

O Protocolo de Vídeo sob Demanda GEBB Sujeito a Restrições de Banda Passante

Rogério M. Zafalão¹ e Nelson L. S. da Fonseca²

¹Instituto de Pesquisas Eldorado
Rodovia SP-340 Campinas-Mogi-Mirim, Km 118,5
13086-902 Campinas, SP

²Instituto de Computação - Universidade Estadual de Campinas
Caixa Postal 6176 - 13083-970 Campinas, SP

rogerio.zafalao@eldorado.org.br, nfonseca@ic.unicamp.br

Abstract. *Periodic Broadcast Protocols require a constant amount of bandwidth regardless the number of viewers. In this paper, a new protocol based on Greedy Equal Bandwidth Broadcasting Protocol for users under bandwidth constrain is proposed. Moreover, a comparison between optimally-structured and non-optimally structured periodic broadcasting protocols with limited client bandwidth is pursued.*

Resumo. *Protocolos de difusão periódica requerem largura de banda constante, sendo indicados para a transmissão dos vídeos mais frequentemente acessados. No entanto, estes protocolos assumem que não existe limitação de banda passante. Neste trabalho, é introduzido um novo protocolo otimamente estruturado com limitação de banda no cliente, o GEBB-LBU. Também são realizadas comparações entre protocolos de difusão periódica que permitam sua utilização por clientes com largura de banda limitada.*

1. Introdução

Protocolos de difusão periódica são vistos como uma opção escalável em aplicações de Vídeo sob Demanda (VoD), uma vez que requerem largura de banda constante independentemente do número de usuários. Nestes protocolos, o servidor divide um vídeo em vários segmentos, e os transmite em um conjunto de canais dedicados à transmissão destes segmentos. Cada cliente recebe dados de vários canais ao mesmo tempo para armazenar/exibir os quadros posteriores. Todos os protocolos que utilizam difusão periódica possuem uma organização similar: dividem cada vídeo em n segmentos, transmitidos simultaneamente via difusão em diferentes canais lógicos.

A largura de banda é um dos principais recursos que devem ser levados em consideração para a implantação de um sistema de VoD. Quanto menor for a demanda de banda passante, mais usuários poderão ser agregados ao sistema. Entretanto, a maioria dos protocolos não foi concebida para suportar clientes com limitações de largura de banda no receptor. Os protocolos otimamente estruturados, protocolos que minimizam a demanda de banda do servidor, requerem que todos os canais sejam recebidos ao mesmo tempo, podendo exigir mais banda passante do que a disponível pelo cliente.

A contribuição do presente trabalho consiste na extensão do protocolo otimamente estruturado GEBB (*Greedy Equal Bandwidth Broadcasting Protocol*, ou Protocolo de

Difusão Guloso com Canais de Mesma Largura de Banda), permitindo usuários com limitação de banda e demandando o mínimo possível de largura de banda do servidor. Além disso, analisa-se os compromissos entre os parâmetros e recursos envolvidos nos protocolos de difusão com limitação de banda no cliente.

Em [9], estudou-se o protocolo poliharmônico sujeito a restrições de banda passante. O presente trabalho difere de [9] por adotar uma nova função objetivo que minimiza a quantidade de banda passante utilizada pelo servidor ao invés do mapa de banda passante. Experimentos realizados e não reportados no presente artigo indicam que a função objetivo adotada aqui consegue um melhor desempenho do que a função objetivo adotada em [9]. Em [8] estendeu-se o protocolo de difusão poliharmônica com a mesma função objetivo usada neste artigo, para a realização da comparação entre os protocolos de difusão periódica com limitação de largura de banda do usuário.

O presente trabalho está organizado da seguinte forma: a Seção 2. aborda as primeiras adaptações feitas a protocolos não otimamente estruturados. Na Seção 3., introduz-se o conceito de protocolo otimamente estruturado. O protocolo GEBB e sua extensão são apresentados na Seção 4.. Na Seção 5. há uma comparação entre os protocolos otimamente e não otimamente estruturados. Considerações finais são apresentadas na Seção 6..

2. Protocolos Não Otimamente Estruturados

Em [6], discute-se um novo método para adaptação dos protocolos com o objetivo de transpor a limitação de largura de banda do cliente, adaptando os protocolos de Difusão Rápida (PDR) e o Novo Protocolo de Difusão em Pagode (NPDPa). Enquanto um protocolo convencional utiliza C canais para a transmissão dos segmentos, este método consiste na utilização máxima de k canais simultâneos. A recepção dos $C - k$ canais restantes é adiada pelos clientes e cada canal inicia imediatamente após a liberação de um canal já utilizado. Desta forma, o i -ésimo canal (onde $i \geq k + 1$) inicia sua recepção logo após o término da recepção do $(i - k)$ -ésimo canal. O método é válido para todos os protocolos que utilizem canais com mesma largura de banda e segmentos de mesmo tamanho.

Entretanto, em qualquer protocolo com limitação de largura de banda para recepção, reduzir esta largura de banda acarreta em um aumento da latência de exibição. Nos casos dos protocolos PDR e NPDPa, isto fica mais evidente devido à redução do número total de segmentos alocados nos canais, cuja consequência imediata é justamente o aumento da latência. O PDR, por exemplo, utiliza 255 segmentos com oito canais, enquanto suas versões com limitação de banda do usuário utilizam 176 e 223 segmentos para $k = 3$ e $k = 4$, respectivamente.

