

Uma Nova Abordagem para Recepção de Vídeo VBR em Camadas

Arthur de Castro Callado¹, Carlos Alberto Kamienski¹, Judith Kelner¹

¹Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7851 – CEP 50.732-970 – Recife – PE – Brazil

{acc2, cak, jk}@cin.ufpe.br

Resumo. *A utilização de vídeo está se tornando cada vez mais popular na Internet. Nesse artigo são apresentadas algumas das principais abordagens de codificação e de transmissão de vídeo utilizadas na Internet. Um arcabouço combinando as abordagens de vídeo codificado em camadas com o descarte seletivo na transmissão e com a adaptação pelo receptor foi sugerido para obter um efeito sinérgico considerável. Um estudo de caso realizado através de simulação estocástica com o Network Simulator 2 mostrou que a técnica proposta obtém ganhos significativos no aumento da qualidade de vídeo recebido e na diminuição da taxa de perda de pacotes.*
Palavras-chave: *codificadores de vídeo, vídeo em camadas, vídeo adaptativo, qualidade de serviço (QoS), Serviços Diferenciados.*

Abstract. *Video is becoming increasingly popular in the Internet. In this paper, we present some of the commonly used approaches to coding video and transporting it over computer networks. We also propose a framework that combines layered video coding with selective packet dropping and with receiver-driven adaptation in order to obtain a reasonable synergetic effect. A stochastic simulation-based evaluation made with Network Simulator 2 showed that the proposed scheme is able to offer significant gains in terms of improving video quality and decreasing the packet loss ratio.*

1. Introdução

Para a Internet evoluir efetivamente para uma rede multi-serviços é fundamental que ela seja capaz de tratar adequadamente tráfego multimídia, como voz e vídeo. Atualmente, a Internet oferece somente o serviço de melhor esforço, que dá o mesmo tratamento para todos os usuários e aplicações. Os pacotes em excesso são descartados sem levar em consideração a sua importância para as aplicações que os geraram. Para que seja possível dar um tratamento especial a tais aplicações é necessário introduzir novos serviços, baseados em garantias de Qualidade de Serviço (QoS) [10].

Nesse contexto, a arquitetura de serviços diferenciados (DiffServ), definida na RFC 2475 [2], vem sendo apontada como uma das tecnologias capazes de possibilitar a introdução de aplicações multimídia, oferecendo eficiência (pouca ou nenhuma sinalização e pouco processamento nos roteadores) e escalabilidade. Tratando agregados de fluxos em cada nó da rede (*Per-Hop Behavior* - PHB), onde cada tipo de serviço recebe um tratamento diferido, é possível oferecer serviços avançados na Internet.

De acordo com previsões de especialistas [9], vídeo compactado com taxa variável de bits (VBR) tende a se tornar onipresente no futuro, em aplicações tais como vídeo sob demanda (VoD), videoconferência, videofone e televisão digital de alta definição. A utilização de vídeo VBR é importante, pois possibilita a obtenção de ganhos consideráveis de multiplexação estatística. Em conjunto com técnicas como vídeo em camadas, adaptação da aplicação e descarte baseado em prioridade, vídeo VBR obtém um efeito sinérgico capaz de facilitar a sua implantação na Internet.

Nesse artigo é proposto um arcabouço para transmissão de vídeo combinando as técnicas citadas anteriormente. Além disso, é realizada uma avaliação preliminar desta técnica através de simulações. Utilizando o DiffServ, foi possível conferir prioridade de descarte [1] e, através de um mecanismo adaptativo, diminuir a taxa de perda de pacotes. Juntando as duas abordagens, obtém-se um novo mecanismo, robusto e suave.

Na seqüência deste artigo a seção 0 apresenta uma revisão dos principais conceitos de transmissão de vídeo, bem como o mecanismo de adaptação proposto. A seção 0 apresenta o modelo da simulação realizada. Resultados da simulação são apresentados na seção 0. Finalmente, a seção 0 apresenta as conclusões desse trabalho, bem como indicações para trabalhos futuros.

