

ANÁLISE DE ALGORITMOS PARA REPRODUÇÃO EM TEMPO REAL DE VOZ EM REDES DE PACOTES

T.L.P. Bastos * **

L.F.G. Soares *

** Centro Científico Rio IBM Brasil Caixa
Postal 4624 20.001, Rio de Janeiro, RJ - Brasil

* Departamento de Informática Pontifícia
Universidade Católica do RJ R. Marquês de S.
Vicente, 225 22.453, Rio de Janeiro, RJ - Brasil

Sumário

A variação estatística do retardo em uma rede comutada por pacotes pode gerar discontinuidades em uma recepção de voz chegando a afetar seriamente a sua inteligibilidade, ou mesmo inviabilizando a comunicação. Para atenuar estes efeitos indesejados foram desenvolvidos algoritmos para reprodução em tempo real da voz transmitida, cuja proposta é compensar a variação estatística do retardo de transmissão atrasando o início da reprodução a fim de garantir uma reserva razoável de amostras de voz.

No presente trabalho é apresentada uma análise de desempenho comparativa de quatro diferentes algoritmos de reprodução, um dos quais baseado em uma proposta desenvolvida no laboratório de Comunicação de Dados da PUC-Rio. A idéia é determinar a sensibilidade e adequação de cada um dos algoritmos, dadas as características temporais do ambiente em que a comunicação se baseia. Para tal, foram desenvolvidos modelos para simulação do sistema, baseados em redes de filas estendidas, compreendendo a fonte de voz, a rede de transmissão e a recepção de pacotes segundo os diferentes algoritmos. Através da análise comparativa foi possível constatar que a maioria dos algoritmos apresenta uma sensibilidade muito alta em relação ao coeficiente de variação do tempo de resposta da rede. A avaliação dos efeitos da escolha do valor para os parâmetros de ajuste dos algoritmos sobre o desempenho permitiu complementar o perfil de robustez.

Abstract

The stochastic nature of delays experienced by voice packets in packet-switching networks can originate discontinuities in output voice streams affecting the intelligibility and sometimes inviabilizing voice communication. Playout strategies for voice transmission have been proposed to remove such variations in delay. This paper presents a comparison of four different voice playout strategies, one of them proposed by the Data Communication Laboratory at PUC-Rio.

The goal of this work is to investigate the sensibility and adequacy of each strategy with respect to the temporal characteristics of the communication environment. Simulation models of the system were made modelling the voice source, the transmission network and each playout strategy. Through the comparative analysis it was possible to observe the high sensibility of most of the studied strategies in respect to the coefficient of variation of the transmission delay. The adjustment of the strategies' parameters was also investigated in order to complement the study of each strategy.

1 Introdução

O tráfego de voz, apesar de contínuo, exibe uma característica de chaveamento entre períodos de presença efetiva de voz, denominados *surtos de voz*, e períodos de silêncio, denominados *pausas*. As pausas compreendem os intervalos de silêncio entre parágrafos, frases, palavras, sílabas etc. Cerca de 45% do tempo total de duração de um monólogo corresponde a pausas. Constata-se também que, para diálogos esse percentual sobe para 60-65% do tráfego gerado por cada interlocutor [9].

A transmissão de voz por chaveamento de pacotes se mostra atraente devido à melhor utilização da capacidade do canal, se valendo da característica chaveada do sinal de voz, onde apenas os pacotes que carregam segmentos dos surtos de voz são transmitidos, e um período de falta de pacotes chegando ao destino é interpretado na reprodução como sendo uma pausa. Entretanto, apesar da voz ser gerada à taxa constante, o processo que descreve a transmissão via rede de pacotes introduz um retardo aleatório aos pacotes, cujo valor depende da carga a que a rede está momentaneamente submetida. Como a reprodução da voz deve se dar à mesma taxa de geração, a variação estatística do retardo de transmissão pode causar descontinuidades no fluxo da voz reproduzida. A *Figura 1* ilustra o processo de empacotamento, transmissão e recepção dos pacotes de voz no destino. Mesmo que se assuma, por hipótese, que a ordem de geração dos pacotes não é alterada pela transmissão, o tempo entre chegadas se disassocia totalmente do tempo entre geração de pacotes dificultando a recuperação da voz para reprodução em tempo real. Sempre que não houver pacotes disponíveis para reproduzir, o receptor assume que se trata de um período de silêncio. Este silêncio indevidamente introduzido é responsável pela quebra da continuidade da voz.

Para atender às exigências de manutenção da continuidade na reprodução garantindo a comunicação da voz em tempo real, é preciso implementar alguns procedimentos especiais no receptor a fim de compensar a variação estatística dos retardos. A idéia em que se baseiam estes procedimentos é eliminar, ou reduzir a níveis aceitáveis, a ocorrência de falta de pacotes durante o processo de reprodução.

A estratégia utilizada pelos algoritmos existentes para reprodução de voz em tempo real baseia-se fundamentalmente em assegurar uma reserva de pacotes antes de dar início ao processo de reprodução, introduzindo um retardo inicial a cada surto de voz. Aparentemente o problema estaria resolvido escolhendo o retardo inicial bem grande, garantindo que todos os pacotes do surto já estivessem disponíveis na fila da reprodução. Entretanto, o valor deste retardo está limitado pelo máximo retardo absoluto permitido para o sinal de voz sem que haja perda de naturalidade da voz. O problema do retardo absoluto torna-se crítico no caso de conversações, onde cada interlocutor precisa detectar o término do discurso do outro a fim de dar início à sua fala. Cada algoritmo proposto trata de maneira distinta a formação desta reserva e o encaminhamento dos pacotes para a reprodução.

Neste trabalho é apresentado um estudo comparativo do desempenho de quatro diferentes algoritmos de recepção de pacotes visando compensar a variação estatística do retardo sofrido durante o processo de transmissão, sendo um dos algoritmos oriundo de uma proposta apresentada em [1], implementado no laboratório de Comunicação de Dados do Departamento de Informática da PUC-Rio. Com base nos resultados obtidos através de simulação foi traçado o perfil de cada algoritmo para diferentes graus de carga na rede, tornando possível avaliar sua robustez, dadas as características temporais da rede de comunicação. Para avaliação de desempenho foram desenvolvidos modelos em redes de filas estendidas para cada um dos componentes envolvidos no processo de comunicação de voz via rede de pacotes, ou seja, a *origem*, a *rede de comunicação* e a *recepção*. A *origem* compreende uma fonte que gera e empacota o sinal de voz. A *rede de comunicação* é

modelada como um servidor de *Erlang* com parâmetro k variável, permitindo alterações no coeficiente de variação da rede, cujo tempo de serviço é o próprio tempo de resposta da rede. Por fim, a *recepção* encerra o tratamento dado aos pacotes por cada um dos algoritmos antes de submetê-los à reprodução. Quatro são os modelos de *recepção*, cada um descrevendo um dos algoritmos enfocados.

Na segunda seção são apresentadas as características do sinal de voz e os algoritmos de compensação a serem estudados. A terceira seção descreve em linhas gerais o modelo desenvolvido do sistema e o tipo de medidas extraídas de sua simulação. A quarta seção apresenta os resultados levantados com base nas medidas realizadas, incluindo a análise comparativa dos quatro algoritmos. A última seção é reservada à apresentação das conclusões.

2 Compensação da Variação Estatística do Retardo

2.1 Características do Sinal de Voz

A variação estatística do retardo sofrido pelos pacotes vai atuar em diferentes pontos do sinal de voz. Reportando novamente à *Figura 1*, tanto se observa o surgimento de falsas pausas dentro de surtos de voz, quanto alterações na duração de pausas entre surtos. A ocorrência de cada um destes eventos traz como consequência uma descontinuidade no fluxo da voz recebida. O efeito de cada um destes eventos sobre a qualidade do surto de voz deve ser avaliado cuidadosamente a fim de estabelecer padrões aceitáveis que mantenham as características temporais da voz inalteradas. É importante que estes padrões sejam definidos para que se possa avaliar a validade da compensação realizada por um dado algoritmo de reprodução.

Alguns estudos foram desenvolvidos em cima dos parâmetros temporais do sinal de voz, visando principalmente descrever modelos adequados para os surtos de voz e as pausas. Os efeitos da variação estatística de retardos sobre o sinal de voz transmitido foram pouco estudados, principalmente no que se refere à introdução de períodos de silêncio dentro de surtos de voz, pois transmissão de voz via rede de pacotes só agora começa a ser empregada.