3. Protocolos Otimamente Estruturados

Em [1], foram definidas três regras para que protocolos de difusão periódica possam obter uma utilização mais racional da largura de banda. A primeira regra estabelece a transmissão única de um segmento durante o ciclo de recepção do usuário, isto é, desde o momento da requisição até o início da exibição deste segmento. A segunda regra sugere que nenhuma porção desnecessária do vídeo deva ser transmitida durante um ciclo. Estes protocolos aumentam a frequência de transmissão de segmentos para garantir a entrega dos mesmos, transmitindo desta forma dados redundantes. A última regra afirma a

necessidade da entrega de todos os segmentos a tempo, sejam estes armazenados anteriormente ou exibidos simultaneamente à recepção. Esta última regra é vital para o correto funcionamento de qualquer protocolo de difusão periódica.

Os protocolos que seguem estas regras são candidatos a minimizar o desperdício de banda passante e por este motivo são denominados otimamente estruturados (*optimally-structured*). Dentre eles, os únicos protocolos considerados otimamente estruturados são o PDPH e o GEBB.

3.1. Protocolos otimamente estruturados com limitação de banda do usuário

Um dos principais objetivos de qualquer protocolo de difusão periódica é a obtenção da menor latência de exibição possível. Como a latência está diretamente associada ao tamanho do primeiro segmento, torna-se possível reduzi-la através de uma maior segmentação do vídeo, reduzindo o prefixo e, por consequência a latência. Entretanto, o aumento no número de segmentos acaba por demandar uma largura de banda extra para acomodá-los.

Para permitir uma menor latência de exibição no caso da limitação de banda no cliente, utiliza-se a seguinte estratégia: introduz-se atrasos no calendário de recepção de alguns segmentos, a fim de se obter uma solução com a latência desejada e de tal forma que em qualquer instante da recepção o cliente não precise receber mais do que sua capacidade de banda passante.

Um grupo de canais que têm sua recepção de segmentos iniciada simultaneamente é definido como sendo um conjunto de canais. No mapa da Figura 1, por exemplo, vê-se dois conjuntos de canais: c_1 , que consiste nos cinco primeiros canais recebidos imediatamente após a requisição de exibição do cliente, e c_2 , compreendendo os três últimos canais, cuja recepção só se inicia no instante do término da recepção do segundo segmento. A partir daí ocorre um período de transição, onde o cliente passa a receber segmentos de canais de conjuntos adjacentes, até o término da recepção de segmentos do conjunto anterior.

O problema que decorre desta abordagem é que, dentre a vasta gama de possibilidades de configuração dos parâmetros modificadores do calendário (largura de banda dos canais, tempo de atraso dos conjuntos de canais e o número de segmentos de cada conjunto), deve-se definir qual configuração demanda a menor banda passante do servidor. Para determinar a configuração ótima destes parâmetros, utiliza-se problemas de otimização.

A escolha de protocolos otimamente estruturados foi feita devido ao fato destes possuírem uma utilização mais racional da banda disponível. A idéia principal é manter a utilização racional da banda passante do servidor, apesar da restrição de largura de banda do usuário.

3.2. Protocolo de Difusão Poliharmônica

O protocolo de Difusão Poliharmônica (*Polyharmonic Broadcasting*)[4] foi o primeiro protocolo otimamente estruturado criado, e consiste na divisão do vídeo em n segmentos de mesmo tamanho, transmitidos em canais exclusivos com largura de banda decrescente, de acordo com a série harmônica.

Este protocolo possui um parâmetro adicional m , relacionado com a largura de banda do primeiro segmento, cujo valor é igual à taxa de consumo do vídeo b multiplicada pelo m -ésimo termo da série harmônica. Em outras palavras, os canais seguem a série: $\frac{b}{m}, \frac{b}{m+1}, \frac{b}{m+2}, \dots, \frac{b}{m+n-1}$.

A versão com suporte à limitação de largura de banda do cliente [8] utiliza a noção de atrasos a conjuntos de canais, e é modelada através de um problema de otimização que define a configuração dos parâmetros m e n ótimos de cada conjunto de canais para que a largura de banda do servidor seja minimizada.

4. O Protocolo GEBB

O protocolo GEBB (*Greedy Equal Bandwidth Broadcasting Protocol*) parte do princípio que a estrutura ótima de um protocolo de difusão decorre da mesma característica gulosa de recepção que o protocolo de Difusão Poliharmônica possui. Neste protocolo, o usuário deve iniciar a recepção de todos os canais (simultaneamente), a partir do momento de sua requisição. Assim como o PDPH, o GEBB também precisa receber o segmento completo antes do início de sua exibição.

A obtenção da segmentação e da largura de banda de cada um dos canais é feita através de um problema de otimização visando obter uma segmentação ótima (S_i) e a largura de banda de cada canal (b_i) de tal forma que a demanda de largura de banda seja minimizada. O resultado é obtido analiticamente e tem como solução:

$$b_i = b^* = \sqrt[n]{\frac{S}{w} + 1} - 1 \quad (1)$$

$$S_i = wb^*(1 + b^*)^{i-1} \quad (2)$$

4.1. O Protocolo GEBB-LBU

A estratégia do GEBB sujeito a limitação de banda passante no usuário consiste na divisão dos canais em conjuntos, onde um conjunto de canais pode ter um atraso no calendário de recepção do cliente em relação ao conjunto anterior, para que o cliente que possua limitação de largura de banda possa receber uma quantidade máxima possível de dados, como ilustra a Figura 1.

A forma como foi estruturado o problema de otimização pode trazer benefícios em sua implantação, tratando cada conjunto de canais como se fosse um problema GEBB isolado. A Figura 1 mostra o mapa de difusão do GEBB-LBU com $d = 2$, destacando-se em sombreado o primeiro conjunto de canais (que compreende os canais um a quatro), e em hachurado o segundo conjunto de canais (cinco a sete), permitindo a utilização das equações (1) e (2) do GEBB sem restrições, a fim de obter os valores de S_i e b_i de cada conjunto de canais. Torna-se necessário somente fazer adaptações apropriadas de variáveis. Por exemplo, no primeiro conjunto de canais o tamanho do vídeo corresponde apenas aos segmentos entre S_1 a S_4 . Já no segundo conjunto deve-se também ajustar o tempo de espera w para um tempo equivalente à duração dos segmentos S_3 e S_4 .