2. Técnicas de Codificação/Transmissão de Vídeo

A transmissão de vídeo em redes exige uma grande largura de banda. Com uma resolução de 1024 x 768 pixels com 24 bits de cor (16 milhões) e 30 quadros por segundo, o vídeo será codificado com a taxa de saída de 566 Mbps. Para transmissões de televisão digital de alta definição chega-se até 1,2 Gbps. Mas transmitir dados a essas taxas elevadas exige muitos recursos, inviabilizando a sua oferta ao público em geral.

Existem algumas técnicas de codificação e algoritmos de compactação de vídeo que reduzem bastante a taxa gerada pelas aplicações e conseqüentemente a largura de banda exigida para transmissão. Além disso, nem todas as aplicações de vídeo necessitam de resoluções altas nem grande quantidade de quadros por segundo (em aplicações de videoconferência os usuários se mexem pouco e não há mudanças de cena), embora algumas sim (vídeo sob demanda – VoD - é uma aplicação que tem muitas mudanças de cena e movimentos rápidos, exigindo taxas mais altas de transmissão). Nos últimos anos foram criados vários padrões para geração e transmissão de vídeo compactado, com várias finalidades [17]. A seguir, algumas técnicas.

2.1. Vídeo VBR

Uma forma de codificar vídeo é utilizar o mesmo algoritmo para cada quadro, gerando uma taxa constante de bits (CBR). A codificação CBR gera uma taxa fixa e previsível para a rede, facilitando o provisionamento de banda. Contudo, possui a desvantagem de desnivelar a qualidade do vídeo gerado. Para uma determinada taxa de transmissão, alguns quadros podem sofrer maior compressão e serão codificados com uma qualidade melhor. Em outros a taxa de compressão é menor, e para não exceder a taxa de transmissão, a qualidade deve ser sacrificada.

Uma alternativa é a codificação com taxa variável de bits (VBR). VBR não tenta manter uma taxa de saída constante do codificador de vídeo, mas a taxa de transmissão varia de acordo com o resultado da compressão. O VBR é possível quando se utilizam técnicas de compensação de movimento e transparência.

A codificação VBR faz com que a taxa de pico de uma fonte atinja valores consideravelmente maiores que a taxa média (de 2 a 5 vezes, dependendo da aplicação). Isso possibilita ao vídeo VBR a obtenção de ganhos de multiplexação estatística [13]. Esse ganho pode ser definido como a razão entre o número de fontes VBR e o número de fontes CBR transmitindo na taxa de pico em uma mesma rede. Essa razão deve sempre estar acima de 1, pois com um grande número de fontes de vídeo, a taxa agregada será aproximadamente igual à taxa média multiplicada pelo número de fontes.

No entanto, o ganho de multiplexação estatística de vídeo VBR apresenta alguns desafios para o provisionamento de recursos na rede (largura de banda e memória). Com vídeo CBR o provisionamento é feito pela taxa de pico acrescido de uma margem de segurança. Para vídeo VBR, o provisionamento pode ser feito para a taxa média, mas o controle de descarte de pacotes torna-se um ponto fundamental. Como não se podem controlar as rajadas das fontes, várias delas podem simultaneamente transmitir a sua taxa de pico. Mesmo que por curtos períodos de tempo, essa ocorrência leva a perdas de pacotes. Provisionar a rede corretamente é um dos desafios da transmissão VBR.

Existem centenas de codificadores de vídeo com o objetivo de transformar vídeo analógico (de uma fonte como câmera filmadora, vídeo cassete ou televisão) em vídeo digital para transmissão ou para armazenamento. Existem também codificadores para transformar vídeo digital para outros formatos. Alguns dos codificadores mais utilizados são: MPEG [11] (*Moving Picture Experts Group*), DivX [4] e Indeo [12].