Com relação às pausas originalmente presentes no sinal de voz transmitido, alterações de $\pm 50\%$ na sua duração são toleráveis, não implicando em perda de inteligibilidade da voz. As pausas de curta duração, compreendendo silêncio entre sílabas e palavras, merecem um cuidado especial, uma vez que podem sofrer uma variação de até 100% , ou seja, serem suprimidas. Uma possível solução para este problema é agregar aos pacotes uma informação que permita verificar a duração das pausas existentes. No entanto, a transmissão de voz se dá sem recuperação de erros, já que o sinal de voz apresenta uma boa tolerância a perdas. Com isso, basta que pacotes sejam perdidos para que informações sobre a duração de algumas pausas se percam. Por este motivo é adotada uma outra técnica para resguardar as pausas curtas de alterações muito grandes em sua duração. A solução alternativa é submeter o sinal de voz à operação de *hangover* antes de ser transmitido. A operação de *hangover* consiste em reduzir a duração de todas as pausas de um fator n , que é chamado intervalo de *hangover* [9]. Com isso, todas as pausas menores ou iguais a n desaparecem, sendo incorporadas aos surtos de voz vizinhos, todos os surtos de voz terminando com um período de silêncio de duração n . Escolhendo um valor adequado para n é possível evitar que, em média, a variação do retardo cause alterações além do permitido na duração das pausas existentes [5].

O efeito decorrente da introdução de períodos de silêncio indevidamente durante a reprodução de surtos de voz não foi completamente investigado. Os levantamentos de que

se dispõe não são completos o bastante para que se estabeleça um padrão fechado que garanta a qualidade da voz reproduzida. A frequência com que pausas são introduzidas indevidamente deve ser mantida baixa de forma a garantir a boa qualidade de voz. Estudos subjetivos indicam que a introdução de períodos de silêncio com duração inferior a 50ms são toleráveis. Além disso, a introdução de pausas com mais de 50ms de duração não deve ocorrer em mais de 1% dos pacotes [7]. Nota-se, entretanto, que nada é mencionado sobre a distribuição destas pausas ao longo do surto de voz.

Outra característica importante do sinal de voz, como mencionado anteriormente, é sua boa tolerância a perdas. Perdas médias abaixo de 2% da informação contida nos surtos de voz não acarreta problemas na recuperação do sinal de voz para reprodução [5,7]. Esta flexibilidade é de grande utilidade, pois fornece uma margem de segurança para as exigências descritas acima. Por exemplo, pacotes recebidos fora dos padrões aceitáveis não representam problema desde que estejam dentro do percentual de perdas. Os algoritmos de compensação tomam vantagem desta característica da voz.

Uma questão relevante diz respeito ao retardo absoluto que o sinal de voz acaba sofrendo devido à transmissão e espera pela reprodução no destino. Este problema, entretanto, já foi exaustivamente estudado, uma vez que a questão do retardo absoluto surgiu com a transmissão analógica da voz via rede de telefonia. O padrão brasileiro de telefonia estabelece limites para o retardo absoluto de 40ms, para distâncias continentais, e de 80ms, para distâncias intercontinentais, a fim de garantir a naturalidade da voz. Este tipo de efeito indesejável vai interferir no andamento de conversações, dificultando que um interlocutor perceba o término da fala do outro. O algoritmo utilizado na recepção dos pacotes para compensar a variação do retardo dos pacotes de voz tem que se preocupar também em respeitar o limite de 40ms para o retardo absoluto, pensando no ambiente de redes locais. Por este motivo, atrasar a reprodução do primeiro pacote de cada surto de voz a fim de garantir a continuidade não é uma solução aceitável na maioria dos casos, uma vez que o limite para o retardo tem que ser respeitado.

2.2 Algoritmos de Compensação

Como já foi mencionado, os algoritmos de compensação procuram formar uma reserva razoável de pacotes de voz a fim de garantir a continuidade da voz reproduzida, retardando o início da reprodução do primeiro pacote de cada surto de voz, a fim de possibilitar o armazenamento dos próximos pacotes. Como será mostrado mais adiante, cada algoritmo vai apresentar um grau diferente de sensibilidade à escolha deste retardo inicial.

Os algoritmos selecionados para análise de desempenho não exigem que os pacotes de voz carreguem informação sobre o retardo que sofreram na transmissão, não exigindo sincronismo da rede. Esta escolha baseou-se num estudo de simulação descrito em [8], que compara o desempenho de algoritmos de compensação que exigem sincronismo da rede com outros que não precisam de informação sobre o retardo experimentado pelos pacotes. Os resultados mostram que não há diferença substancial no desempenho entre estes dois tipos de algoritmos. Com isso, aqueles que não exigem sincronismo da rede tornam-se, obviamente, mais atraentes.

Segue abaixo uma descrição rápida de cada um dos quatro algoritmos enfocados neste trabalho:

• Algoritmo 1

Este algoritmo, desenvolvido no MIT [7], retarda a reprodução do primeiro pacote de cada surto de voz de um tempo D (Figura 2). Os demais pacotes do surto de voz só são reproduzidos caso o retardo sofrido por eles na rede não exceda o retardo total (retardo na rede somado ao retardo inicial D) sofrido pelo primeiro pacote do surto.

- **Algoritmo 2**

Esta proposta surgiu no Instituto de Ciência da Informação da USC [7] (*Figura 3*). A reprodução do primeiro pacote de cada surto também é retardada de um tempo D . Os outros pacotes do surto de voz são reproduzidos respeitando a ordem de chegada, independente do retardo sofrido na rede.

- **Algoritmo 3**

Este algoritmo foi empregado no projeto UNIVERSE desenvolvido no laboratório de Computação da Universidade de Cambridge [6] (*Figura 3*). Ele reproduz os pacotes seguindo o mesmo critério empregado pelo *Algoritmo 2* com a diferença de que o retardo inicial D é ajustado dinamicamente. Durante a reprodução é utilizado o conceito de *reserva*, que compreende o número de amostras de voz na fila de reprodução. A reserva negativa (ou *déficit*) indica a quantidade de *gaps* introduzidos no fluxo da voz desde a reprodução do último pacote. A intervalos regulares o mecanismo de ajuste é ativado, verificando as *reservas* e *déficits*. Se existem *déficits*, nenhuma ação é tomada, apenas são zeradas as filas de *reserva* e *déficit*. Se não há *déficits* e a *reserva* é grande o bastante, a reprodução é acelerada do tempo equivalente à metade do tamanho da *reserva*, encurtando o próximo intervalo de silêncio deste valor.

- **Algoritmo 4**

Este algoritmo foi desenvolvido no laboratório de Comunicação de Dados da PUC-Rio [1] (*Figura 4*). Sua principal característica é fornecer um limite superior para o retardo de cada pacote na fila de espera pela reprodução. Como os demais algoritmos, o primeiro pacote do surto de voz é retardado de um tempo D . A cada chegada de um novo pacote do surto de voz, é estimado seu tempo de espera até poder ser reproduzido, baseado no estado atual da fila de reprodução. Se a espera estimada ultrapassar $2D$, a reprodução em andamento é avançada da quantidade necessária a fim de garantir que o tempo de espera do pacote recém chegado seja de $2D$. Os segmentos de voz são descartados em ordem decrescente de antiguidade, visando preservar sempre a informação de voz mais recente. Caso o novo pacote não encontre nenhum pacote em reprodução, este começa a ser reproduzido imediatamente. Como no *Algoritmo 3* existe também um mecanismo de ajuste dinâmico de D [11], mas que não é levado em consideração no presente artigo.

Todos os algoritmos apresentados encaram o retardo inicial D como um parâmetro que lhes é fornecido por um agente externo, de acordo com as necessidades do ambiente em que está baseada a comunicação de voz. Nas *Figuras 2, 3 e 4*, as setas abaixo dos pacotes indicam a chegada do respectivo pacote, que é identificado por um número. As setas acima dos pacotes assinala o início da reprodução do respectivo pacote.

3 Modelagem do Sistema

3.1 Hipóteses do Modelo

Para estudar cada um dos algoritmos foi necessário modelar o processo de comunicação de voz via rede de pacotes. O modelo foi separado em três módulos independentes, a *origem*, a *rede de comunicação* e a *recepção*, que se comunicam através de suas entradas e saídas. Os modelos foram desenvolvidos na linguagem de simulação RESQ, baseada em redes de filas estendidas.