No GEBB-LBU são utilizados dois tipos de variáveis. O primeiro tipo engloba as que se aplicam a um único canal ou segmento, como no caso das variáveis S_i e b_i

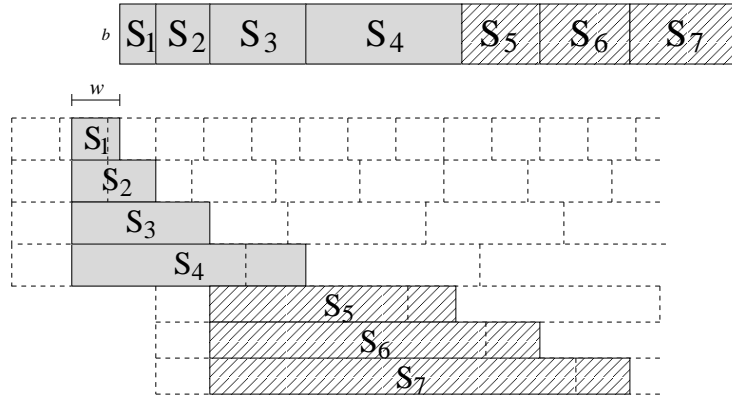


Figura 1. Um problema do GEBB-LBU tratado como d problemas do GEBB

($i = 1, \dots, n$), que indicam o tamanho do i -ésimo segmento e a largura de banda alocada para sua transmissão, respectivamente. Já o segundo tipo compreende as variáveis que se aplicam aos conjuntos de canais. É o caso das variáveis $s^{(c)}$, $n^{(c)}$ e $a^{(c)}$, onde c estende-se de 1 a d , que aparecem com uma notação diferenciada: o índice aparece sobrescrito e entre parênteses, para facilitar a diferenciação da “natureza” das variáveis. $s^{(c)}$ e $n^{(c)}$ dizem respeito, respectivamente, à soma do tamanho dos segmentos que compõem o c -ésimo conjunto de canais, e ao número de segmentos acumulados até o referido conjunto de canais. A definição da variável $a^{(c)}$ é relacionada ao momento de término de recepção do último segmento anterior ao conjunto de canais c . O seu significado corresponde ao número de segmentos com atraso de recepção em relação ao primeiro conjunto de canais.

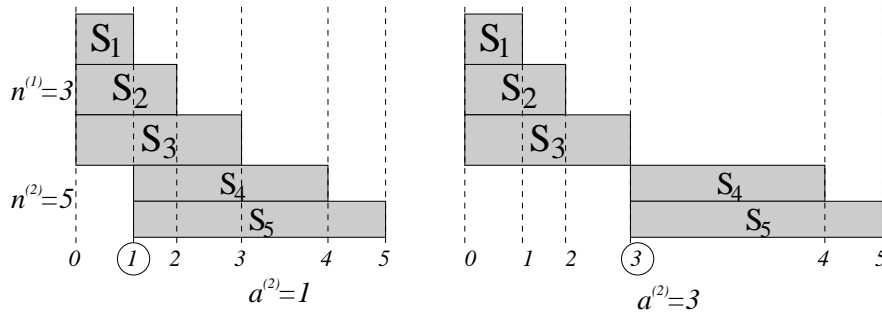


Figura 2. Dois possíveis valores para a variável $a^{(2)}$ ($n^{(1)} = 3$, $n^{(2)} = 5$)

Na Figura 2, são ilustrados dois valores possíveis de $a^{(2)}$ para uma configuração de cinco canais no total, sendo $n^{(1)} = 3$ e $n^{(2)} = 5$ (há dois canais no segundo conjunto, que se acumulam aos três primeiros). Cada variável $a^{(c)}$, em outras palavras, deve estar entre zero (quando então o protocolo funciona como se o conjunto c fundisse com o conjunto anterior, ocorrendo o mesmo sempre que $a^{(c)} = a^{(c-1)}$) e n_i . Por convenção, assume-se que $a^{(1)}$ deva sempre ser igual a zero, uma vez que não existem conjuntos de canais anteriores ao primeiro.

O problema de otimização que define o GEBB-LBU é dado a seguir:

$$\min \sum_{i=1}^n b_i \quad (3)$$

sujeito às restrições:

$$b_i \left(w + \sum_{j=1}^{i-1} S_j \right) = S_i \quad i = 1, 2, \dots, n^{(1)} \quad (4)$$

$$b_i \left(\sum_{j=a^{(c)}}^{i-1} S_j \right) = S_i \quad i = n^{(c-1)} + 1, n^{(c-1)} + 2, \dots, n^{(c)},$$

$$\forall c \in 2 \dots d \quad (5)$$

$$\sum_{j=a^{(c)}+1}^{n^{(c)}} b_j \leq k \quad c = 1, \dots, d \quad (6)$$

$$\begin{array}{ccccccc} b_1 & = & b_2 & = & \dots & = & b_{n^{(1)}} \\ b_{n^{(1)}+1} & = & b_{n^{(1)}+2} & = & \dots & = & b_{n^{(2)}} \\ \vdots & & \vdots & & \dots & & \vdots \\ b_{n^{(d-1)}+1} & = & b_{n^{(d-1)}+2} & = & \dots & = & b_{n^{(d)}} \end{array} \quad (7)$$

$$\sum_{i=1}^{n^{(1)}} S_i = s^{(1)} \quad (8)$$

$$\sum_{i=n^{(c-1)}+1}^{n^{(c)}} S_i = s^{(c)}, \quad c = 2, \dots, d \quad (9)$$

$$\sum_{c=1}^d s^{(c)} = S \quad (10)$$

$$0 = a^{(1)} \leq a^{(2)} \leq \dots \leq a^{(d)} \quad (11)$$

$$a^{(i)} \leq n^{(i-1)} \quad i = 2, \dots, n \quad (12)$$

$$\begin{array}{l} w > 0, \quad 1 < n, \quad 1 < d \\ 0 < S_i < S, \quad i = 1, 2, \dots, n \\ 0 < b_i, \quad i = 1, 2, \dots, n \end{array} \quad (13)$$

A adição de c conjuntos de canais ao conjunto original e as variáveis criadas correspondentes a cada conjunto geram várias restrições para relacionar estas variáveis com as variáveis que quantificam os canais e segmentos (respectivamente, b_i e S_i).

