

Caracterização das Propriedades Dinâmicas da Estrutura da Rede Sobreposta do SopCast *

Kênia Carolina Gonçalves¹, Alex Borges², Jussara Almeida¹, Ana Paula Silva²,
Humberto Marques-Neto³, Sérgio Vale Campos¹

¹Departamento de Ciência da Computação – Universidade Federal de Minas Gerais
Belo Horizonte – MG – Brazil

²Departamento de Ciência da Computação – Universidade Federal de Juiz de Fora
Juiz de Fora – MG – Brazil

³Pontifícia Universidade Católica de Minas Gerais
Belo Horizonte – MG – Brazil

keniacarolina@dcc.ufmg.br, alex.borges@ufjf.edu.br, jussara@dcc.ufmg.br
anapaula.silva@ufjf.edu.br, humberto@pucminas.br, scampos@dcc.ufmg.br

Resumo. *Sistemas Peer-to-Peer (P2P) de transmissão de vídeo ao vivo estão se tornando cada vez mais populares. Apesar de vários estudos sobre o comportamento dos usuários e sobre as propriedades da rede sobreposta formada em tais sistemas, pouco é o conhecimento sobre como a estrutura desta rede evolui com o tempo durante uma transmissão de conteúdo. Nesse trabalho apresentamos uma caracterização das propriedades dinâmicas da estrutura desta rede sob dois pontos de vista: características individuais dos nós e características da rede como um todo. Para isso, foram utilizadas métricas de redes que avaliam a centralidade do nó e a topologia da rede. Os resultados mostram três perfis de nós com métricas de centralidade distintas na rede e pouca variabilidade dos nós entre esses perfis. Além disso, as medidas da topologia da rede também tendem a permanecer estáveis. Finalmente, os resultados obtidos neste trabalho podem ser usados por pesquisas futuras com simulações mais realistas com estratégias otimizadas que representem o dinamismo real da rede sobreposta.*

Abstract. *Peer-to-Peer video streaming systems are becoming increasingly popular. Despite several studies on network overlay and user behavior, little is still known about how the overlay network evolves during a live streaming transmission. In this work we present a network dynamic characterization under two different perspectives: the individual nodes and the overlay network characteristics. We have used complex network metrics such as node centric metrics and network topology properties. Our results show that network nodes may be grouped into 3 distinct clusters. Moreover, network topology metrics tend to be stable during the transmission. The characterization we show may be used as a substrate to future research and system simulation, as it can be used to generate a more realistic synthetic workload.*

*Esta pesquisa é parcialmente financiada pelo Instituto Nacional de Ciência e Tecnologia para a Web - INCTWeb (MCT/CNPq 573871/2008-6), CNPq, FAPEMIG e CAPES.

1. Introdução

A popularidade de aplicações de transmissão de vídeos na Internet cresce a cada dia. As pesquisas recentes mostram que em 2013 o número de usuários desse tipo de aplicação pode chegar a 83 milhões [Borges et al. 2010a]. Uma das arquiteturas de rede utilizadas por essas aplicações na transmissão de vídeo ao vivo pela Internet é a rede Par-a-Par (P2P). O SopCast¹, o PPLive² e o PPStream³ são exemplos deste tipo de aplicação. O SopCast tem apresentado uma grande popularidade e em 2010 teve um maior volume de tráfego mundial se comparado com as aplicações PPLive e PPStream [Google 2010]. Por isso, foi escolhido para ser analisado neste trabalho.

No SopCast, a transmissão de vídeo ocorre em uma rede P2P, na qual os clientes recebem o conteúdo transmitido por outros participantes, criando assim o que é conhecido como rede sobreposta, ou seja, uma rede lógica de transmissão de dados. Nesta rede, os nós representam clientes e uma aresta direcionada entre dois nós representa a troca de conteúdo entre clientes, ou seja, eles estabelecem uma parceria. Estas parcerias são estabelecidas e desfeitas dinamicamente em resposta tanto ao comportamento dinâmico dos pares quanto ao desempenho (qualidade do vídeo recebido) observado por um par.

Embora existam na literatura alguns trabalhos que tenham caracterizados diferentes aspectos de aplicações de transmissão de vídeo ao vivo baseada em redes P2P [Hei et al. 2007, Borges et al. 2010b, Huang et al. 2008], incluindo alguns estudos sobre as propriedades da rede sobreposta criada em tal aplicação [Silverston and Fourmaux 2007, Oliveira et al. 2009], nenhum deles caracterizou as propriedades dinâmicas desta rede. Dentre alguns aspectos que já foram analisados destacam-se o número e a duração das parcerias estabelecidas por um par [Tang et al. 2009, Stutzbach et al. 2008]. Embora relevantes, estes aspectos fornecem apenas uma visão parcial sobre como a rede de parcerias evolui ao longo do tempo. Uma outra questão em aberto é o quão dinâmica é a estrutura da rede ao longo do tempo de uma transmissão. Uma rede muito dinâmica pode ocasionar em problemas de imprevisibilidade, dificuldade de manutenção, alocação de recursos e entre outros.

O objetivo desse trabalho é prover uma caracterização inicial das propriedades dinâmicas da rede sobreposta de um sistema de transmissão P2P de vídeo ao vivo, isto é, de como as características dos nós e da rede evoluem com o tempo. Essa caracterização busca apresentar as características individuais dos nós e da topologia da rede. Para analisar os nós individualmente foram utilizadas métricas de centralidade que permitem entender o comportamento do nó dentro da estrutura da rede. São elas: Grau, *Betweenness* e *Closeness*. Já para a topologia da rede foram utilizadas as seguintes métricas para compreender a estrutura da rede no qual os nós estão inseridos: Diâmetro, Coeficiente de Agrupamento, Grau Máximo e Caminho Mínimo Médio.