O primeiro módulo, a *origem* da comunicação, modela a fonte de voz e o processo de empacotamento. A modelagem do sinal de voz baseia-se no estudo descrito em [9], que levantou as distribuições e parâmetros temporais de surtos de voz e pausas característicos de monólogos, abrangendo também considerações sobre o efeito da aplicação de *hangover* sobre o sinal de voz. Modelar a fonte de voz com as características temporais do

monólogo não invalida os resultados obtidos na avaliação de desempenho, principalmente assumindo-se que apenas um interlocutor fala por vez, ou seja, é desprezível a fala simultânea. A operação de *hangover* foi considerada, pois, como mencionado, visa resguardar as pausas menores da variação estatística do retardo. Apenas os surtos de voz são empacotados e transmitidos. O tamanho dos pacotes é fixo e os últimos pacotes de cada surto podem, eventualmente, carregar silêncio pertencente à próxima pausa. As amostras de voz são empacotadas imediatamente após sua geração, não introduzindo retardos no empacotamento.

De acordo com as considerações acima, seguem abaixo as distribuições e médias utilizadas para modelar a *origem*.

- Duração das Pausas: distribuição *exponencial* e média de 600 milissegundos.
- Duração dos Surtos de Voz: distribuição *exponencial-mais-constante*; a média da exponencial é 2.3 segundos e a contante, que compreende o intervalo de *hangover* utilizado, é 200 milissegundos.

O segundo módulo compreende a *rede de comunicação*, que recebe como entrada a voz devidamente empacotada vinda da *origem*. A rede é modelada como um centro de serviço ativo com um único servidor, cujo tempo de serviço corresponde ao tempo de resposta da rede. A variação estatística do retardo de transmissão é uma medida da distorção no fluxo de pacotes de voz que chega ao destino, estando associada ao coeficiente de variação da distribuição do tempo de serviço deste servidor. Como a idéia é investigar o desempenho dos algoritmos em compensar as variações de retardo dos pacotes transmitidos, optou-se pela distribuição de *Erlang-k*, cujo coeficiente de variação é dado por $1/\sqrt{k}$. Isto permite que se altere convenientemente o valor do coeficiente de variação mexendo unicamente no parâmetro *k* sem afetar o valor da média. A fim de que as médias utilizadas no modelo da rede fossem próximas da realidade, os valores utilizados para o tempo médio de resposta da rede foram extraídos de um estudo sobre retardos experimentados por pacotes de voz transmitidos via uma rede *Token-ring* [2] (vide *Figura 5*). Assume-se, ainda, que há apenas tráfego de voz na rede e que o número de canais de voz está maximizado. Além disso, não há perdas de pacotes na transmissão.

O último módulo compreende a *recepção* dos pacotes de voz que já passaram pelo servidor de *Erlang-k* e foram retardados. Existem quatro modelos diferentes de *recepção*, um para cada algoritmo de compensação. Em todos os modelos de recepção, o *buffer* de recepção é suposto ilimitado e os retardos envolvidos no processo de recepção dos pacotes da rede são desprezíveis face ao retardo de transmissão e de espera na fila de reprodução.

3.2 Medidas de Desempenho

As medidas extraídas das simulações têm como finalidade verificar a qualidade do sinal de voz reproduzido, ou seja, avaliar se a compensação estatística do retardo conseguiu trazer para padrões razoáveis as distorções no sinal de voz. De acordo com as características levantadas do sinal de voz, foram efetuadas seis medidas em cima da voz recebida e tratada pelos diferentes algoritmos. As medidas são descritas nos tópicos abaixo:

• Reprodução dentro dos Limites de Retardo Absoluto

Como já foi salientado anteriormente, o retardo absoluto do sinal de voz não deve ultrapassar um certo valor máximo, sob o risco de perda da naturalidade da voz. Espera-se, então, que os algoritmos de compensação reproduzam apenas pacotes de voz cujo retardo absoluto não ultrapasse este limite. Esta medida vai registrar o percentual da informação de voz dentro dos padrões aceitáveis de retardo absoluto reproduzido por cada algoritmo.

• Superposição de Surtos de Voz

Esta medida avalia a interferência de surtos de voz consecutivos na reprodução. Todos os algoritmos ajustam seus parâmetros, a maioria retornando a condições iniciais, assim que recebem o primeiro pacote de um surto de voz. O modelo não pressupõe independência entre os surtos de voz, o percentual de superposição de surtos vai indicar com que frequência a reprodução de um surto de voz foi suspensa devido à chegada do primeiro pacote do surto de voz seguinte. A preferência é dada à informação de voz mais recente, ou seja, assim que o primeiro pacote de um surto de voz entra na fila de reprodução, qualquer reprodução em andamento é interrompida para que se possa dar início à reprodução do surto de voz recém chegado.

• **Duração de um Gap e Frequência de Gaps**

Devido à variação estatística do retardo de transmissão, pode ocorrer falta de pacotes depois de iniciado o processo de reprodução de um surto de voz. Isto ocasiona, como já mencionado, a introdução de trechos de silêncio que não existiam no sinal de voz original. Estes intervalos de silêncio ou *gaps* interferem na continuidade da voz reproduzida, podendo vir a comprometer sua inteligibilidade. As medidas descritas neste tópico vão reportar a média da duração observada para um *gap* e a frequência com que estes aparecem no surto de voz.

• **Varição das Pausas**

Também são levantados os percentuais de alteração observados na duração das pausas. No caso específico de verificar a deterioração do sinal de voz devido a alterações na duração das pausas, importa apenas o percentual e não a natureza da alteração, ou seja, saber se a mesma aumentou ou diminuiu. Por isso, a medida escolhida foi a média da variação das pausas levantadas durante a simulação. Outra medida importante e relacionada com a anterior diz respeito à maior variação observada na duração das pausas, permitindo avaliar se o percentual médio está muito distante do máximo observado. Com o emprego do *hangover* de 200ms, espera-se que as alterações na duração das pausas esteja abaixo dos $\pm 50\%$, lembrando que uma pausa dura em média 600ms.

4 Resultados obtidos

Para uma análise comparativa dos algoritmos de compensação é necessário assegurar que cada um esteja operando nas condições mais favoráveis. Para satisfazer este critério, a simulação se dividiu em duas etapas. Na primeira etapa foi executada uma simulação preliminar, cuja finalidade foi ajustar, para cada algoritmo, o valor do retardo inicial do primeiro pacote de cada surto de voz D , conforme as características da rede, a fim de garantir seu melhor desempenho sobre os critérios de retardo absoluto e introdução de *gaps* na reprodução.

Os resultados destas simulações preliminares mostraram que, para os *Algoritmos 1, 2 e 3*, o maior percentual de reproduções dentro dos padrões de retardo absoluto e introdução de *gaps* foi alcançado escolhendo um retardo inicial tal que $D < a_{m\acute{a}x} - t_{pac} - t_{rede}$, onde, de acordo com as hipóteses assumidas para o modelo do sistema, t_{pac} é constante (tamanho dos pacotes fixos), $a_{m\acute{a}x}$ é constante (40ms). O retardo médio da transmissão t_{rede} é constante para cada função de distribuição que modela a rede, sendo, na prática, o único componente difícil de se determinar.

Para o *Algoritmo 4* observou-se que:

- Caso $(t_{pac} + t_{rede}) < (a_{m\acute{a}x}/2)$, então o valor do retardo inicial deve se situar próximo de $t_{pac}/2$.
- Caso $(t_{pac} + t_{rede}) > (a_{m\acute{a}x}/2)$, então o valor do retardo inicial deve ser escolhido bem pequeno, ou seja, mais próximo de 0.

Os valores de retardo inicial determinados nesta simulação preliminar foram utilizados nas simulações executadas na segunda etapa, a fim de levantar as medidas descritas na seção anterior. Cada execução do modelo, tanto na primeira quanto na segunda etapa, avaliou o desempenho médio dos algoritmos após o tratamento de 700 a 800 surtos de voz. O intervalo de confiança escolhido, para o tempo de espera na fila de reprodução, foi de 95%. Todos os resultados apresentados se encontram dentro deste intervalo de confiança.