Para os canais do primeiro conjunto, (4) é utilizada para fazer a correspondência entre a recepção e a exibição de um segmento, onde o momento de término da recepção de um segmento é o mesmo do início de sua exibição. Para os conjuntos restantes, utiliza-se (5).

A restrição seguinte (6) compreende na verdade d restrições, uma para cada conjunto de canais. Cada uma delas limita a largura de banda máxima no instante no qual um conjunto de canais tem sua recepção iniciada, quando o requisito de banda passante é o máximo possível.

As equações contidas em (7) refletem a relação existente entre as variáveis individuais e as variáveis de conjuntos de canais. Especificamente, a largura de banda é a mesma para todos os canais pertencentes a um conjunto. O mesmo ocorre com as restrições no tamanho dos segmentos, que seguem duas regras. A primeira delas diz que a soma dos tamanhos dos segmentos de um conjunto equivale ao tamanho do conjunto, dadas por (8) e (9). Na segunda, o total do tamanho dos conjuntos deve ser igual ao tamanho total do vídeo.

Quando $a^{(c)} < a^{c-1}$, (12) faz com que o conjunto c tenha um atraso em relação ao conjunto anterior. No caso de $a^{(c)} = a^{(c-1)}$, o conjunto é aglutinado com o anterior. Quando $a^{(1)} = a^{(2)} = \dots = a^{(d)} = 0$, a formulação é idêntica à do GEBB-LBU com somente um conjunto de canais.

Um aspecto a se salientar é que a obtenção dos parâmetros ótimos só é necessária na primeira vez que um vídeo entra no sistema ou ainda quando o mesmo tem necessidades de restrição modificadas. Ao receber estes parâmetros, o cliente pode calcular rapidamente a duração de cada segmento através de (2) e a recepção é feita da mesma forma que qualquer protocolo de difusão periódica.

4.2. Uma avaliação da efetividade do GEBB-LBU

Exemplos numéricos foram derivados para avaliar a efetividade do protocolo GEBB-LBU. Foram realizados estudos sobre o compromisso entre a largura de banda do cliente, o tempo de espera máximo ao qual o cliente é submetido e a largura de banda do servidor.

Sendo os problemas de otimização GEBB-LBU de natureza inteira e não-linear, utilizou-se algoritmos genéticos (AG) [7, 3] como heurística para sua resolução.

A estrutura de um cromossomo para uma solução GEBB-LBU é composta pelos parâmetros que definem cada conjunto de canais, consistidos das variáveis $a^{(c)}$, $n^{(c)}$ e $s^{(c)}$, para cada conjunto de canais c_1, c_2, \dots, c_d . Há também um gene k_{ini} , que define a porção de largura de banda inicial a ser utilizada pelo primeiro conjunto de canais.

A população inicial é gerada de forma que cada gene de um indivíduo receba um valor aleatório entre sua faixa de valores possíveis. Depois da geração dos genes, a factibilidade da solução é verificada, considerando-se as restrições do problema. Caso contrário, o indivíduo é descartado e não entra na população. Desta forma, os indivíduos são adicionados até atingir o tamanho da população inicial, definido neste trabalho em

500 indivíduos. Se ocorrerem 100.000 tentativas improdutivas para gerar um indivíduo, o melhor indivíduo gerado até então é apresentado como solução do problema. A mesma verificação das restrições realizada na fase de geração dos indivíduos é feita após a geração de novos indivíduos por cruzamento e mutação, com taxas de 60% e 5%, respectivamente [2]. Neste trabalho, utiliza-se 100 gerações para a obtenção do resultado final.

A fase de seleção dos indivíduos foi feita utilizando-se o método da roleta com função de aptidão (*fitness*) igual a $\frac{1}{F_{obj}}$. Também foi utilizada a reprodução seletiva, que garante que a sobrevivência de uma solução ótima local pela próxima geração.