A caracterização foi realizada com base nos resultados de experimentos que utilizaram entre 200 e 465 clientes SopCast executados no PlanetLab⁴. Estes clientes foram conectados a um canal de grande popularidade na China, o CCTV-1, e monitorados durante 40 minutos em um horário de pico local (20 horas). Os resultados encontra-

¹www.sopcast.org

²www.pptv.com

³www.ppstream.com

⁴www.planet-lab.org

dos mostram que os nós podem ser agrupados em três perfis com valores de métricas de centralidade distintos na rede: *Centralidade Alta* (CA), *Centralidade Intermediária* (CI) e *Centralidade Baixa* (CB). O perfil CA contém nós mais centrais dentro da estrutura da rede com medidas altas para as métricas avaliadas. Já os nós do perfil CB são mais periféricos na visão da rede sobreposta coletada com baixos valores de centralidade. Sendo os nós do perfil CI aqueles com valores intermediários entre o CA e o CB. Apesar de terem sido encontrados três perfis de nós diferentes, os nós apresentam pouca dinâmica entre os perfis. O perfil individual do nó é definido pela estrutura da rede em que ele se conecta. Além disso, os nós tendem a manter em média 70% das parcerias estabelecidas ao longo do tempo. As características estruturais da rede não apresentam variações significativas ao longo do tempo em relação ao Diâmetro e Caminho Mínimo Médio. Os resultados dessa caracterização podem ser explorados por pesquisas futuras em simulações mais realistas com estratégias otimizadas que representem a real dinâmica da rede.

O restante desse trabalho está organizado da seguinte forma: a Seção 2 discute os trabalhos relacionados da literatura e a Seção 3 apresenta o processo de execução dos experimentos. Já a Seção 4 traz um *overview* das métricas de centralidade dos nós e da estrutura da rede que foram utilizadas neste trabalho. Finalmente, a Seção 5 apresenta os resultados obtidos seguida pela Seção 6 com a conclusão desses resultados.

2. Trabalhos Relacionados

Sistemas P2P de transmissão de vídeo ao vivo já foram estudados sob diferentes pontos de vista ao longo dos últimos anos. Um dos principais objetivos é entender o comportamento da rede e de seus usuários a fim de melhorar o desempenho dos sistemas, propondo, por exemplo, novos algoritmos para o estabelecimento de uma rede sobreposta que diminua o atraso de difusão de conteúdo em uma transmissão.

Em Hei et. al. [Hei et al. 2007], os autores mostram que a aplicação PPLive atinge um bom desempenho usando a Internet para a transmissão de vídeo ao vivo. Além disso, eles caracterizam o comportamento do usuário, bem como as informações que são trocadas entre os nós. No entanto, o dinamismo da rede sobreposta não foi explorado. Em outro trabalho dos mesmos autores [Hei et al. 2008], foi utilizada uma metodologia para analisar o comportamento dos usuários. Esta análise baseou-se no desempenho do sistema e na qualidade percebida pelos usuários finais, sem considerar as características da estrutura topológica da rede sobreposta.

O trabalho de Siverston et. al. [Silverston and Fourmaux 2007] analisa as características do tráfego gerado pelas aplicações SopCast, PPLive, TVAnts e PPStream, utilizando um conjunto de dados coletados durante a Copa do Mundo de 2006. Apesar de analisarem algumas métricas, como a taxa de *upload* e *download*, bem como o tamanho médio dos pacotes trocados, os autores não abordaram as características da rede sobreposta formada para a transmissão de conteúdo e como elas evoluem com o tempo.

Em [Stutzbach et al. 2008], Stutzbach et. al. propõem uma metodologia para analisar a evolução da rede sobreposta em uma aplicação de compartilhamento de arquivos em redes P2P. Porém, neste trabalho é aplicada a metodologia em redes P2P de transmissão de vídeo ao vivo. A contribuição deste trabalho é diferente da apresentada em [Stutzbach et al. 2008] dado que a transmissão ao vivo de vídeo em aplicações

P2P têm características diferentes, como prazo de entrega de pacotes, que influenciam na metodologia de caracterização.

A evolução das propriedades de redes também foi estudada por Watts [Watts and Strogatz 1998] em que ele apresenta o fenômeno do mundo pequeno como um dos resultados de experimentos em redes dinâmicas. Para realizar tal análise Watts usou métricas de redes como o diâmetro, grau dos nós e coeficiente de agrupamento. O livro recente de Easley et. al. [Easley and Kleinberg 2010] também apresenta vários trabalhos que buscam compreender a dinamicidade do crescimento e da redução de clientes em redes populares. No entanto, não é abordado como é a evolução da estrutura de uma rede P2P para a transmissão de vídeo ao vivo.

O trabalho que está sendo apresentado aqui segue o trabalho de Borges et al. [Borges et al. 2010b], onde os autores estudaram o comportamento dos usuários na aplicação SopCast. Os resultados obtidos dos autores foram estendidos neste trabalho, porém novos dados e novos resultados foram gerados sobre a perspectiva da estrutura da rede sobreposta e seu dinamismo. A rede sobreposta do SopCast é estudada pelo autores Tang et. al. [Tang et al. 2009] com *logs* de clientes SopCast executados e monitorados no PlanetLab. Os autores mostram a correlação entre o grau de saída e a média de *upload* em fotografias da rede. Eles mostram a relação entre pacotes de controle e de dado transmitidos entre os parceiros. Os dados utilizados são coletados de 51 nós do PlanetLab com duração de 40 minutos em apenas um dia.

Em comparação com a caracterização de Tang et. al. [Tang et al. 2009] e as outras de sistemas P2P anteriores [Silverston and Fourmaux 2007, Hei et al. 2007, Huang et al. 2008, Tang et al. 2009, Borges et al. 2010b, Stutzbach et al. 2008], neste trabalho foi utilizado um número maior de clientes SopCast (trabalhos anteriores como [Fallica et al. 2008] utilizaram no máximo 70 clientes) na tentativa de se obter uma representatividade maior da rede sobreposta do SopCast. Além disso, este trabalho apresenta uma análise diferente da rede que busca representar o dinamismo das propriedades sob dois pontos de vista: os nós com diferentes perfis de centralidade e a estrutura da rede.