2.2. Vídeo em Camadas

Vídeo em camadas [2] foi uma técnica de codificação proposta para transmissão de vídeo com difusão seletiva (*multicast*). O objetivo é possibilitar que vários usuários possam receber o mesmo vídeo com qualidades e taxas diferentes, optando estaticamente, de acordo com a sua rede de acesso ou o preço, ou dinamicamente, de acordo com as condições momentâneas da rede [1]. Nesse artigo, vídeo em camadas não está sendo utilizado no contexto de difusão seletiva, mas para descartar seletivamente os pacotes (por camada) e para utilizar um mecanismo de adaptação, no qual o receptor avalia a qualidade recebida e guia o transmissor (seção 0).

Essa técnica utiliza algoritmos que particionam o sinal básico de vídeo em várias camadas. A camada base possui uma imagem primitiva e cada camada mais alta possui informações que refinam a qualidade da imagem. Por esse motivo, a recepção de uma camada n somente faz sentido se a camada $n - 1$ tiver sido recebida. Para cada camada adicional que se deseja transmitir, a taxa de transmissão cresce de acordo com uma relação que depende do algoritmo utilizado.

2.3. Mecanismo de Adaptação

Algoritmos que adaptam a taxa de transmissão são uma abordagem alternativa para obter maior eficiência na utilização dos recursos da rede [3]. Em conjunto com vídeo em camadas, possibilita a adaptação da taxa transmitida sem afetar a codificação.

Para não desperdiçar banda, esse artigo propõe a utilização do seguinte mecanismo de adaptação: o receptor, periodicamente avalia sua qualidade de recepção de vídeo (em função da taxa de perda de pacotes por camada) e faz o transmissor alterar a sua taxa de transmissão se necessário (em número de camadas enviadas). Toda a decisão é feita no receptor, cabendo ao transmissor somente efetivá-las.

O receptor define os limiares mínimo e máximo de perda na recepção de cada camada. Assim, a taxa de perdas em cada camada deve ser medida e comparada com estes limiares. Se estiver abaixo do limiar mínimo, pode tentar receber a camada superior; se estiver acima do limiar máximo, deve subtrair aquela camada. Nas camadas mais baixas as perdas de pacotes têm efeito mais nocivo à construção da imagem, necessitando de limiares mais rígidos (menores). Para conferir alguma estabilidade ao método, o receptor deve fazer algumas observações da taxa de perda antes de tomar uma decisão. Nesse artigo, utilizamos valores baseados em avaliações preliminares feitas em função da adaptabilidade e da estabilidade do mecanismo de transmissão.

3. Modelo de Simulação

Uma avaliação preliminar do mecanismo de adaptação de vídeo em camadas foi realizada através de simulação utilizando o simulador de redes ns-2 (Network Simulator 2) [13], para o qual foi criado um gerador/transmissor de vídeo VBR em camadas com marcação de prioridade nos pacotes e um receptor de vídeo que monitora a qualidade de recepção e envia ao transmissor um pedido de alteração da taxa se necessário.

Como o objetivo foi somente avaliar o desempenho do mecanismo na rede, não foi realizada uma simulação do comportamento completo do codificador de vídeo. Apenas a geração de tráfego sintético (que poderia ser a leitura de vídeo armazenado em disco) e o comportamento (este completo) do transmissor foram implementados. Diferentemente do que é mostrado em outros trabalhos [1][3][6][8], o modelo utilizado visou contemplar simultaneamente o descarte por prioridade e o mecanismo de adaptação guiado pelo receptor (otimização da utilização dos recursos de rede).