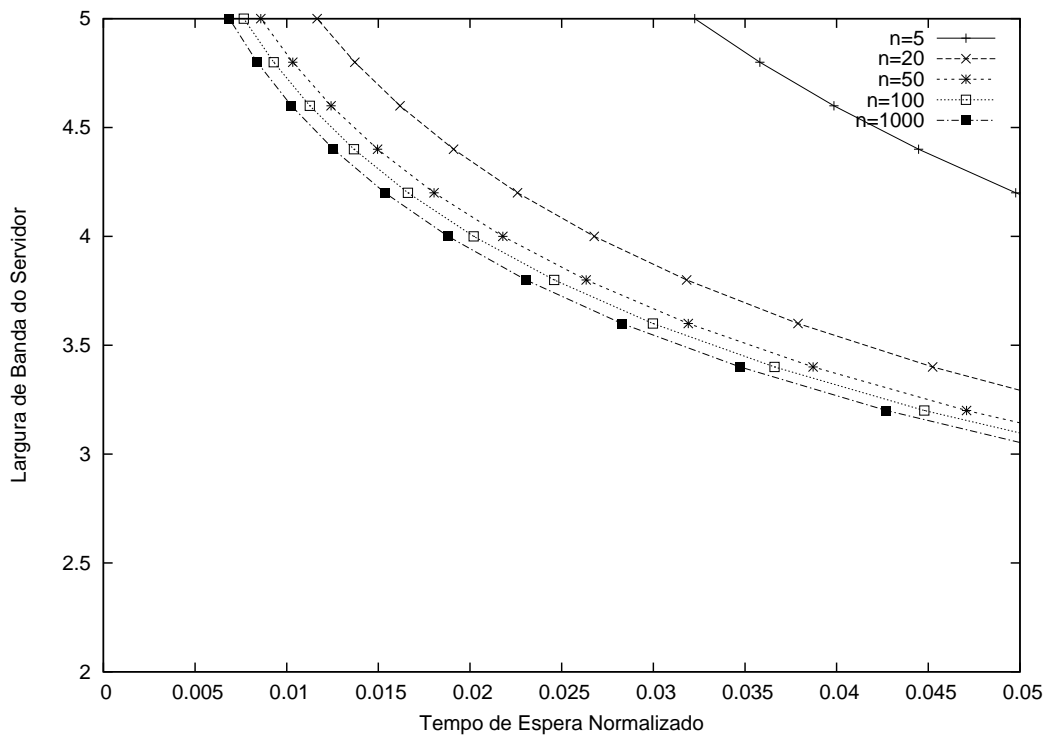


Figura 3. Influência do parâmetro n no GEBB-LBU ($d = 1, k = 5$)

A Figura 3 mostra a influência do parâmetro n sobre a latência de espera para uma limitação de largura de banda a cinco canais ($k = 5$). Considera-se apenas um conjunto de canais. Um aspecto interessante é o ganho de desempenho devido à diminuição da demanda de largura de banda do servidor, advindo do aumento da segmentação. Entretanto, este ganho diminui conforme o aumento de n . O ganho de $n = 5$ para $n = 20$ é muito maior, tanto em termos de latência de exibição quanto de banda passante, do que de $n = 20$ para $n = 100$. O mesmo ocorre para segmentações iguais a 50, 100 ou 1000, seguindo um comportamento logarítmico; a latência obtida com uma segmentação $n = 5$ é de 0,0323 e reduzindo drasticamente se a segmentação aumenta para $n = 20$ (0,0166), obtendo-se um ganho de 0,0206 em relação ao tamanho do vídeo, ou seja, o tempo de espera diminui para dois minutos e meio para um vídeo de duas horas. Por outro lado, para uma segmentação $n = 100$ consegue-se uma latência igual a 0,0077. Caso a segmentação seja dez vezes maior, a latência obtida passa para 0,0069, o que dá um ganho de apenas 0,0008, correspondendo a uma redução na latência de menos de seis segundos para um vídeo de duas horas.

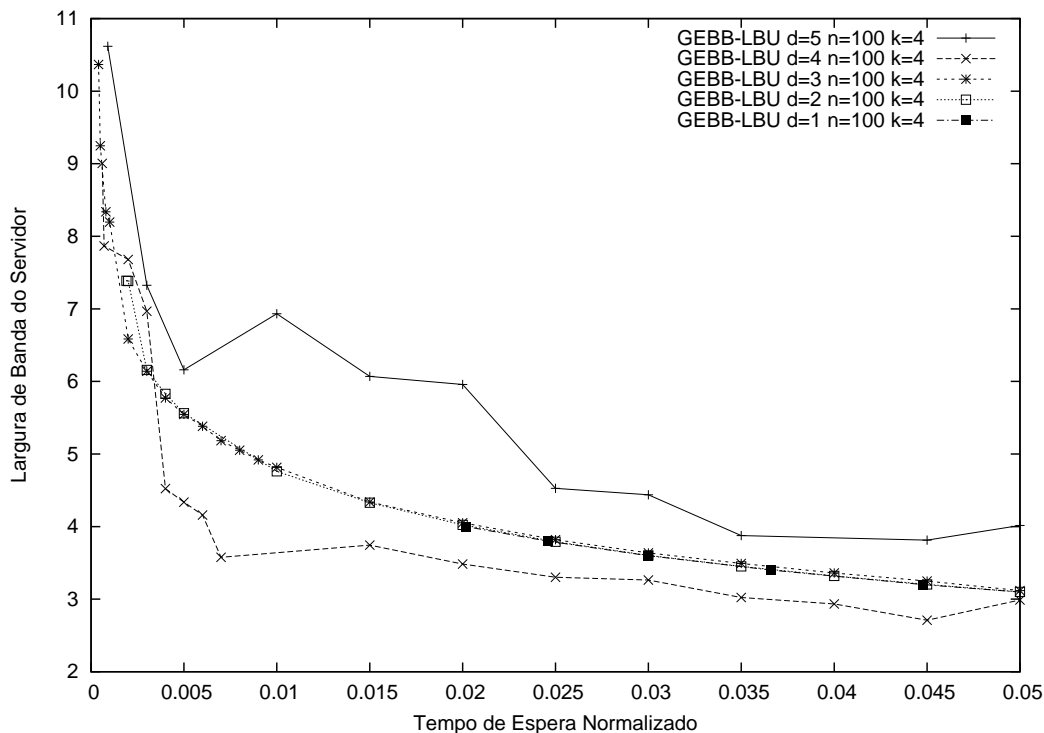


Figura 4. Influência do parâmetro d no GEGB-LBU

A Figura 4 ilustra a influência exercida pelo parâmetro d no protocolo GEGB-LBU. O gráfico mostra melhorias na latência para uma dimensão três; para quatro conjuntos de canais, percebe-se uma grande melhoria na utilização da banda passante, com exceção à faixa entre 0,0020 e 0,0035 (oscilação esta devida à geração de soluções por AG). Já para $d = 5$, nota-se uma queda de rendimento excepcional referente à largura de banda, uma vez que as soluções com esta restrição requerem em média entre $0,5b$ e b a mais do que as soluções utilizando-se $d = 2$ ou $d = 3$, chegando a uma diferença de mais de $2b$ para uma latência de 1% ($7b$ para a solução de $d = 5$ e $4,8b$ quando $d = 3$). Embora tenha conseguido boas soluções de tempo de espera, não houve uma melhoria de latência em relação aos outros valores de d .