Os resultados apresentados a seguir foram obtidos durante a segunda etapa da simulação e se baseiam nas medidas descritas na seção anterior. A apresentação dos resultados está organizada pelo tipo de parâmetro avaliado. Todos os gráficos foram traçados em função do retardo médio da rede, que por sua vez está relacionado ao tamanho do pacote, para uma maximização do número de canais de voz, conforme ilustra a *Figura 5*[2]. Para cada parâmetro avaliado, o coeficiente de variação da função distribuição do retardo na rede recebeu os valores 1, 0.71, 0.5, 0.25 e 0.14. Os gráficos apresentados a seguir dizem respeito aos coeficientes de variação 1, 0.5 e 0.14.

• Reprodução dentro dos Limites de Retardo Absoluto

A diminuição do coeficiente de variação da função distribuição do retardo na rede vai causar um aumento no percentual da voz reproduzida dentro dos padrões aceitáveis de retardo absoluto. Para um coeficiente de variação menor ou igual a 0.25, todos os algoritmos passam a reproduzir praticamente todos os pacotes de voz recebidos, dentro do padrão aceitável. Para coeficientes de variação acima de 0.25, o percentual de pacotes reproduzidos dentro do padrão diferem para cada algoritmo, de acordo com o tamanho dos pacotes de voz. No entanto, pode-se observar uma certa uniformidade na forma das curvas levantadas para cada algoritmo. Nestas curvas pode-se identificar uma região de quebra dividindo-as em dois segmentos principais. No primeiro segmento é onde se verificam os mais altos percentuais de reproduções dentro do padrão de retardo absoluto. Os *Algoritmos 1, 2 e 3* apresentam perfil praticamente idêntico nesta parte da curva. No segundo segmento, ou seja, para pacotes maiores, os *Algoritmos 2 e 3* continuam muito próximos, mas distanciam-se do *Algoritmo 1*. A quebra que ocorre no *Algoritmo 4* é bem mais sutil do que nos demais algoritmos, mesmo para o maior coeficiente de variação. A região de quebra acontece praticamente para os mesmos valores de retardo na rede para todos os algoritmos. A região de quebra acentua-se mais para o coeficiente de variação $c=1$. Para pacotes acima de 20ms nota-se uma queda no percentual de pacotes reproduzidos dentro do padrão.

O *Algoritmo 4* apresenta-se o mais robusto, pois alterações no coeficiente de variação afetam pouco seu desempenho. Para exemplificar, o pior resultado é observado para pacotes de 30ms. Para o coeficiente de variação $c=1$, as reproduções aceitáveis atingem 89.13%. Diminuindo o coeficiente de variação em aproximadamente dez vezes, $c=0.14$, as reproduções dentro do padrão atingem 100%, representando uma melhora de 10.87%.

O pior desempenho referente ao retardo absoluto foi observado para os *Algoritmos 2 e 3*. Seu pior resultado também acontece para pacotes de 30ms. Para $c=1$, o percentual de reproduções aceitáveis é de 11.90%. Diminuindo o coeficiente de variação em aproximadamente dez vezes, $c=0.14$, as reproduções dentro do padrão atingem 100%, o que representa um aumento de 88.10%.

Por ser menos sensível ao coeficiente de variação da rede, o *Algoritmo 4* é o que apresenta um desempenho menos dependente da carga à que a rede está submetida momentaneamente. Sua robustez também se mostra em relação aos retardos médios e tamanhos de pacotes de voz, uma vez que seu desempenho não apresenta uma queda expressiva em nenhuma das curvas traçadas. Os *Algoritmos 2 e 3* são os mais sensíveis ao tamanho do pacote e retardo na rede, provavelmente devido ao fato de reproduzirem o surto integralmente, causando um acúmulo de retardos prejudicial aos pacotes finais de cada surto de voz. O *Algoritmo 1*, por sua vez, é menos sensível ao tamanho do

pacote e ao retardo na rede do que ao coeficiente de variação, embora apresente uma queda no percentual de reproduções dentro do padrão para pacotes a partir de 25ms.

- **Superposição de Surtos de Voz**

As Figuras 9, 10 e 11 mostram o percentual do surto de voz (sempre o trecho final) que foi perdido devido à chegada de um novo surto de voz, ou seja, devido à superposição de surtos de voz na reprodução. Com a diminuição do coeficiente de variação, a ocorrência de superposição de surtos de voz diminui. Os percentuais de perda, no pior dos casos, não ultrapassam 0.018%, o que demonstra que a interferência entre surtos de voz na reprodução é mínima. A hipótese de independência entre surtos de voz é, pois, bastante razoável.

- **Introdução de Gaps**

A duração dos *gaps* introduzidos pelo Algoritmo 1 depende unicamente do tamanho dos pacotes de voz, já que este algoritmo mantém a taxa de reprodução fiel à taxa de geração. Os demais algoritmos apresentam uma certa uniformidade com relação à duração média dos *gaps* introduzidos. Com o aumento do coeficiente de variação da rede, a duração dos *gaps* também aumenta, o que pode ser observado nas Figuras 12, 13 e 14. Como as curvas levantadas baseiam-se em médias, não é possível afirmar que os algoritmos não introduzem *gaps* acima de 50ms em mais de 1% do surto de voz. Pode-se afirmar, intuitivamente, que quanto maiores os *gaps* introduzidos, menos natural fica a reprodução, isto é, talvez seja preferível introduzir mais *gaps* de curta duração a poucos de longa duração. Para complementar a análise, é apresentado nas Figuras 15, 16 e 17 o percentual de aumento do surto de voz devido à introdução de *gaps*. Como já mencionado, o Algoritmo 1 introduz *gaps* fixos, do tamanho dos pacotes de voz. Este algoritmo chegou a provocar um percentual de aumento do surto de voz em torno de 30%, para $c=1$ e pacotes de 30ms. Os Algoritmos 2 e 3 ocasionam pequenos aumentos no surto de voz, entretanto, seu desempenho com relação ao máximo retardo absoluto é muito ruim, conforme visto anteriormente. O Algoritmo 4 apresenta a vantagem de espalhar *gaps* de curta duração ao longo do surto de voz. Pode-se notar que os algoritmos fizeram opções distintas para controlar a descontinuidade da voz, baseados no retardo absoluto e/ou na duração média dos *gaps*. Observa-se também, que o aumento na duração do surto de voz causado pelo Algoritmo 4 se mantém constante e independente do tamanho de pacote utilizado. Esta independência com relação ao tamanho do pacote de voz, que não se verifica para o Algoritmo 1, é um aspecto que torna a utilização do Algoritmo 4 bastante vantajosa. Com a diminuição do coeficiente de variação da rede, todos os algoritmos passam a introduzir, em média, menos *gaps*, porém a duração média destes pouco varia.

- **Varição das Pausas**

Conforme mencionado anteriormente, a variação das pausas não deve ultrapassar 50% para que a inteligibilidade da voz não seja afetada. Para avaliar este desempenho foram traçados os gráficos mostrados nas Figuras 18 a 23. As alterações médias verificadas na duração das pausas não chega a 5%, no pior dos casos. Quanto menor o coeficiente de variação da rede, mais próximos estão as alterações médias originadas pelos quatro algoritmos. Para o coeficiente de variação $c=1$, observa-se que os Algoritmos 1 e 4 mantêm uma curva bastante semelhante; o mesmo ocorre em relação aos Algoritmos 2 e 3. Para assegurar que nenhuma pausa teve sua duração alterada em mais do que 50% do seu valor, foram levantados também os valores máximos de suas alterações. O percentual mais alto de alteração não ultrapassou 25%, ocorrendo para o coeficiente de variação $c=1$.

5 Conclusão

A análise de desempenho apresentada no presente trabalho teve a preocupação de abordar o problema da compensação da variação estatística do retardo, levando em consideração o

limite máximo permitido para o retardo absoluto da voz, cujo controle é muito importante para viabilizar comunicações remotas interativas.

Pelos resultados levantados, o *Algoritmo 4* apresentou-se mais robusto tanto em relação à carga na rede quanto ao tamanho do pacote de voz. Outra grande vantagem deste algoritmo é distribuir melhor os *gaps*, que são de curta duração, ao longo do surto de voz. Os demais algoritmos mostram-se bem mais sensíveis à variações da carga na rede, podendo chegar facilmente a níveis muito baixos de reproduções dentro dos padrões de retardo absoluto.

Outra vantagem do *Algoritmo 4* diz respeito à sua simplicidade de implementação. Ele pode ser implementado utilizando um *buffer* circular com tamanho $(2D + t_{pac}) \cdot \text{taxa}$ de codificação. A referência [11] relata uma implementação do algoritmo, fornecendo também extensões ao mesmo no que diz respeito ao ajuste dinâmico do retardo inicial D .