As camadas IP (*Internet Protocol*, para roteamento [14]), UDP (*User Datagram Protocol*, para envio sem conexão [15]) e RTP (*Real Time Protocol*, para controle de tempo real [16]) já existiam no ns-2 [13]. As perdas de pacotes por erro na camada física não foram consideradas na simulação, somente por congestionamento. Isso porque a topologia utilizada não apresenta enlaces sem fio (onde a perda seria um fator relevante). Entretanto, devido à estabilidade do mecanismo, eventuais erros (em fibra ótica, são da ordem de 10^{-12}) não provocariam uma diminuição da taxa de transmissão.

3.1. Topologia de rede

A topologia utilizada foi escolhida pela simplicidade e pela necessidade de haver compartilhamento de um enlace, de modo a investigar como os mecanismos de transmissão e recepção de vídeo se comportam quando há um gargalo na rede. Essa topologia é conhecida como alteres (*dumbbell*).

O provisionamento do enlace (4 Mbps) foi feito baseado na taxa de pico de 6 transmissores, i.e., com 6 transmissores o descarte de pacotes da rede é nulo. Para isto, calculamos a taxa de transmissão média de cada camada: 32, 64 e 128 Kbps. Como a taxa de pico é de 3 vezes a taxa média, um transmissor pode enviar até 672 Kbps em uma rajada. Assim, para 6 transmissores, teremos uma taxa máxima de 4,032 Mbps. Mas devido ao ganho de multiplexação estatística, mesmo com 10 transmissores descartes de pacotes são raros. No estudo de caso (Seção 0), foram realizadas simulações para 10, 12, 14, 16, 18 e 20 transmissores.

3.2. Modelo de tráfego

Foi utilizado um modelo de tráfego baseado em [1] e [5] para a geração de tráfego sintético VBR (i.e., somente a taxa gerada é fiel à realidade, não o conteúdo dos dados), onde a camada base é gerada em intervalos de 1 segundo. Em cada intervalo, n pacotes são transmitidos, onde n é escolhido de acordo com a seguinte distribuição aleatória:

$$n = \begin{cases} 1, & \text{probabilidade} = 1 - \frac{1}{P} \\ PA + 1 - P, & \text{probabilidade} = \frac{1}{P} \end{cases}$$

P representa a intensidade das rajadas e fornece uma indicação da razão entre taxa de pico e taxa média (peak-to-mean ratio). A é o número médio de pacotes por intervalo e foi configurado para 4 pacotes de 1KB nessa avaliação. Os n pacotes são transmitidos em uma única rajada, iniciando em um tempo aleatório (distribuição uniforme) dentro do intervalo. Para cada camada mais alta c , o intervalo é quebrado em 2^c subintervalos e n pacotes são enviados em um tempo aleatório em cada um desses.

Nessa avaliação, o gerador de vídeo VBR em camadas foi implementado para utilizar um máximo de 3 camadas. A taxa média de geração é de 4, 8 e 16 pacotes para a 1^a, a 2^a e a 3^a camadas, respectivamente. O tamanho do pacote foi fixado em 1000 bytes. Assim, a taxa média gerada por um transmissor utilizando as 3 camadas é de $(4 + 8 + 16) \times 8 \times 1000 = 224$ Kbps. Já a razão entre a taxa de pico (rajada) e a taxa média que utilizamos foi de 3. Assim, a taxa de pico é de 672 Kbps.

3.3. Configurações DiffServ

Foi utilizada a implementação de DiffServ para o ns-2 desenvolvida pela Nortel e incorporada à distribuição padrão do ns-2. Os roteadores foram configurados para utilizar o mecanismo DropTail em simulações que utilizam descarte uniforme e dsRED (DiffServ-enabled Random Early Detection – onde o descarte é feito em função da prioridade do pacote e dos tamanhos médio e atual da fila) ao utilizar descarte prioritário. A configuração dos parâmetros do dsRED (baseado no RED [5]) foi feita através de um PHB AF (mecanismo de encaminhamento de pacotes de Serviço Assegurado [7]) com 3 prioridades de descarte (fila real com 3 filas virtuais), uma para cada camada. O limiar de descarte do dsRED para cada prioridade foi definido de forma proporcional à taxa média da camada correspondente, sendo a probabilidade de descarte de pacotes de uma camada nula abaixo de seu limiar e igual a 1 acima dele. O peso do tamanho atual da fila no cálculo do tamanho médio da fila foi de 1.