3. Base de Dados

A base de dados utilizada neste trabalho foi obtida através de experimentos realizados no PlanetLab com clientes da aplicação SopCast conectados a um canal aberto, o CCTV-1. Na Seção 3.1 é apresentada uma breve descrição sobre a aplicação SopCast e seu funcionamento. Na Seção 3.2 apresentamos a descrição de como os experimentos foram feitos. Os dados utilizados estão sumarizados na Seção 3.3.

3.1. O Sistema SopCast

O SopCast é um das aplicações mais populares de transmissão de vídeo ao vivo que utiliza redes P2P, apresentando um maior volume de tráfego em relação as aplicações PPLive e PPStream em 2010 [Google 2010]. Para que ocorra transmissões de vídeo na rede P2P do SopCast são utilizadas redes sobrepostas. Em cada rede sobreposta ocorre a transmissão de um determinado vídeo, ou seja, trata-se de uma rede lógica formada exclusivamente para a transmissão de conteúdos específicos. A rede sobreposta de transmissão de um vídeo está relacionada na aplicação por *canal* em que o cliente está conectado.

Os clientes SopCast ao usarem a aplicação se conectam a um determinado canal, passando a fazer parte da rede sobreposta que está sendo transmitindo o vídeo. Os canais do SopCast podem ser abertos ou fechados. Um canal aberto é aquele que ele está disponível na aplicação para que qualquer cliente se conecte a rede sobreposta de transmissão. Por outro lado, o canal fechado é aquele que para o cliente ter acesso à rede que está sendo transmitido o vídeo é necessário que ele tenha permissão. Neste trabalho os clientes SopCast foram conectados a um dos mais populares canais abertos da China, o CCTV-1. A China é um dos países com o maior número de clientes SopCast [Google 2010].

As redes sobrepostas para a transmissão do conteúdo de um canal SopCast é composta por um servidor de transmissão ao vivo, servidor de *boot* (*bootstrap*) e clientes(nós) da rede. Neste contexto, o servidor de transmissão ao vivo é a fonte inicial do conteúdo e participa da rede sobreposta como um cliente especial do canal que ele está transmitindo. Já o servidor de *boot* é responsável por manter de forma centralizada o registro dos clientes por rede sobreposta. Quando um novo cliente se conecta a um determinado canal, o servidor de *boot* envia uma lista de parceiros candidatos para esse cliente. Os parceiros são clientes que já estão na rede sobreposta do canal desejado e que podem estabelecer uma parceria para a troca de dados. Uma parceria é, portanto, definida pelo fato de um cliente receber dados de outro cliente. Estas parcerias podem ser estabelecidas e desfeitas dinamicamente de acordo com o comportamento dos clientes e de acordo com a qualidade do vídeo transmitido entre parceiros.

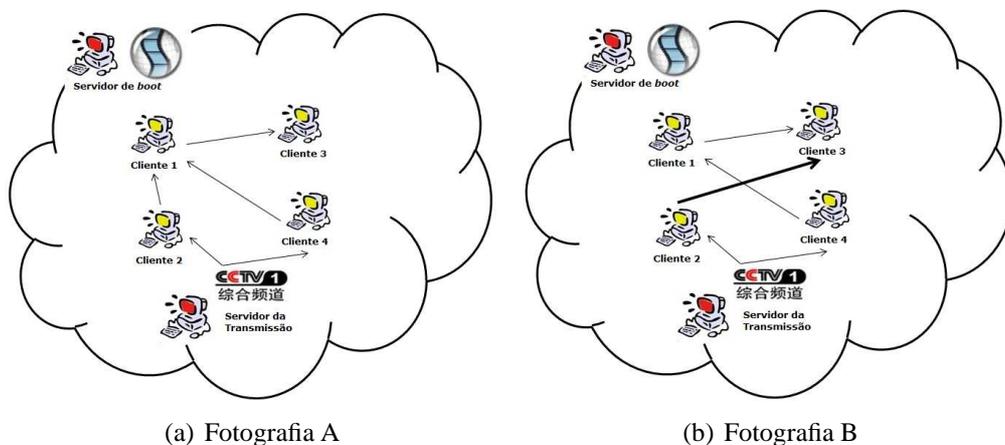


Figure 1. Representação da rede sobreposta do SopCast em duas fotografias

Para representar a dinamicidade das parcerias, a Figura 1 mostra duas fotografias consecutivas (*A* ocorre antes de *B*) de uma rede sobreposta hipotética formada para a transmissão do canal CCTV-1 no SopCast. Na Figura 1 é possível observar que no momento *B* o *Cliente 1* não possui mais parceria com o *Cliente 2* e que o *Cliente 2* estabelece uma nova parceria com o *Cliente 3* diferentemente do momento *A*. Entender a formação das propriedades dinâmicas da rede sobreposta e como suas parcerias evoluem com o tempo é o principal objetivo desta caracterização.

3.2. Metodologia para Realização dos Experimentos

Os resultados deste trabalho são baseados em um conjunto de E experimentos realizados com o SopCast. Cada experimento foi realizado utilizando um número N de computadores do PlanetLab. Cada computador foi configurado para executar um cliente SopCast

cada um e armazenar todas as trocas de dados realizadas ao longo de uma transmissão. Atuando como clientes SopCast eles se conectaram ao canal aberto da China, o CCTV-1, durante o horário de pico local (20 horas). O CCTV-1 é um canal muito popular da emissora estatal chinesa e possui uma qualidade de transmissão de vídeo (600 kbps) maior do que quase todos os outros canais do SopCast. A metodologia para a realização dos experimentos deste trabalho é composta basicamente por duas etapas: configuração dos computadores do PlanetLab e monitoramento dos mesmos durante a transmissão de vídeo no SopCast.

Durante a etapa de configuração dos computadores não foram impostas quaisquer restrições, sejam de armazenamento, taxa de *upload* e *download*. Todas as versões dos softwares utilizados foram atualizadas para a mesma versão em todos os computadores. Além do próprio SopCast, em particular, foi utilizado o Wireshark⁵ (*tcpdump*) para capturar o tráfego de rede observado durante o monitoramento de uma transmissão de vídeo.