Como o modelo foi implementado na forma modular, fica fácil substituir o módulo da rede por outro a fim de avaliar o comportamento dos algoritmos em outro ambiente de comunicação. Atualmente o desempenho dos algoritmos vem sendo avaliado pensando em um ambiente de redes de alta velocidade, utilizando o modo assíncrono de transmissão (ATM).

Referências Bibliográficas

- [1] SOARES, L.F.G. e MARTINS, S.L. "Um Algoritmo para Compensação da Variação Estatística do Retardo em Redes Comutadas por Pacotes". Anais do 8o. Simp. Bras. de Redes de Comp., 1991.
- [2] SUDA, T. e BRADLEY, T.T. "Packetized Voice/Data integrated Transmission on a Token Passing Ring Local Area Network". IEEE Transactions on Communications, vol.37, no.3, março de 1989.
- [3] LEE, H.H. e UN, C.K. "A Study of On-Off Characteristics of Conversational Speech". IEEE Transactions on Communications, vol.com-34, no.6, junho de 1986.
- [4] GOPAL, P.M. e BHARATHI KADABA. "A Simulation Study of Network Delay for Packetized Voice". IEEE (CH2298), 27.2.1-27.2.7, 1986.
- [5] GRUBER, J.H. e STRAWCZYNSKI, I. "Subjective Effects of Variable Delay and Speech Loss in Dynamically Managed Voice Systems". IEEE Trans. Commun., vol.com-33, agosto de 1985.
- [6] ADAMS, C. e ADES, S. "Voice Experiments in the UNIVERSE Project". Proc. Int. Conf. Commun., 29.4.1-29.4.9, 1985.
- [7] GOPAL, P.M., WONG, J.W. e MAJITHIA, J.C. "Analysis of Playout Strategies for Voice Transmission Using Packet Switching Techniques". Performance Evaluation, no.4, fevereiro de 1984.
- [8] SUDA, T., YEMINI, Y., MIYAHARA, H. e HASEGAWA, T. "Performance Evaluation of a Packetized Voice System Simulation Study". Proc. Int. Conf. Commun., C5.2.1-C5.2.5, Boston 1983.
- [9] GRUBER, J.G. "A Comparison of Measured and Calculated Speech Temporal Parameters Relevant to Speech Activity Detection". IEEE Transactions on Communications, vol.com-30, no.4, abril de 1982.
- [10] WEINSTEIN, C.J. e HOFSTETTER, F.M. "The Tradeoff Between Delay and TASI Advantage in a Packetized Speech Multiplexer". IEEE Transactions on Communications, vol.com-27, no.11, novembro de 1979.
- [11] FARIA, A.L.A. "Floor Control em Teleconferência por detecção de silêncio". Dissertação de Mestrado do Depto. de Engenharia Elétrica da PUC-Rio, a publicar.

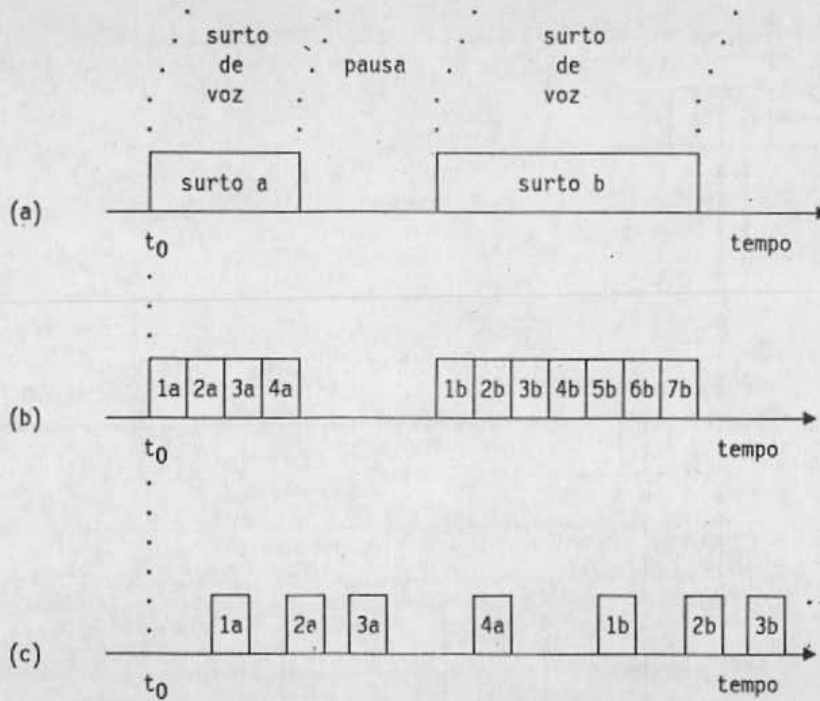


Figura 1 - Empacotamento de surtos de voz para transmissão via rede.(a) segmento de voz a ser transmitido;(b) empacotamento dos surtos de voz;(c) efeito da variação estocástica do retardo na chegada dos pacotes no destino.

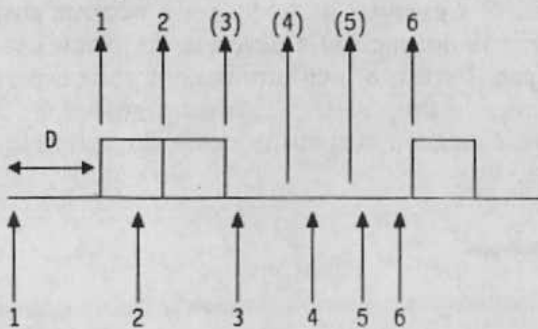


Figura 2 - Exemplo de funcionamento do Algoritmo 1, onde as setas abaixo indicam a chegada do pacote (identificado por um número) ao destino e a seta acima indica o início da sua reprodução. O número de um pacote entre parênteses indica a falta do mesmo para reprodução.

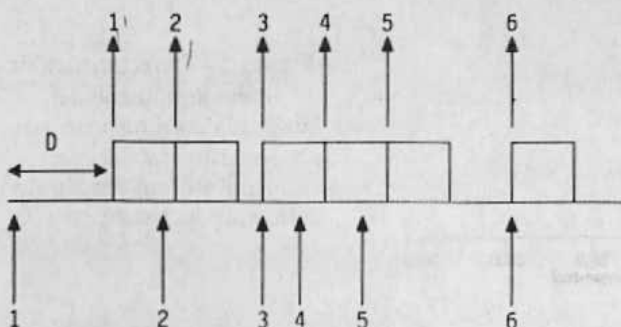


Figura 3 - Exemplo de funcionamento dos Algoritmos 2 e 3, onde todos os pacotes são reproduzidos.

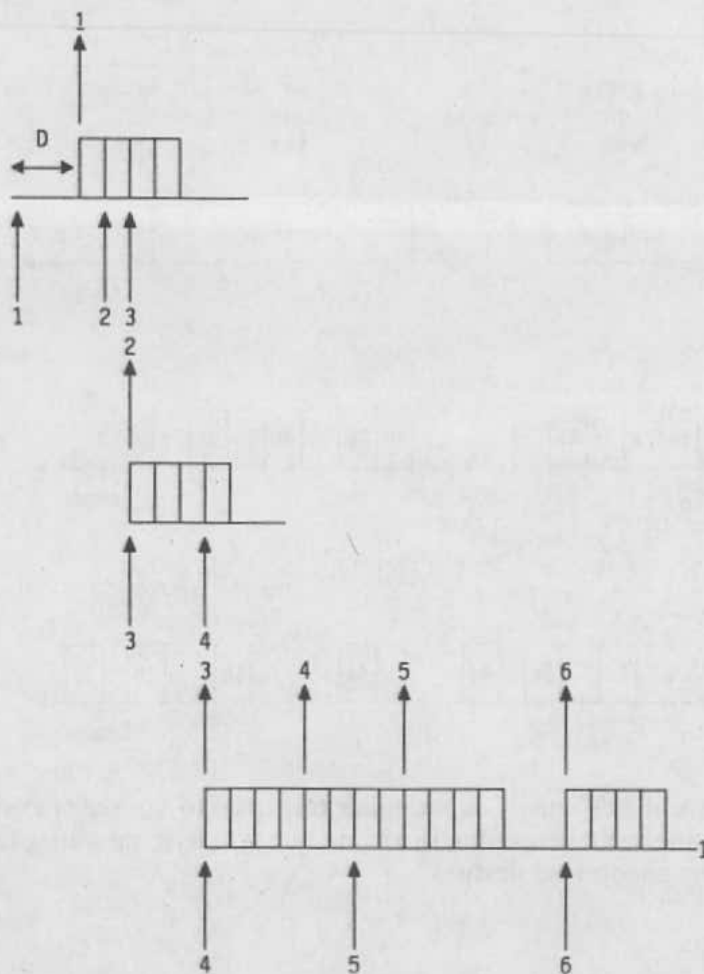


Figura 4 - Exemplo de funcionamento do *Algoritmo 4*, onde pode ocorrer suspensão da reprodução em andamento (parte hachuriada do pacote) e descarte de pacotes já na fila de reprodução. No exemplo acima, $D = 2t_{pac}$. Portanto, nenhum pacote pode esperar mais do que t_{pac} na fila pela reprodução. A chegada do *pacote 3* causou a suspensão da metade final do *pacote 1*; já a chegada do *pacote 4* causa a suspensão de $\frac{1}{4}$ do *pacote 2*; os demais pacotes são reproduzidos completamente.