As restrições impostas por problemas de PDPH-LBU e GEGB-LBU com dimensões grandes acabam trazendo problemas durante a fase de geração da população. Uma vez que esta é totalmente aleatória, e em algumas situações não se consegue gerar toda a população para execução do restante do algoritmo. Isto acontece, no exemplo anterior, onde soluções com $d = 4$ aparecem fora da tendência da função, cujas soluções inferiores a soluções com menos de quatro conjuntos de canais (na figura, uma diferença de $0,3b$ para um w/S igual a 0,002 e $0,8b$ para uma fração 0,003).

Outro possível fator para a dificuldade na obtenção de soluções para altas dimensões pode estar na granularidade das soluções. Neste trabalho, utiliza-se como valor mínimo para um segmento o equivalente a uma fração de 0,01% do tamanho total de um vídeo (exemplificando, 0,72s para um vídeo de duas horas). Considerando um fluxo MPEG-2 de 6Mbps, o segmento passa a ter um tamanho mínimo de 4,32Mb, o que indica a possibilidade de utilização de uma granularidade ainda maior.

5. Comparação

Nesta seção, é realizada uma comparação entre os protocolos que possuem limitação na largura de banda. Considera-se a largura de banda demandada pelo servidor e a latência mínima como parâmetros de comparação. Protocolos diferentes são comparados somente quando possuem a mesma limitação de banda passante do usuário.

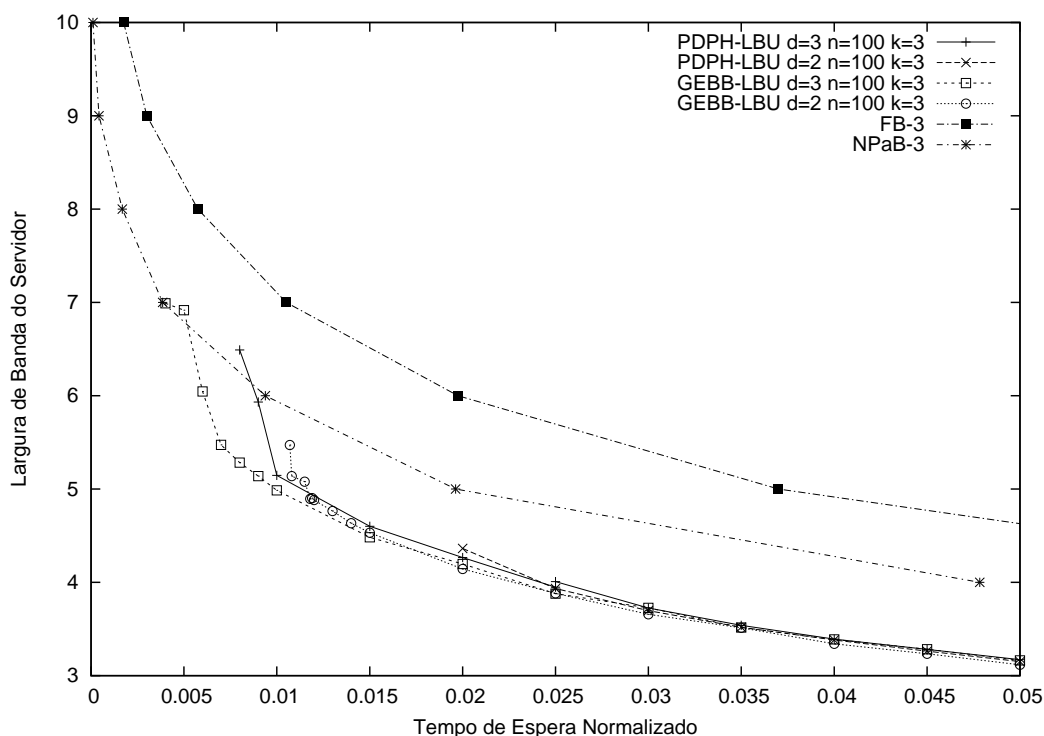


Figura 5. Comparação entre PDPH-LBU, GEGB-LBU, PDR-3 e NPDPa-3

Na Figura 5, uma comparação dos protocolos otimamente estruturados com outros protocolos que utilizam limitação de largura de banda dos usuários, PDR-3 e NPDPa-3, ambos limitados a uma recepção máxima de três canais.

A versão do GEGB-LBU com um conjunto de canais, que teve como sua melhor fração de latência o valor 0,0550, foi excluída do gráfico. Pode-se considerar um valor razoável levando-se em conta que esta versão não possui atraso na recepção do cliente. O mesmo ocorre com a versão do PDPH-LBU, cuja melhor latência obtida foi de 0,0582.

A versão do GEGB-LBU com dois conjuntos de canais obteve um resultado bem mais expressivo que a versão com apenas um canal (0,0107), ou seja, um minuto e 17 segundos para um vídeo de duas horas. O ganho do PDPH-LBU também é bem expressivo, obtendo uma latência igual a 0,02.