O Wireshark foi configurado para capturar apenas o tráfego com porta de origem e destino iguais as portas configuradas para o sistema SopCast (por exemplo, porta UDP/TCP 3908). Essa captura é armazenada em um arquivo de *log* e contém a data e a hora (na granularidade de 1 segundo) de cada pacote enviado pelo próprio computador ou recebido por ele na porta configurada ao longo da transmissão de vídeo. Para que todos os computadores estivessem sincronizados nos experimentos foi utilizado o *Network Time Protocol*⁶ (NTP) com a configuração do mesmo servidor entre eles [Pathak et al. 2008]. Essa sincronização garante que a diferença de fuso horário e de tempo entre os computadores seja muito pequena (menos de um segundo).

Após a etapa de configuração de todos os computadores, eles passaram a executar o SopCast e a ferramenta de monitoramento Wireshark. De forma distribuída ao longo do tempo inicial T esses computadores, agora clientes SopCast, foram conectados ao canal CCTV-1. Como clientes comuns do SopCast, estes passaram a monitorar e registrar nos seus respectivos arquivos de *log* todas as parcerias realizadas, ou seja, todas as trocas de dados entre eles mesmos e entre outros clientes. A duração de um experimento é dada pela soma do tempo inicial T e a duração da transmissão monitorada D . Nas análises deste trabalho, foi considerado apenas os registros dos *logs* realizados após todos os clientes do SopCast do experimento já terem se conectado ao canal CCTV-1, ou seja, depois o tempo inicial T .

Finalizado o período D de monitoramento, todos os arquivos de *logs* dos clientes SopCast foram organizados em um único arquivo de *log* usado para reconstruir a dinâmica da rede sobreposta do SopCast durante a transmissão de vídeo. Baseando-se nas informações da data, da hora, do IP de origem e do IP de destino foram retiradas S fotografias (*snapshots*) da rede em intervalos de I segundos com janelas deslizantes de L segundos entre cada uma delas. Dado que a caracterização deste trabalho em relação as propriedades dos nós e da rede considera uma granularidade de tempo de um segundo, a diferença de tempo entre os dados dos Cliente SopCast pode ser considerada insignificante.

A rede deste trabalho foi reconstruída utilizando clientes executados no PlanetLab

⁵www.wireshark.org

⁶www.ntp.org

que foram configurados para armazenar todas as trocas de pacotes entre clientes SopCast, trata-se, portanto, de uma visão parcial da rede SopCast do canal aberto em que os clientes foram conectados. Todas as métricas analisadas neste trabalho são em função dessa visão da redesobreposta. A abordagem utilizada neste trabalho não é nova e muitas vezes é aplicada na literatura [Vu et al. 2007, Fallica et al. 2008]. No entanto, uma das principais contribuições deste trabalho está no número de clientes SopCast conectados na rede ao mesmo tempo em relação aos trabalhos da literatura que utilizaram no máximo 70 clientes [Fallica et al. 2008]. Além disso, essa abordagem busca por uma reconstrução da rede mais realista por apresentar uma maior probabilidade de uma visão completa da rede analisando a visão dos nós deste trabalho ao mesmo tempo em pontos diferentes do mundo.

3.3. Sumário dos Experimentos

Neste trabalho são apresentados 7 experimentos realizados entre os dias 28/10/2010 e 17/11/2010 com duração de transmissão de vídeos de 40 minutos em cada dia. Foram utilizados em cada experimento entre 200 a 465 clientes SopCast. Essa variação do número de clientes SopCast acontece devido a instabilidade do ambiente de testes, no entanto a diferença entre o número de participantes nestes experimentos é de aproximadamente 1%. Estes clientes foram conectados ao canal aberto da China (CCTV-1) no horário de pico local (20 horas) ao longo de um tempo inicial de 10 minutos. E então, monitoraram os tráfegos de rede após todos os clientes já estarem conectados na rede durante 40 minutos. Os *logs* de cada cliente SopCast foram agrupados por experimento e retirados 118 fotografias (*snapshots*) em que cada uma delas representa 60 segundos e a diferença entre elas de 20 segundos. Estes dados estão sumarizados na Tabela 1.

Table 1. *Sumário dos Experimentos Realizados*

Número de experimentos realizados (E)	7
Período dos experimentos	28/10/2010 a 17/11/2010
Número de computadores por experimento (N)	200 a 465
Canal aberto utilizado	CCTV-1
Horário de pico local	20:00 hrs
Tempo inicial de conexão (T)	10 minutos
Duração da transmissão (D)	40 minutos
Número de fotografias da rede por experimento (S)	118
Duração de cada fotografia (L)	60 segundos
Janelas deslizantes entre as fotografias (I)	20 segundos

4. Métricas Utilizadas

A rede sobreposta formada pelos clientes conectados no SopCast durante um experimento pode ser representada por um grafo direcionado $G = (V, E)$, onde V é o conjunto de $|V|$ vértices e E é o conjunto de arestas que conectam dois elementos de V . Uma aresta direcionada $(v_i, v_j) \in E$ representa uma parceria estabelecida entre dois clientes representados pelos nós v_i e v_j ($v_i, v_j \in V$) para a transmissão de dados de v_i para v_j . As características deste grafo podem variar ao longo do tempo à medida que as parcerias são feitas e desfeitas.

O objetivo deste trabalho é exatamente caracterizar como as características deste grafo evoluem com o tempo durante uma transmissão. Essa caracterização é feita sobre

duas perspectivas: características dos nós e características da rede como um todo. Para tanto, foram utilizadas métricas de nós (Seção 4.1) e métricas da rede (Seção 4.2). Por simplificação nas subseções seguintes uma fotografia da rede será abordada como instante t , a rede sobreposta desse instante como grafo G , os nós como vértices v e as parcerias estabelecidas por estes nós como arestas e .

4.1. Métricas dos Nós

Para analisar os nós dentro da estrutura da rede foram utilizadas métricas que avaliam a centralidade do nó nessa estrutura, são elas: grau, *betweenness* e *closeness*.