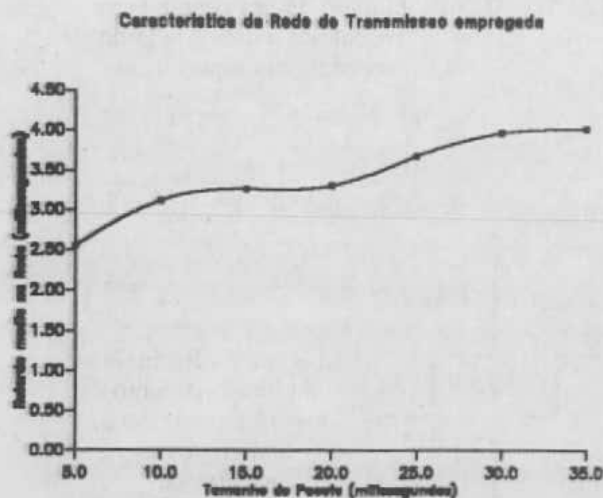


Figura 5 - Característica da rede de transmissão utilizada com respeito aos retardos médios de transmissão em função do tamanho do pacote de voz.

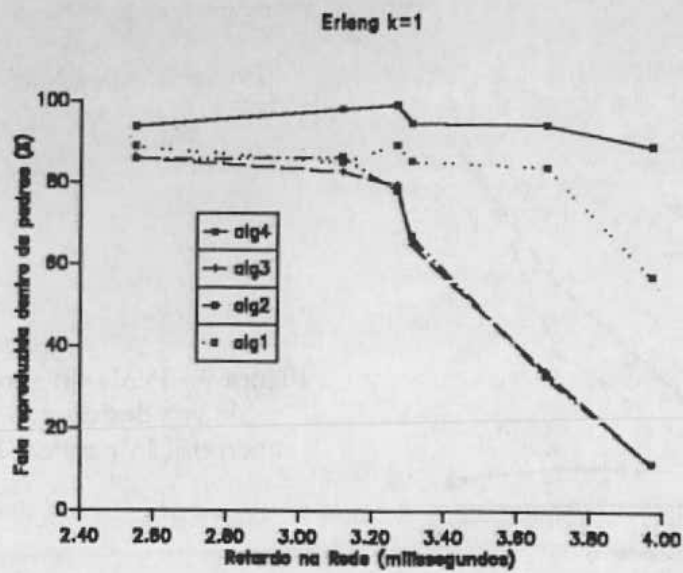


Figura 6 - Reprodução dentro dos limites de máximo retardo para $c = 1$

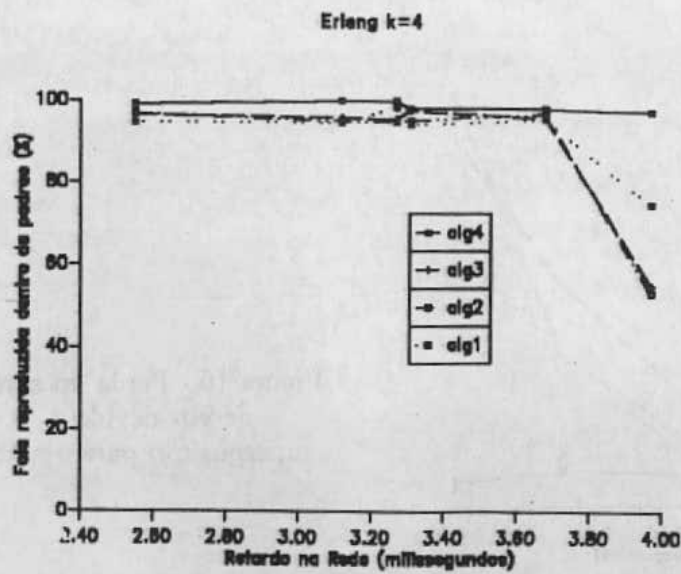


Figura 7 - Reprodução dentro dos limites de máximo retardo para $c = 0.5$

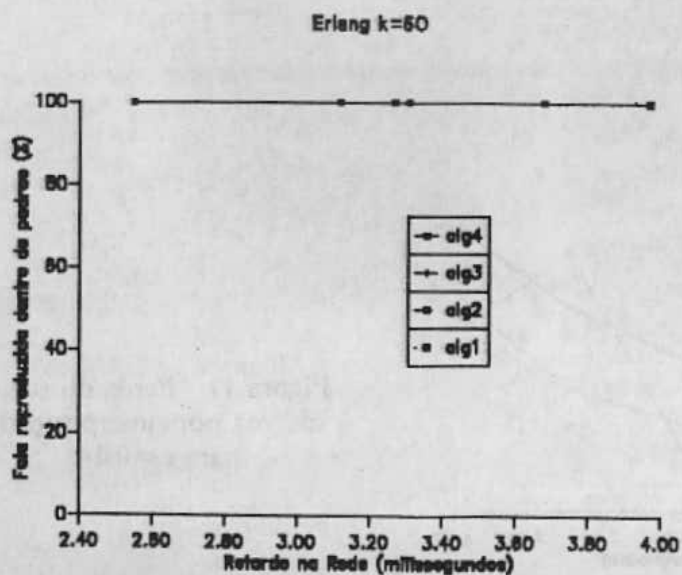


Figura 8 - Reprodução dentro dos limites de máximo retardo para $c = 0.14$

Erlang k=1

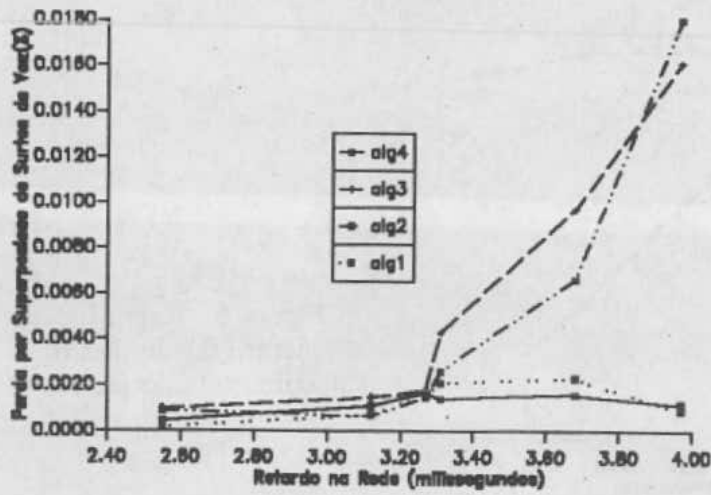


Figura 9 - Perda do surto de voz devido à superposição para $c = 1$

Erlang k=4

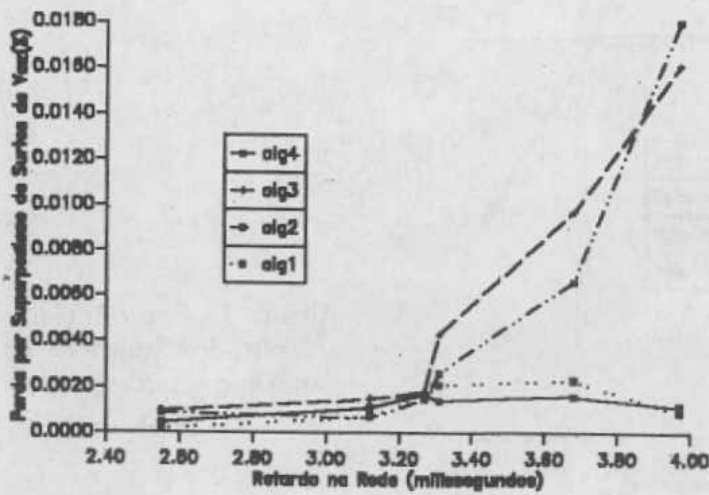


Figura 10 - Perda do surto de voz devido à superposição para $c = 0.5$