Já com a versão do GEGB-LBU com três conjuntos de canais ($d = 3$) obteve-se um ganho de latência em relação à versão com dois conjuntos. Para resultados na faixa de latência acima de 0,01, obteve-se em média uma demanda de largura de banda aproximadamente 1,27b a menos que o PDR-3 e 0,7b a menos que o NPDPa-3. O desempenho do PDPH-LBU é ligeiramente inferior.

Para valores de latência entre 0,01 e 0,005 há um grande aumento na demanda

de banda, tanto no PDPH-LBU quanto no GEBB-LBU. Os protocolos PDR e NPDPa limitados são os que conseguem apresentar solução factível para latências menores que 0,5% (36 segundos para um filme de duas horas).

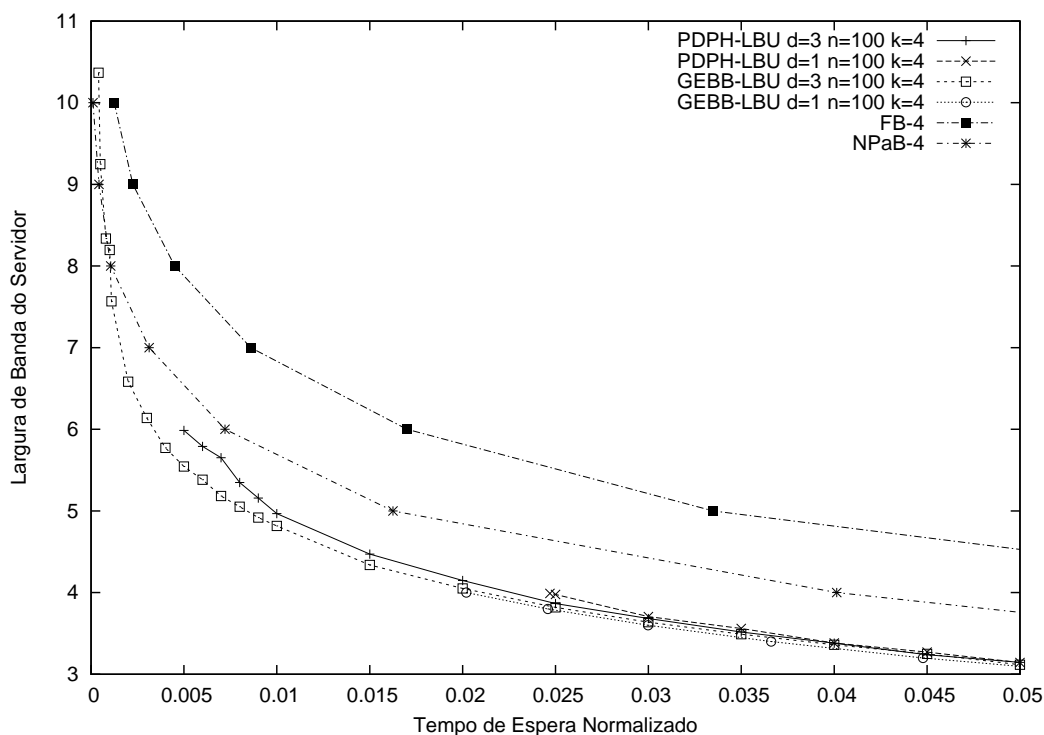


Figura 6. Comparação entre PDPH-LBU, GEBB-LBU, PDR-4 e NPDPa-4

O gráfico da Figura 6 ilustra a mesma comparação que o gráfico anterior, agora com uma limitação na largura de banda do cliente de quatro canais, onde se pode notar uma melhoria no desempenho do protocolo GEBB-LBU. Utilizando somente um conjunto de canais. O GEBB-LBU não consegue uma latência mínima boa, pois não há atrasos na recepção dos conjuntos, sendo 2,5% do vídeo (o equivalente a três minutos de espera para um vídeo de duas horas) a menor latência encontrada. Mesmo assim, vê-se neste ponto do gráfico a ótima demanda de largura de banda do servidor, apenas $4b$, aproximadamente b e $1,8b$ a menos que a banda requerida pelo servidor nos protocolos NPDPa-4 e PDR-4, respectivamente. Para um tempo de espera equivalente a 1% (72 segundos para um vídeo de duas horas), o GEBB-LBU requer $4,98b$, enquanto que o NPDPa-4 requer $6b$ e o PDR-4 $7b$. O PDPH-LBU apresenta resultado semelhante, demandando apenas pouca quantidade de banda passante adicional. Além disto, este protocolo demonstrou diminuição na latência em comparação com a versão limitada a $k = 3$.

Já no caso do GEBB-LBU com três conjuntos de canais, observa-se uma nítida redução no tempo de espera obtido. A menor latência obtida com um resultado superior ao NPDPa-4 é uma fração de 0,0011 (requerendo uma largura de banda de $7,568b$) do vídeo utilizando 100 segmentos, o que equivale a um tempo de espera de 7,92 segundos para um vídeo de duas horas. Este valor é satisfatório, correspondendo a um compromisso entre complexidade de gerenciamento, latência e utilização de largura de banda. O protocolo PDPH-LBU não conseguiu obter uma latência tão reduzida (0,005) quanto o restante dos

protocolos. Mesmo assim, demanda banda do servidor aproximadamente $2b$ a menos que o PDR-4 e $0,7b$ a menos que o NPDPa-4.

Comparando apenas os protocolos otimamente estruturados, PDPH-LBU e GEBB-LBU, nota-se uma clara vantagem do GEBB-LBU, que obtém sempre uma demanda por largura de banda do servidor ligeiramente menor (aumentando a diferença conforme a diminuição da fração w/S). A vantagem do GEBB-LBU na latência obtida é percebida mais claramente, como no caso da utilização de três conjuntos de canais. O GEBB-LBU consegue obter um expressivo tempo de espera de 0,04%, contra 0,5% obtido pelo PDPH-LBU. Esta diferença decorre principalmente devido ao tipo de segmentação utilizado pelos protocolos. Embora ambos sejam protocolos otimamente estruturados, uma segmentação crescente como a do GEBB-LBU é muito mais eficaz que a segmentação utilizada pelo PDPH-LBU, pois com o mesmo número de segmentos obtém-se um primeiro segmento de tamanho menor, reduzindo-se, assim, a latência de exibição.