No transmissor, para especificar a qualidade de serviço do pacote, cada pacote é marcado com o *codepoint* referente a sua camada. Utilizamos os valores 10, 11 e 12 para as camadas 1, 2 e 3, respectivamente.

3.4. Funções de avaliação

O cálculo da função utilidade é feito pelo receptor com base nos pacotes recebidos e perdidos em cada camada. Cada camada tem um limiar mínimo de recepção para ser considerada (camadas com muitas perdas não podem ser aproveitadas) e as camadas inferiores têm um peso maior na qualidade do vídeo recebido, pois servem de base para as camadas superiores (codificação incremental em camadas). Em uma amostra, para

cada usuário i temos a taxa de chegada dos pacotes na camada c , definida como a razão entre os pacotes recebidos e os transmitidos. Com esta taxa e com os limites l_c de aproveitamento da camada c , definimos o aproveitamento da camada c de um usuário i .

$$t_{ic} = \frac{C_{ic}}{T_{ic}} \quad a_{ic} = \begin{cases} 1, & t_{ic} \geq l_c \\ 0, & t_{ic} < l_c \end{cases}$$

Assim, podemos definir matematicamente a função utilidade média (parametrizando o resultado para o intervalo real $[0;10]$ para o melhor entendimento dos resultados) em função do número de receptores n , do aproveitamento e do peso de cada camada. O cálculo da perda de pacotes (em porcentagem) é dado pela média ponderada do total de pacotes perdidos por cada receptor i pelo total de pacotes transmitidos por cada transmissor i (parametrizado para uma porcentagem, para melhor comparação).

$$\bar{u} = 10 \times \frac{\sum_{i=1}^n \sum_{c=1}^3 a_{ic} p_c}{n \sum_{c=1}^3 p_c} \quad \bar{p} = 100 \times \frac{\sum_{i=1}^n \sum_{c=1}^3 P_{ic}}{\sum_{i=1}^n \sum_{c=1}^3 T_{ic}}$$

4. Resultados

Nesta seção, serão apresentados os principais resultados obtidos com os métodos, cenários e técnicas descritos. Comparamos os mecanismos de transmissão de vídeo em camadas em função da qualidade do vídeo recebido (cálculo da função utilidade) e da perda de pacotes. O tempo de simulação para cada configuração foi de 6000 segundos, com amostragem a cada 60 (100 amostras por configuração). Os gráficos foram construídos com intervalos de confiança assintóticos no nível de 99,9% (barra vertical).

4.1. Descarte Uniforme x Prioridade de Descarte

Essa seção compara o descarte uniforme de pacotes (Internet atual), com o descarte prioritário (DiffServ). Ao compará-los (Figura 1), verificamos que a função de utilidade registra rápida queda na qualidade do método uniforme com o aumento da carga. Como as perdas de pacotes ocorrem uniformemente nas camadas, o vídeo perde dados mais importantes, diminuindo sua qualidade. No método baseado em prioridade, a queda é mais suave, pois nas camadas mais baixas os pacotes têm preferência.