Grau: O grau g de um vértice $v_i(t)$ no instante t é definido pela soma do grau de entrada e saída de $v_i(t)$. O grau de entrada de v_i é o número de v para os quais ele aponta, enquanto o grau de saída é o número de v que apontam para v_i . [Easley and Kleinberg 2010]. O grau de um vértice é uma das propriedades estruturais mais básicas que avalia o número de arestas adjacentes.

Betweenness: O *betweenness* de um vértice v_i é a razão de $\sigma_{st}(v_i, t)$ por σ_{sw} , em que $\sigma_{sw}(v_i, t)$ representa o número de caminhos mínimos de v_s a v_w que passam por v_i e σ_{sw} o número de caminhos mínimos entre v_s e v_w . O *betweenness* de um vértice é a probabilidade dele estar no caminho mínimo entre dois outros vértices quaisquer.

$$Betweenness(v, t) = \sum_{s \neq w \neq i \in V} \frac{\sigma_{st}(v_i, t)}{\sigma_{sw}}$$

Closeness: O *closeness* de um vértice $v_i(t)$ no instante t é usado para avaliar o caminho mínimo médio, definido por l , entre o vértice v_i e todo vértice v_w alcançável a partir de v_i tal que $(v_i, v_w) \in V$. Com essa métrica é possível identificar o quão próximo um vértice está dos demais vértices da rede. Em outras palavras, no SopCast, se um nó com um valor alto de *closeness* possui alguma informação na rede sobreposta, essa informação será encontrada pelos demais nós rapidamente.

$$Closeness(v, t) = \frac{|V| - 1}{\sum_{i \neq w \in V} l(v_i, v_w)}$$

4.2. Métricas da Rede

A estrutura da rede sobreposta como um todo foi avaliada utilizando as seguintes métricas: caminho mínimo médio, diâmetro, coeficiente de agrupamento e grau máximo da rede.

Diâmetro: O diâmetro de um grafo é definido pela distância máxima entre quaisquer dois nós, considerando que não haja ciclo. Em outras palavras, o diâmetro de um grafo é a maior distância m definida pelo número de arestas entre dois vértices (v_i, v_j) , considerando todos os pares de vértices no grafo G do instante t . Com essa métrica é possível ter uma idéia da dispersão do grafo. Logo, ela pode ser utilizada para fazer avaliações aproximadas sobre latência, impactos na rede, entre outros.

$$Diâmetro(G, t) = \max_{v_i, v_j \in V} m(v_i, v_j)$$

Caminho Mínimo Médio: O caminho mínimo médio de um grafo G no instante t é definido pelo menor caminho l de um vértice v_i para todos os outros vértices da rede utilizando busca em largura para calcular l .

$$CaminhoMínimoMédio(G, t) = \frac{\sum_{i \neq w \in V} l(v_i, v_w)}{|V|}$$

Coeficiente de Agrupamento: O coeficiente de agrupamento de um grafo G no instante t é dado pela média do coeficiente de agrupamento de todos os vértices Cv . Dados três vértices v_i, v_j e v_w com arestas definidas entre $(v_i, v_j) \in E$ e $(v_j, v_w) \in E$, o coeficiente de agrupamento de um vértice Cv é definido pela probabilidade de existir uma aresta entre $(v_i, v_w) \in E$.

$$CoeeficienteDeAgrupamento(G, t) = \frac{\sum_{v \in V} Cv(v, t)}{|V|}$$

Grau Máximo: Por definição o grau máximo de um grafo G no instante t é dado pelo maior grau de um vértice $g(v)$ é dado por:

$$GrauMáximo(G, t) = \max_{v_i \in V} g(v_i)$$

5. Resultados

Para cada experimento realizado foram caracterizadas as propriedades da rede e as centralidades individuais dos nós na estrutura desta rede. Na Seção 5.1 são apresentados os resultados obtidos pela caracterização dos nós enquanto que os resultados das propriedades estruturais da rede estão na Seção 5.2.

5.1. Caracterização dos Nós

A caracterização dos nós decorrentes de suas respectivas centralidades foi feita extraindo as medidas de Grau(g), *Betweenness*(b) e *Closeness*(c) de cada nó por fotografia da rede. Com isso em cada fotografia de um experimento, um nó v_i possui um vetor $v^{vi} = (g, b, c)$ com três posições correspondentes aos valores medidos da centralidade dele. Para identificar perfis de nós com centralidades distintas, os vetores de todos os nós de um experimento foram submetidos a um algoritmo de agrupamento, o *k-means* [Wan et al. 1988]. Este algoritmo atua nos dados de modo não supervisionado utilizando a distância Euclidiana como medida de distância entre dois vetores v^{vi} e v^{vj} . O número de perfis em cada experimento foi definido através da relação β_{cv} que avalia a variação da separação dos vetores em perfis. Esta avaliação é feita através da razão da distância média dos vetores dos perfis e seus centróides pela distância média entre os próprios centróides [Menasce and Almeida 2007].

Utilizando esta estratégia foram encontrados em todos os experimentos três perfis de centralidade dos nós. Os perfis foram denominados de *Centralidade Alta* (CA), *Centralidade Intermediária* (CI) e *Centralidade Baixa* (CB). A Tabela 2 sumariza os valores médios e os coeficientes de variação (CV) das medidas de centralidade dos nós calculados sobre todas as fotografias de cada experimento realizado. Note que, apesar de encontrados sempre três perfis (centróides), os valores exatos que definem cada centróide variam de acordo com o experimento.

Um nó do perfil CA é aquele que possui Grau e *Betweenness* mais elevado do que os demais, em contrapartida o *Closeness* entre os perfis não apresenta uma variação significativa. Os nós do perfil CA são aqueles que dentro da estrutura topológica da rede possuem maior centralidade. Já os nós do perfil CB são aqueles que apresentam baixos valores de Grau e *Betweenness*, podendo ser considerados nós com menor centralidade. Já os nós do CI são aqueles com características intermediárias ao CA e ao CB.