Pode-se dizer que protocolos otimamente estruturados beneficiam-se enormemente de restrições (de clientes) mais fracas de largura de banda, principalmente na faixa entre $4b$ e $5b$, enquanto possuem um menor desempenho para restrição de largura de banda de clientes variando entre $k = 2$ e $k = 3$, onde não se obteve uma latência tão boa quanto para valores de k entre 4 e 5.

Uma vantagem dos protocolos otimamente estruturados é que se pode atribuir qualquer valor a k , permitindo assim qualquer restrição do usuário. A metodologia utilizada pelos protocolos da família Pagode utilizam sempre valores de k inteiros, pois estes alocam C canais no servidor de largura de banda b e os clientes recebem dados de até k canais simultâneos, o que nem sempre corresponde à realidade.

A utilização de mensagens de direitos autorais ou mesmo inserções comerciais durante o tempo de espera do vídeo[4] é uma das alternativas para a ocupação do tempo de espera. Protocolos otimamente estruturados têm nesta abordagem uma facilidade adicional, uma vez que a latência de exibição dos mesmos é sempre fixa.

Contudo, os novos protocolos apresentados também apresentam uma desvantagem. Como todos os outros protocolos com clientes limitados em banda passante, tanto o PDPH-LBU quanto o GEBB-LBU são específicos para um valor único de k . Em uma rede heterogênea, com clientes utilizando diferentes restrições de canais, há a necessidade de se utilizar transmissões exclusivas para cada caso ($k = 3, k = 4, \dots$). Por outro lado, o NPDPa atingiu níveis de latência não alcançados pelos protocolos otimamente estruturados para uma limitação de três canais simultâneos. Já o PDR sempre obtém as mesmas latências que NPDPa, embora sempre requeiram mais banda do servidor (b , em média).

Entretanto, cabe salientar o aumento de complexidade dos protocolos PDR e NPDPa limitados em função da diminuição da latência. Para reduzir o tempo de espera em um protocolo da família Pagode, deve-se obrigatoriamente aumentar o número de segmentos, aumentando também a complexidade de gerenciamento dos mesmos. Isto não acontece com o PDPH-LBU nem tampouco com o GEBB-LBU, onde a limitação de complexidade é feita através da limitação do parâmetro n_{max} .

6. Conclusão

Este artigo introduziu um novo protocolo de difusão periódica sujeito à limitação de banda passante, denominado GEBB-LBU, que forma juntamente com o PDPH-LBU uma classe de protocolos otimamente estruturados com limitação de banda passante estudada no presente trabalho. Além disto, comparou-se os protocolos desta classe com protocolos não otimamente estruturados. O protocolo GEBB-LBU destaca-se pelo ótimo desempenho principalmente com tempos de 1% do tempo de duração do vídeo para uma limitação de $k = 3$ e de 0,11% para uma limitação de $k = 4$.

A técnica de pré-carregamento dos prefixos (*prefix caching*) [5] pode ser uma boa alternativa para prover acesso instantâneo a vídeos populares transmitidos por um protocolo de difusão periódica. É interessante também realizar um estudo sobre a integração desta técnica com protocolos com limitação de banda do usuário, como o PDPH-LBU e o GEBB-LBU.

Agradecimentos

Este trabalho foi financiado parcialmente pelo CNPq e pela CAPES.

Referências

- [1] Ailan Hu. Video-on-Demand broadcasting protocols: A comprehensive study. In *Proceedings of the IEEE Infocom 2001 Conference*, Anchorage, Alaska EUA, abr. 2001.
- [2] John R. Koza, Forrest H. Bennett III, David Andre, and Martin S. Keane, editors. *Genetic Programming III: Darwinian invention and problem solving*. San Francisco: Morgan Kaufmann, 1999.
- [3] Scott R. Ladd. *Genetic Algorithms In C++*. New York: M&T Books, 1996.
- [4] Jehan-François Pâris, Steven W. Carter, and Darrell D. E. Long. A low bandwidth broadcasting protocol for video on demand. In *Proceedings of the 7th IC3N'98*, Lafayette, LA, EUA, pp. 690-697, out. 1998.
- [5] Jehan-François Pâris, Darrell D. E. Long, and Patrick E. Mantey. Zero-delay broadcasting protocols for video-on-demand. In *Proceedings of 7th ACM International Conference on Multimedia*, Orlando, FL, Estados Unidos, pp. 189-197, out./nov. 1999.
- [6] Jehan-François Pâris and Darrell D. E. Long. Limiting the receiving bandwidth of broadcasting protocols for video-on-demand. In *Proceedings of the Euromedia 2000 Conference*, Antuérpia, Bélgica, pp. 107-111, maio 2000.
- [7] Darrell Whitley. An overview of evolutionary algorithms: Practical issues and common pitfalls. *Information and Software Technology*, 43(14):817–831, 2001.
- [8] Rogério M. Zafalão and Nelson L. S. Fonseca. Um protocolo para a difusão periódica de vídeo sob limitação de banda passante. In *XXI SBT*, Belem, PA, set. 2004.
- [9] Rogério M. Zafalão, Nelson L. S. Fonseca, and Cid C. Souza. O protocolo polyharmonic broadcasting sujeito a limitação de banda passante. In *XXI SBRC, SBC*, volume I, pages 397–410, Natal, RN, maio 2003.