Erlang k=60

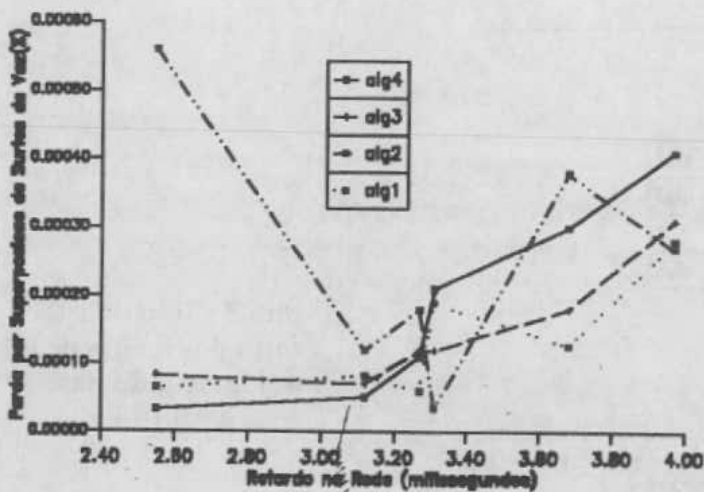


Figura 11 - Perda do surto de voz por superposição para $c = 0.14$

Erlang k=1

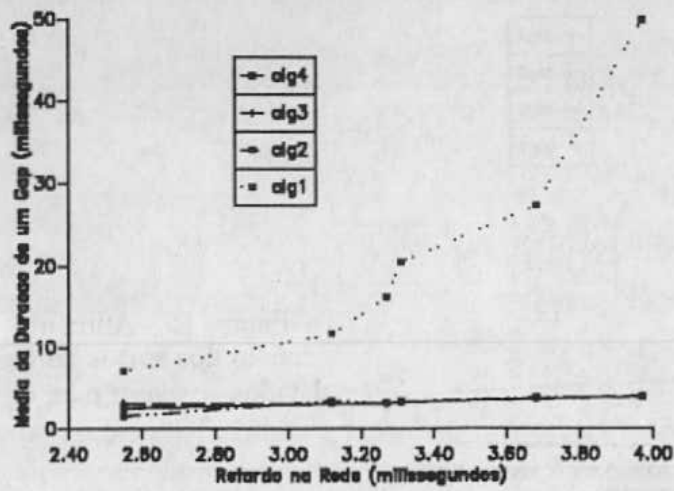


Figura 12 - Duração média dos gaps introduzidos para $c = 1$

Erlang k=4

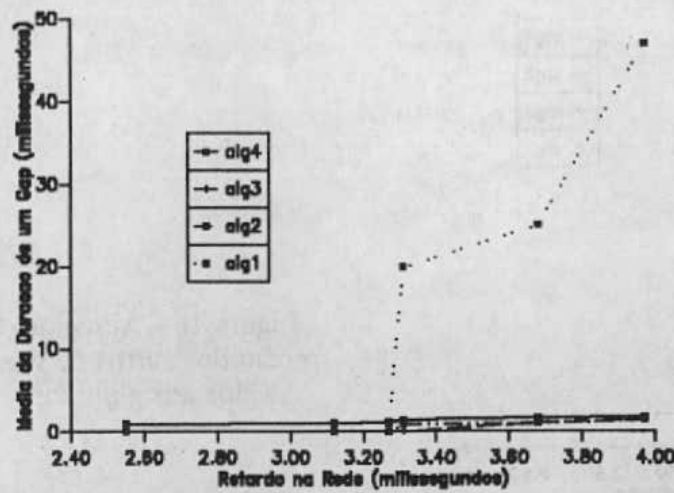


Figura 13 - Duração média dos gaps introduzidos para $c = 0.5$

Erlang k=60

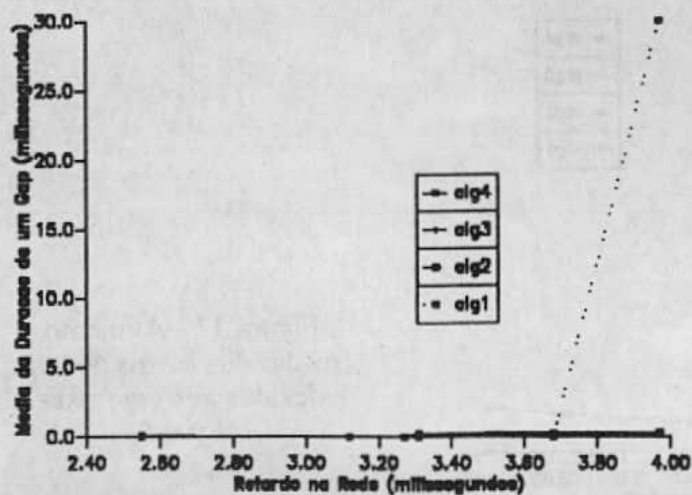


Figura 14 - Duração média dos gaps introduzidos para $c = 0.14$

Erlang k=1

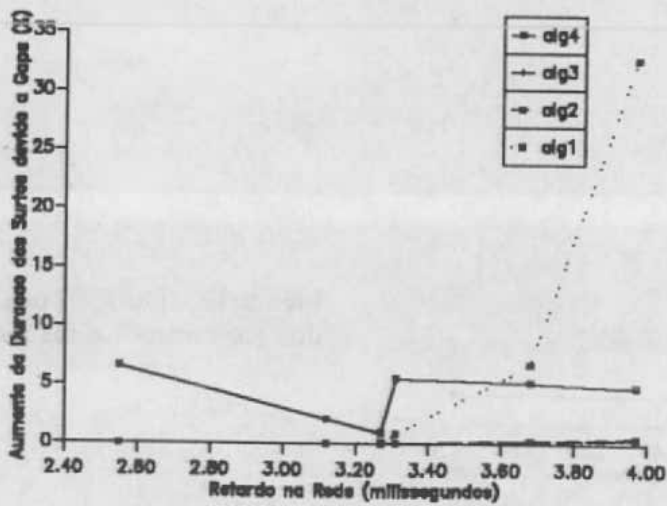


Figura 15 - Aumento médio dos surtos de voz devidos aos *gaps* para $c = 1$

Erlang k=4

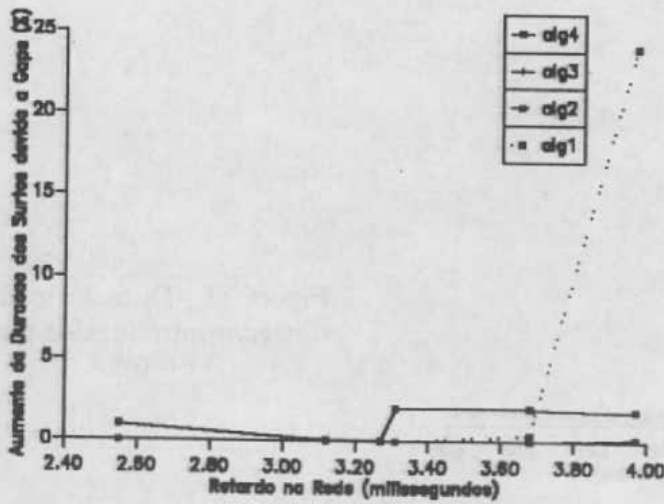


Figura 16 - Aumento médio dos surtos de voz devidos aos *gaps* para $c = 0.5$

Erlang k=50