Table 2. Definição dos perfis dos nós

Experimento			% de Nós	Grau		Betweenness		Closeness	
				Média	CV	Média	CV	Média	CV
				1	CA	4.76%	282.83	0.17	3312.52
	CI	32.69%	257.99	0.20	1212.52	0.30	0.008	1.20	
	CB	62.53%	86.94	0.82	129.45	1.60	0.005	1.16	
2	CA	6.98%	334.89	0.21	6585.63	0.33	0.005	0.61	
	CI	45.85%	224.13	0.17	2258.33	0.32	0.006	0.64	
	CB	47.16%	36.10	1.22	58.73	3.22	0.003	0.88	
3	CA	3.81%	361.95	0.18	8604.33	0.30	0.003	0.52	
	CI	17.55%	240.43	0.21	2556.77	0.36	0.004	1.06	
	CB	78.62%	56.83	1.14	135.47	2.30	0.003	0.90	
4	CA	2.92%	298.33	0.24	11430.93	0.43	0.004	0.56	
	CI	14.61%	230.04	0.20	3443.58	0.38	0.007	0.82	
	CB	82.46%	56.83	1.14	157.44	2.64	0.004	0.97	
5	CA	6.76%	322.39	0.24	6898.68	0.56	0.004	0.79	
	CI	40.97%	230.42	0.19	2134.54	0.36	0.006	0.89	
	CB	52.25%	56.83	1.14	120.73	2.24	0.003	1.04	
6	CA	2.69%	241.23	0.20	11865.11	0.63	0.004	0.12	
	CI	8.99%	252.241	0.24	3159.50	0.42	0.007	0.75	
	CB	88.31%	86.20	1.04	282.39	1.61	0.007	1.08	
7	CA	19.79%	298.42	0.16	1751.81	0.31	0.006	1.42	
	CI	30.03%	261.86	0.16	833.93	0.22	0.007	1.37	
	CB	50.17%	123.47	0.68	159.52	1.13	0.005	1.61	

As diferenças entre os perfis de centralidade são apresentadas na Figura 2 pelas distribuições acumuladas (CDFs) encontradas para o Experimento 1 (vide Tabela 2), escolhido apenas para exemplificação. Nesta figura é possível observar que nós do perfil CA apresentam uma maior probabilidade de terem valores de Grau maior do que os demais perfis. Cerca de 20% dos nós do perfil CB possuem Grau aproximadamente igual a 0. Para a métrica *Betweenness*, as três curvas seguem também apresentam uma clara diferença entre os perfis, ou seja, a distribuição para CA tem uma tendência maior para valores mais altos seguido por CI e então por CB. Já o *Closeness* não segue o mesmo comportamento apresentado variações de valores entre os perfis o que não permite uma diferenciação clara entre os perfis.

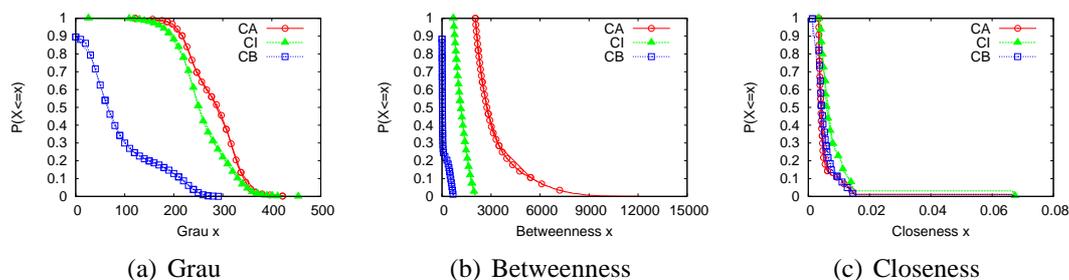


Figure 2. Métricas dos nós por perfil

Ao longo de uma transmissão, a medida que as parcerias entre os nós se alteram,

um mesmo nó pode apresentar diferentes perfis. Ou seja, em um dado momento, o nó pode ter uma centralidade mais alta e mais adiante, a medida que suas parcerias se alteram, ele pode ter uma centralidade mais baixa. Por exemplo: um nó que no início de uma transmissão é caracterizado pelo perfil *CA*, no decorrer do tempo, suas parcerias se alteram e ocorre a transição para perfil *CI*. O interesse deste trabalho é caracterizar este comportamento dinâmico. Para tal, foi utilizado uma técnica de modelagem denominada *Customer Behavior Model Graph* (CBMG) [Menasce and Almeida 2007]. Um CMBG é um grafo utilizado para representar a dinamicidade em um modelo de comportamento. No contexto deste trabalho, o uso de um CMBG tem por objetivo mapear as alterações das centralidades dos nós ao longo do tempo. Para isso, no CMBG cada estado (retângulo) representa um perfil de centralidade do nó e as arestas entre os estados representam a probabilidade de um nó ter sua centralidade alterada de um perfil para o outro. A Figura 3 (a) apresenta o CMBG encontrado para o Experimento 1 (Tabela 2) em que o mesmo padrão de transição entre os perfis pode ser observado para os outros experimentos. A Figura 3 (b) mostra o CMBG médio de todos os sete experimentos realizados.