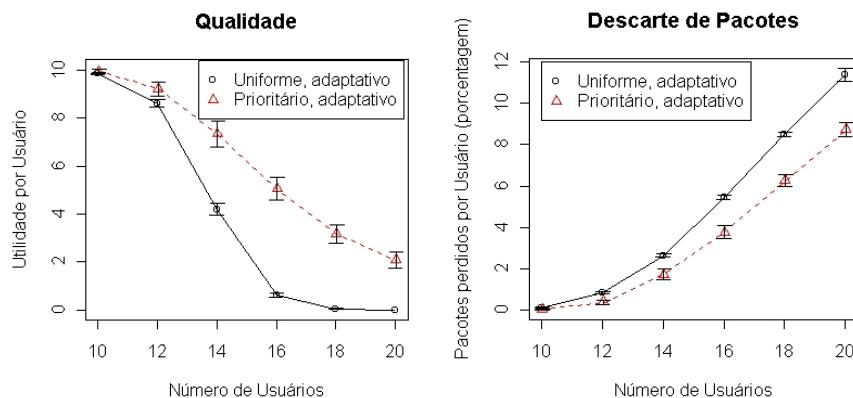


Figura 1 – Qualidade e nível de descarte: descarte prioritário versus uniforme

O ganho de qualidade pelo descarte com prioridade sobre o descarte uniforme é de até 5 vezes. Como os pontos no gráfico se referem à média das observações e os intervalos de confiança são disjuntos, podemos afirmar que, baseado nessas simulações, os benefícios do descarte com prioridade compensam o aumento de complexidade.

Com relação ao descarte, a Figura 1 mostra que o método com prioridade descarta sistematicamente menos pacotes do que o método uniforme. No método de descarte uniforme os pacotes são descartados aleatoriamente em todas as camadas, prejudicando-as igualmente. Já no método prioritário, os descartes iniciam nas camadas mais altas, gerando taxas de perdas mais altas nessas camadas e conseqüentemente levando o receptor a tomar a decisão de redução do número de camadas mais cedo.

4.2. Adaptativo x Não-Adaptativo

Comparando a função de utilidade para os mecanismos de descarte prioritário adaptativo e não adaptativo (Figura 2), pode-se observar que o mecanismo adaptativo perde qualidade mais lentamente que o não-adaptativo, mas ambos têm uma perda de qualidade gradual, devido ao descarte prioritário. A figura mostra também que a rede está provisionada corretamente somente para 10 usuários transmitindo na taxa máxima. Vale lembrar que a rede foi provisionada para 6 usuários na taxa de pico, o que gera um ganho de multiplexação estatística de 1,67. Com essa quantidade pequena de usuários, esse valor pode ser considerado promissor.

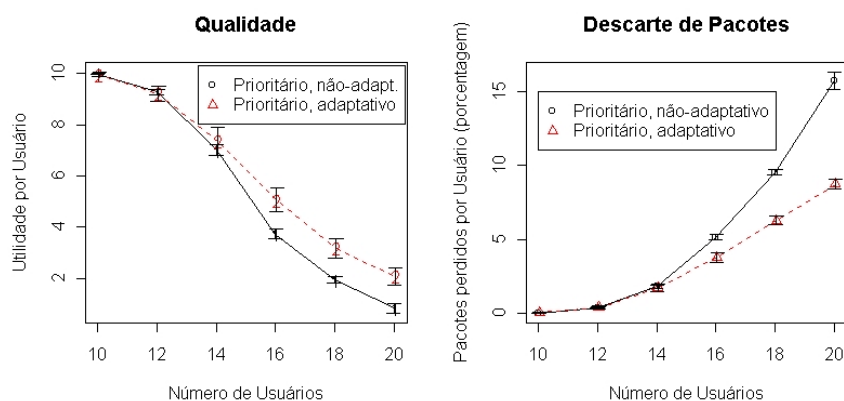


Figura 2 – Qualidade e nível de descarte: adaptativo versus não-adaptativo

Porém, ao observarmos o descarte de pacotes, vemos que este cresce exponencialmente no mecanismo não-adaptativo, enquanto que no adaptativo, após o limite de 14 usuários, cresce linearmente. Isso mostra a necessidade do mecanismo de adaptação para vídeo em camadas com descarte por prioridade, porque sem adaptação a quantidade de pacotes submetida à rede é muito grande e pode afetar outros fluxos com prioridade inferior (e.g., www, ftp, etc) sem ganho de qualidade no vídeo.

