede de pacotes.

A continuidade do trabalho em curso envolve a integração do Módulo desenvolvido em uma ferramenta de VoIP que trabalhe com técnicas de controle e recuperação de perdas. Esta integração já vem sendo realizada, através da implementação de modificações no aplicativo Gphone [11]. Estas modificações permitem ao Gphone avaliar as saídas do módulo apresentado neste trabalho, utilizando-as de forma adaptativa no controle de retransmissão de pacotes e inserção de informação redundante nos mesmos (geração de redundância FEC). A figura 6 mostra esta integração.

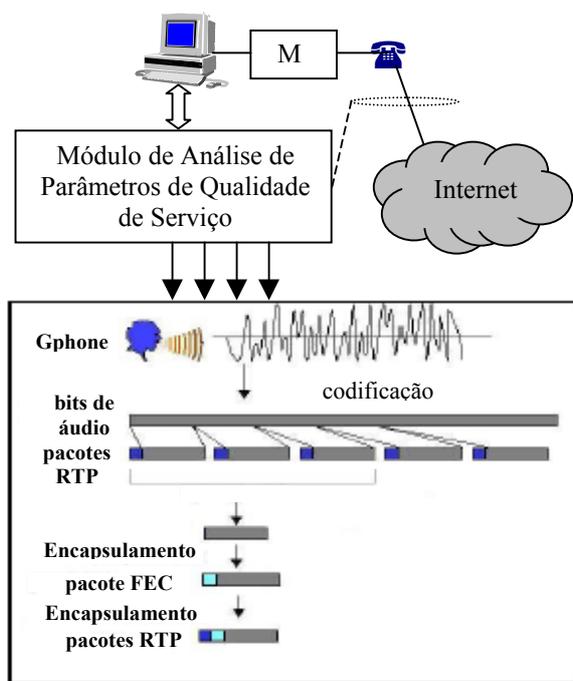


Figura 6. Integração do Módulo de Análise de Parâmetros de Qualidade de Serviço com a ferramenta Gphone modificada.

6. Referências Bibliográficas

[1] Rosenberg, J. D.; Qiu L.; Schulzrinne, H. Integrating packet FEC into adaptive voice playout buffer algorithms on the Internet". In *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, 2000. URL:

- [2] Sanneck, H.; Stenger, A.; Younes, K. B.; Girod, B. A New Technique for audio Packet Loss Concealment. *Proceedings IEEE Global Internet* (Jon Crowcroft e Henning Schulzrinne, eds.), pp. 48-52, Londres, 1996.
- [3] Sanneck H.; Carle, G.; Kooddli., R. A framework model for packet loss metrics based on loss runlengths. *SPIE/ACM SIGMM Multimedia Computing and Networking Conference*, 2000.
- [4] ETSI TR 101 329 v2.1.1 (1999-06): "Telecommunications and Internet Protocol Harmonization over Networks (TIPHON); General Aspects of Quality of Service (QoS)"
- [5] Rosenberg, J. D.; Schulzrinne, H. An RTP Payload Format for Generic Forward Error Correction. *RFC2733 IETF*, December, 1999. <ftp://ftp.ietf.org/rfc/rfc2733.txt>.
- [6] Hardman, V.; Sasse M.; Handley, M.; Watson, A. Reliable Audio for use over the Internet. *Proceedings INET*, 1995.
- [7] Floyd, S.; Jacobson, V. "Random Early Detection Gateways for Congestion Avoidance". *IEEE/ACM Transactions on Networking*, 1(4): 397-413, August, 1993.
- [8] Garcia, A. V. Control Mechanisms for Audio Transmission over the Internet. PhD Thesis, Nice University, Nice, France, October, 1996.
- [9] Shacham, N.; Mckennney, P. Packet Recovery in High-Speed Networks using Coding and Buffer Management. *Proceedings ACM SIGCOMM*, pp 124-131, San Francisco, California, June, 1990.
- [10] Wenyu, J.; Schulzrinne, H. QoS Measurement of Internet Real-Time Multimedia Services. Technical reports. URL: http://www.cs.columbia.edu/~wenyu/research/qos_tools.tar.gz.
- [11] GPHONE - Gnome-o-Phone Internet telephone. <http://gphone.sourceforge.net>