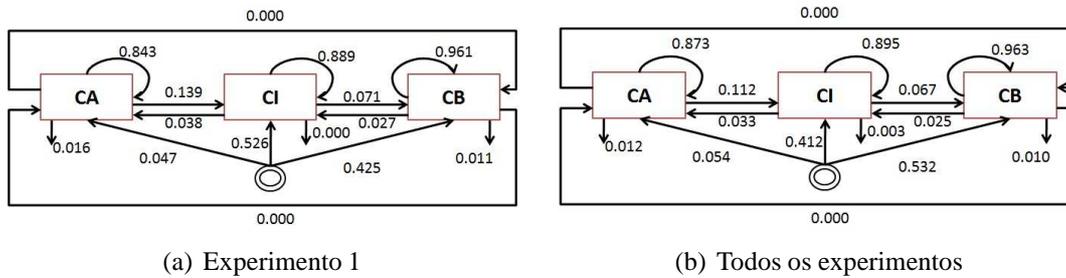


Figure 3. CMBG

Nos dois CMBGs da Figura 3 é possível visualizar a transição dos nós entre os perfis de centralidade observados nos experimentos. A probabilidade de um nó permanecer no mesmo perfil tende a ser bem alta. Por exemplo, considerando o CMBG médio (Figura 3 (b)) um nó que inicialmente possuía centralidade baixa|alta tende a permanecer no perfil *CB|CA* com probabilidade de 0.963|0.873. As transições entre dois perfis extremos como o *CA* e *CB* é zero dentro da precisão de três casas adotada neste trabalho, indicando que não há uma redução brusca de conectividade de um nó. Outro fato que pode ser observado é que a probabilidade de um nó diminuir sua centralidade é maior que a probabilidade dele aumentar, visto que de *CA* para *CI* a probabilidade é de 0.112, *CI* para *CB* de 0.067 enquanto que de *CB* para *CI* é 0.025 e *CI* para *CA* 0.033. Logo, os nós ao se conectarem em uma rede sobreposta possuem maior probabilidade de permanecerem no mesmo perfil.

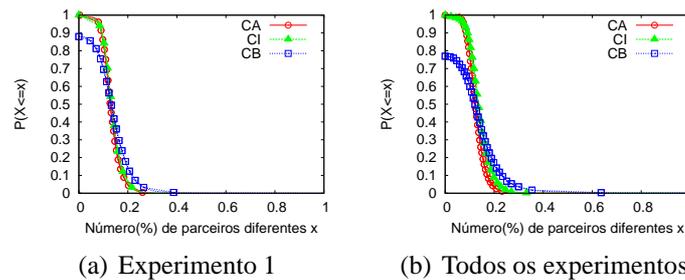
Sendo assim, outra questão que pode ser considerada é que apesar de tenderem a manter o mesmo perfil de centralidade, os nós tendem a variar seus parceiros ao longo do tempo. Note que isto pode acontecer sem que as propriedades dos nós sejam alteradas. A Figura 4 (a) apresenta para cada perfil a CDF da porcentagem p de parceiros diferentes de um mesmo nó i em duas fotografias consecutivas $t1$ e $t2$ da rede. Em outras palavras:

$$p_i = \frac{p_i^{t1} \cap p_i^{t2}}{p_i^{t1} \cup p_i^{t2}}$$

Table 3. Medidas das métricas da rede

		Diâmetro		Caminho Mínimo Médio		Coeficiente de Agrupamento		Grau Máximo	
		Média	CV	Média	CV	Média	CV	Média	CV
Experimento	1	4.11	0.07	1.98	0.03	0.24	0.34	361.47	0.08
	2	4.13	0.09	2.08	0.03	0.15	0.28	431.51	0.11
	3	4.33	0.13	2.09	0.02	0.15	0.46	429.71	0.06
	4	4.27	0.10	2.16	0.03	0.12	0.56	389.88	0.09
	5	4.22	0.09	2.14	0.02	0.17	0.21	398.68	0.12
	6	4.23	0.10	2.10	0.07	0.18	0.64	339.83	0.11
	7	4.05	0.05	2.00	0.02	0.32	0.11	377.00	0.05

A porcentagem de parceiros diferentes entre fotografias consecutivas tende a ser muito baixa nos três perfis. Em particular, para os perfis *CA* e *CI* a porcentagem máxima observada foi de 20%. A Figura 4 (b) mostra para cada perfil a CDF correspondem considerando todos os experimentos. Assim como o Experimento 1, a porcentagem de parceiros diferentes entre fotografias consecutivas para todos os experimentos também tende a ser baixa.

**Figure 4. Probabilidade de novos parceiros por perfil de nó**

Portanto, os nós possuem poucos parceiros diferentes entre fotografias consecutivas da rede e além disso tendem a se manter em um mesmo perfil. Os perfis foram encontrados pelas métricas de Grau e *Betweenness*. O *Closeness* apesar de ser uma métrica do nó não apresentou uma variação significativa entre os perfis. A definição de cada perfil foi feita de acordo com o experimento, foi possível identificar que em todos os experimentos foram encontrados três perfis distintos de centralidade.

5.2. Caracterização da Rede Sobreposta

Sobre a perspectiva das propriedades estruturais da rede sobreposta, as métricas de rede foram extraídas de cada fotografia dos experimentos. Foi feita uma análise primeiramente do valor médio medido em todas as fotografias de cada experimento. A Tabela 3 sumariza as medidas encontradas por experimento com as médias e os coeficientes de variação (CV) dos valores encontrados.

Em todos os experimentos realizados os valores do Diâmetro e Caminho Mínimo Médio apresentam valores em torno de 4 e 2 respectivamente. Os Graus Máximos dos experimentos apresentam valores médio variando entre 360 e 430 e Coeficientes de Agrupamento entre 0.12 e 0.32. Para entender essa variação do Grau Máximo e do Coeficiente de Agrupamento da rede foi estudada como ocorre a evolução desses valores ao longo do

tempo, ou seja, nas fotografias de um experimento. A Figura 5 mostra as quatro métricas da rede ao longo do tempo do Experimento 1. Nesta análise é possível observar que o Coeficiente de Agrupamento da rede decresce ao longo do tempo, Figura 5(d). Sendo que o número de nós analisados é o mesmo por todas as fotografias, o coeficiente de agrupamento decresce devido a novas parcerias estabelecidas entre os nós ao longo de todo o tempo.