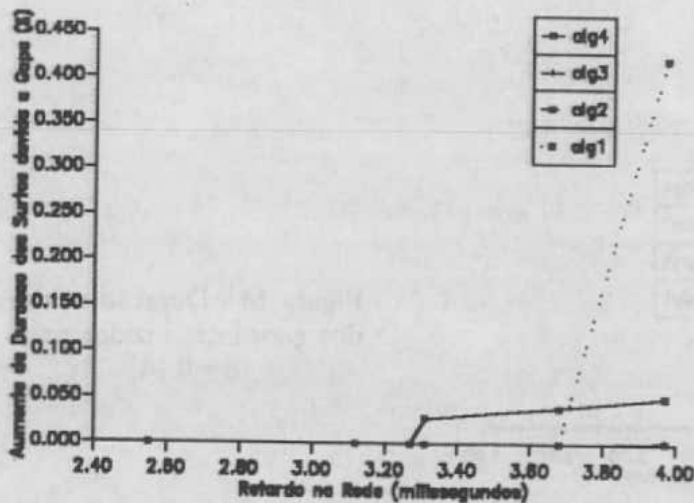


Figura 17 - Aumento médio dos surtos de voz devidos aos *gaps* para $c = 0.5$

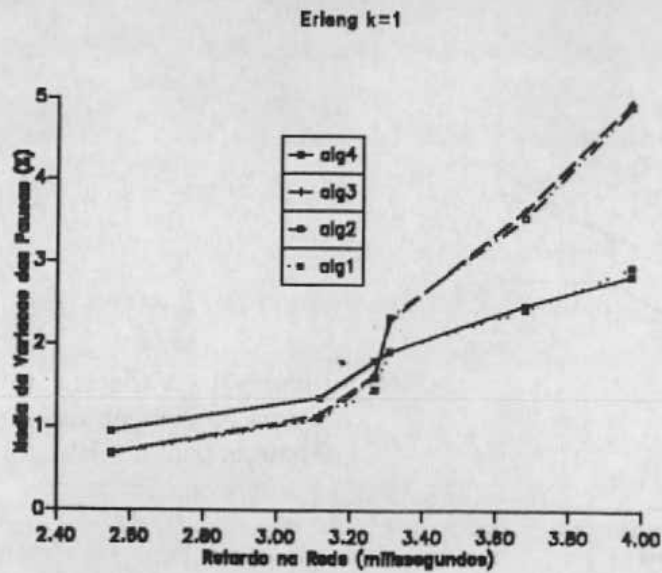


Figura 18 - Variação média da duração das pausas para $c = 1$

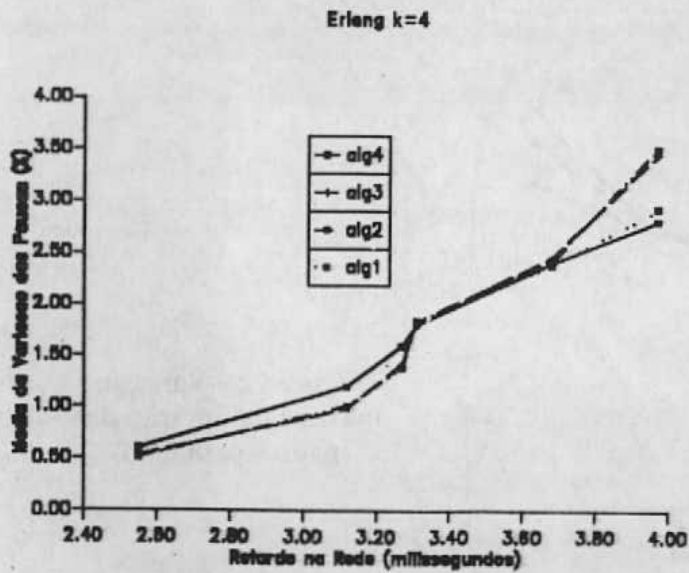


Figura 19 - Variação média da duração das pausas para $c = 0.5$

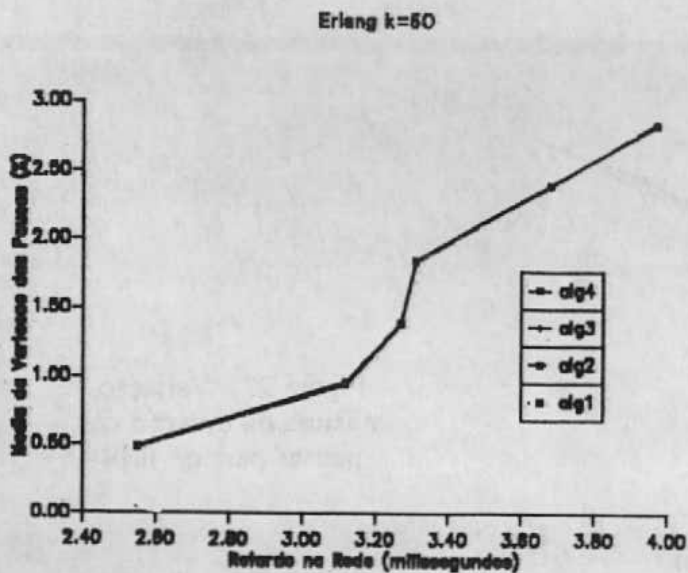


Figura 20 - Variação média da duração das pausas para $c = 0.14$

Erlang k=1

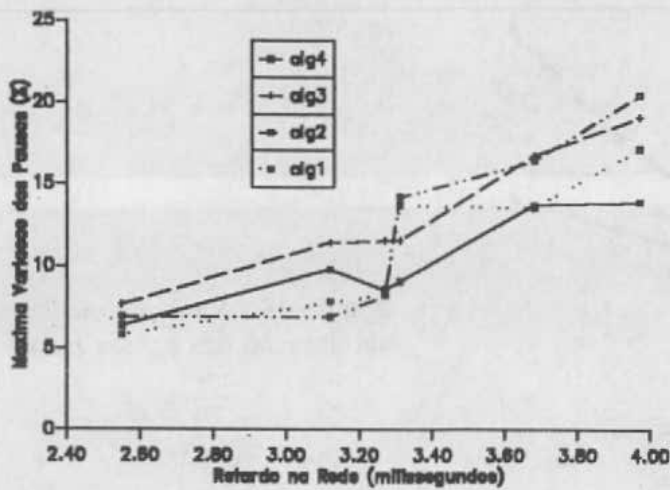


Figura 21 - Variação máxima da duração das pausas para $c = 1$

Erlang k=4

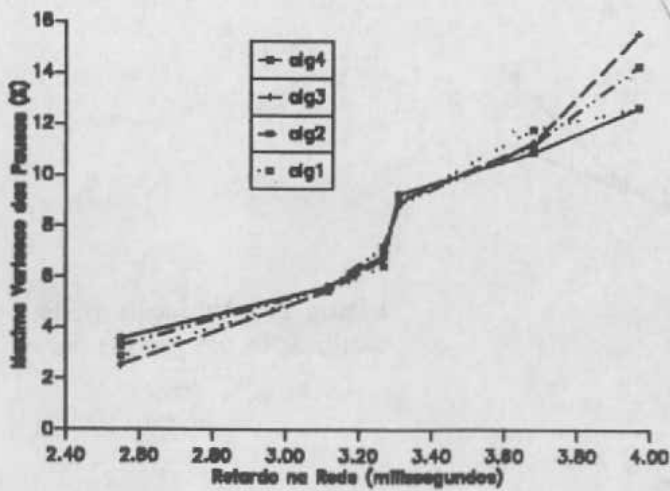


Figura 22 - Variação máxima da duração das pausas para $c = 1$

Erlang k=80

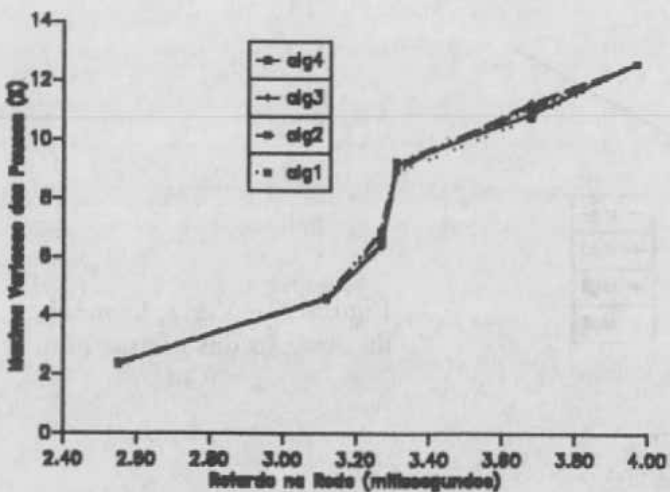


Figura 23 - Variação máxima da duração das pausas para $c = 0.14$