5. Conclusões

As simulações realizadas nos mostraram que a combinação de estratégias de recepção de vídeo pode ser uma boa solução para melhorar a qualidade de recepção de vídeo e tornar mais eficiente a utilização dos recursos da rede. A função de utilidade que mede a qualidade da recepção mostrou ganhos significativos tanto com a utilização de descarte prioritário quanto com a adaptação de vídeo pelo receptor.

Embora ainda não seja utilizado em larga escala na Internet, DiffServ é uma tendência para fornecer QoS às aplicações multimídia avançadas, como voz e vídeo. Com aplicações cientes da qualidade necessária e de como melhor aproveitar os recursos disponíveis tomando decisões em tempo real, diminui a complexidade da rede (sem marcadores de pacotes) e tem-se um ganho de qualidade na utilização da mesma.

Os resultados apresentados se referem a uma avaliação preliminar do desempenho da técnica proposta. Existem alguns melhoramentos que podem ser feitos na técnica de transmissão de vídeo e no cenário avaliado. O mecanismo de decisão do receptor pode ser melhorado. Uma alternativa seria incorporar uma função estatística que evitasse a oscilação do número de camadas ao mesmo tempo em que procurasse transmitir o número ideal de camadas para melhor aproveitar a banda. Outros trabalhos futuros incluem a escolha dos parâmetros de acordo com vários cenários.

6. Referências

- [1] BAJAJ, S., BRESLAU, L. & SHENKER, S., “Uniform versus Priority Dropping for Layered Video”, ACM SIGCOMM’98, Setembro 1998.
- [2] BLACK, D. et. al., ”An Architecture for Differentiated Services”, RFC 2475, Dezembro 1998.
- [3] BOLOT, J.-C. & TURLETTI, T., “A rate control mechanism for packet video in the Internet”, IEEE Infocom, Junho 1994.
- [4] HERON, Robert, “First Look: DivX 5.0”, Março 2002, <http://www.techtv.com/products/software/story/0,23008,3378072,00.html>
- [5] FLOYD, S. & JACOBSON, V., “Random Early Detection Gateways for Congestion Avoidance”, IEEE/ACM Transactions on Networking, Agosto 1993.
- [6] GOPALAKRISHNAN, R. et. al., “A Simple Loss Differentiation Approach to Layered Multicast”, IEEE Infocom, Março 2000.
- [7] HEINAMEN, J. et. al., “Assured Forwarding PHB Group”, RFC 2597, Junho 1999.
- [8] HOFFMANN, D & SPEER, M., “Hierarchical video distribution over internet style networks”, IEEE International Conference on Image Processing, Setembro 1996.
- [9] IZQUIERDO, M., REEVES, D., “A survey of statistical source models for variable-bit-rate compressed video”, Multimedia Systems, Volume 7, Number 3, 1999.
- [10] KAMIENSKI, C.A. & SADOK, D., “Qualidade de Serviço na Internet”, 18o. SBRC. Maio 2000.
- [11] KOENEN, Rob. “MPEG-4 Overview”. Março 2002, <http://mpeg.telecomitalialab.com/standards/mpeg-4/mpeg-4.htm>
- [12] Ligos Corporation, “INDEO Video 5 Technical Overview”, Agosto 2001, <http://indeo.ligos.com/pi=197.php?>
- [13] Network Simulator (versão 2.1b8a), Outubro 2001, <http://www.isi.edu/nsnam/ns/>.
- [14] POSTEL, J., “Internet Protocol”, RFC 791, Setembro 1981.
- [15] POSTEL, J., “User Datagram Protocol”, RFC 768, Agosto 1980.
- [16] SCHULZRINNE, H. et. al., “RTP: A Transport Protocol for Real-Time Applications”, RFC 1889, Janeiro 1996.
- [17] TANENBAUM, A. S., “Redes de Computadores”, 3ª edição, Editora Campus, 1997.