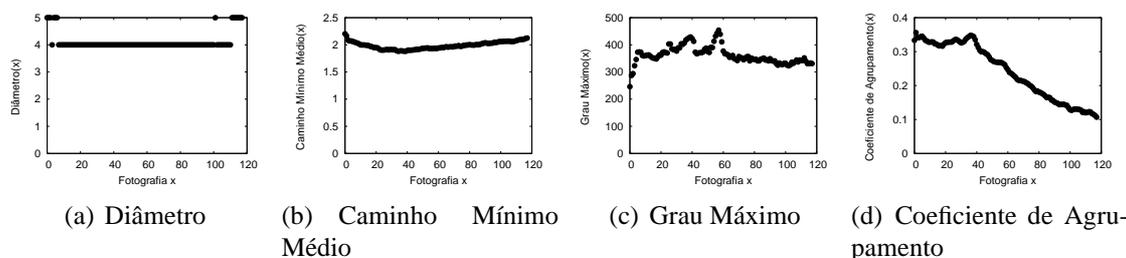


Figure 5. Métricas de rede ao longo do tempo do Experimento 1

Assim, as métricas da rede Diâmetro (Figura 5(a)) e Caminho Mínimo Médio (Figura 5(b)) tendem a permanecer estáveis. Já o Coeficiente de Agrupamento decresce ao longo do tempo e o Grau Máximo (Figura 5(c)) possui variações porém dentro do Grau observado pelo perfil CA dos nós. Não foi encontrada uma relação direta dos perfis dos nós com as métricas da rede, exceto pela comportamento estável dos nós nos perfis e pelas métricas Diâmetro e Caminho Mínimo Médio.

6. Conclusões e Trabalhos Futuros

Neste trabalho foi apresentado uma caracterização das propriedades dinâmicas da estrutura de redes sobrepostas do SopCast. Foram realizados uma série de sete experimentos utilizando computadores do PlanetLab que se conectaram ao canal aberto CCTV-1 do SopCast e passaram a atuar na rede de transmissão do canal como clientes comuns. Estes computadores monitoraram o tráfego de rede realizado para a transmissão do vídeo e armazenaram as informações em arquivos de *log*. Estes arquivos foram unificados em um único arquivo que pode ser utilizado para representar uma visão da rede sobreposta para a transmissão de vídeo. Analisando os nós individualmente em cada experimento, foram retiradas as métricas de centralidade dos clientes SopCast.

Os valores destas métricas de cada experimento foram submetidos ao algoritmo de agrupamento *k-means*, onde foi possível identificar três perfis distintos de centralidade: *Centralidade Alta*, *Centralidade Intermediária* e *Centralidade Baixa*. Além de gerar as distribuições acumuladas dos valores obtidos pelas métricas por perfil, foram gerados CBMGs (*Customer Behavior Model Graph*) para representar a transição de um nó com um dado perfil de centralidade para outro perfil ao longo do tempo. Com o CBMG foi possível concluir que os nós possuem baixa probabilidade de transição entre os perfis de centralidade. Além disso, foi analisado entre fotografias consecutivas de um experimento a porcentagem de parceiros diferentes realizados pelos nós. Já na perspectiva da estrutura da rede foi possível observar que o Diâmetro e o Caminho Mínimo Médio da rede tendem a serem mais estáveis que o Grau Máximo e o Coeficiente de Agrupamento.

Com os resultados obtidos neste trabalho foi possível identificar *insights* da dinamicidade da estrutura da rede sobreposta de um sistema de transmissão de vídeo ao

vivo. Estes *insights* podem ser utilizados em pesquisas futuras e na simulação de sistemas para que cargas de trabalho sintéticas ao serem geradas reflitam o real dinamismo da estrutura da rede sobreposta para a transmissão de vídeo em redes P2P.

References

- Borges, A., Almeida, J., and Campos, S. (2010a). Can malicious attacks slow peer-to-peer live iptv systems adoption? *Submetido para periódico*.
- Borges, A., Gomes, P., Nacif, J., Almeida, J. M., and Campos, S. (2010b). Characterizing sopcast client behavior. *Submetido para periódico*.
- Easley, D. and Kleinberg, J. (2010). *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, New York, NY, USA.
- Fallica, B., Lu, Y., Kuipers, F., Kooij, R., and Mieghem, P. V. (2008). On the quality of experience of sopcast. *International Conference on Next Generation Mobile Applications, Services and Technologies*.
- Google (2010). Google trends.
- Hei, X., Liang, C., Liang, J., Liu, Y., and Ross, K. (2007). A measurement study of a large-scale p2p iptv system. *IEEE Transactions on Multimedia*.
- Hei, X., Liu, Y., and Ross, K. (2008). Iptv over p2p streaming networks: the mesh-pull approach. *IEEE Communications Magazine*.
- Huang, Y., Fu, T. Z., Chiu, D.-M., Lui, J. C., and Huang, C. (2008). Challenges, design and analysis of a large-scale p2p-vod system. *SIGCOMM Comput. Commun. Rev.*
- Menasce, D. and Almeida, V. (2007). *Scaling for E-Business: Technologies, Models, Performance, and Capacity Planning*. Prentice Hall.
- Oliveira, J. F. A., de C. Gomes, P., Vieira, A. B., and Campos, S. V. A. (2009). Centralidade em redes p2p de transmissão ao vivo. *SBRC-09*.
- Pathak, A., Pucha, H., Zhang, Y., Hu, Y. C., and Mao, Z. M. (2008). A measurement study of internet delay asymmetry. In *PAM'08*, Berlin, Heidelberg.
- Silverston, T. and Fourmaux, O. (2007). Measuring p2p iptv systems. *NOSSDAV'07: Proceedings of Network Operating Systems Support for Digital Audio Video*.
- Stutzbach, D., Rejaie, R., and Sen, S. (2008). Characterizing unstructured overlay topologies in modern p2p file-sharing systems. *IEEE/ACM Trans. Netw.*
- Tang, S., Lu, Y., Hernández, J. M., Kuipers, F., and Mieghem, P. (2009). Topology dynamics in a p2ptv network. In *NETWORKING '09*, Berlin, Heidelberg.
- Vu, L. H., Gupta, I., Liang, J., and Nahrstedt, K. (2007). Measurement and modeling of a large-scale overlay for multimedia streaming. In *QShine International Conference*, Vancouver, Canada.
- Wan, S. J., Wong, S. K. M., and Prusinkiewicz, P. (1988). An algorithm for multidimensional data clustering. *ACM Trans. Math. Softw.*
- Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*.