

**CARACTERIZAÇÃO DAS PROPRIEDADES
DINÂMICAS DA REDE SOBREPOSTA EM UMA
APLICAÇÃO DE TRANSMISSÃO PAR-A-PAR DE
VÍDEO AO VIVO**

KÊNIA CAROLINA GONÇALVES

CARACTERIZAÇÃO DAS PROPRIEDADES
DINÂMICAS DA REDE SOBREPOSTA EM UMA
APLICAÇÃO DE TRANSMISSÃO PAR-A-PAR DE
VÍDEO AO VIVO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

ORIENTADORA: JUSSARA DE ALMEIDA
CO-ORIENTADOR: ALEX BORGES VIEIRA

Belo Horizonte
31 de julho de 2012

© 2012, Kênia Carolina Gonçalves.
Todos os direitos reservados.

G635c Gonçalves, Kênia Carolina
Caracterização das Propriedades Dinâmicas da Rede
Sobreposta em uma Aplicação de Transmissão
Par-a-Par de Vídeo ao Vivo / Kênia Carolina
Gonçalves. — Belo Horizonte, 2012
xx, 66 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal de
Minas Gerais

Orientadora: Jussara de Almeida

Co-orientador: Alex Borges Vieira

1. Computação - Teses. 2. Videodigital - Teses.
3. Sistemas de transmissão de dados - Teses.
I. Orientadora. II. Coorientador. III. Título.

CDU 519.6*84(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

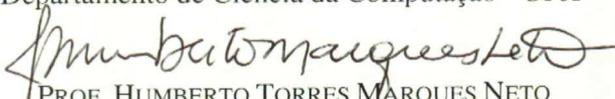
Caracterização das propriedades dinâmicas da rede sobreposta em uma aplicação de transmissão par-a-par de vídeo ao vivo

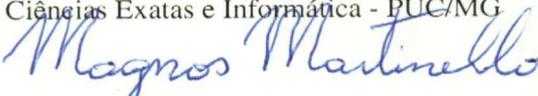
KÊNIA CAROLINA GONÇALVES

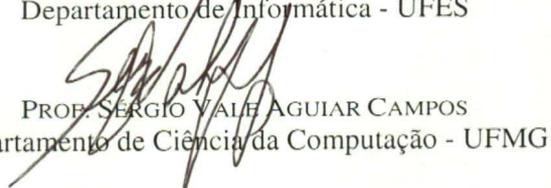
Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:


PROFA. JUSSARA MARQUES DE ALMEIDA GONÇALVES - Orientadora
Departamento de Ciência da Computação - UFMG


PROF. ALEX BORGES VIEIRA - Co-orientador
Departamento de Ciência da Computação - UFJF


PROF. HUMBERTO TORRES MARQUES NETO
Instituto de Ciências Exatas e Informática - PUC/MG


PROF. MAGNOS MARTINELLO
Departamento de Informática - UFES


PROF. SÉRGIO VALE AGUIAR CAMPOS
Departamento de Ciência da Computação - UFMG

Belo Horizonte, 31 de julho de 2012.

Agradecimentos

Deus: Obrigada por permitir que eu sonhe e realize.

Jussara: Obrigada por me ajudar e pela paciência na realização deste trabalho. Eu não teria conseguido sem sua ajuda. Em especial, espero conseguir ter um pouco da competência que vejo em você.

Alex: Obrigada pelo apoio no desenvolvimento deste trabalho, pelas idéias e pelo ouvido “on-line” nos momentos de nervosismo.

Humberto: Obrigada pela credibilidade em mim. Você foi o responsável por eu acreditar e correr atrás deste mestrado na UFMG.

Cleber: Obrigada pela paciência nos dias de nervosismo e de euforia. Mais que isso, por dividir comigo sua sabedoria e sensatez contribuindo para meu crescimento profissional. Você me ajudou e me ensinou a tomar decisões com maior consciência e maturidade. Obrigada por acreditar em mim, por estar ao meu lado e por me apoiar!

Nilda e Maurilio: Obrigada, sem vocês eu certamente não estaria aqui hoje.

VoD: Agradeço a todos meus amigos de laboratório pelo carinho que tiveram comigo. Principalmente, pela ajuda e paciência durante o meu estágio em docência. Em especial, agradeço ao Glauber, Henrique, Fabiano, Flávio e João que sempre me auxiliaram nas minhas dúvidas.

Amigos: Agradeço pela ajuda e credibilidade com o meu trabalho. Em especial, à Andrea e Ionara por me ouvirem e me apoiarem sempre. Fabrícia, muito obrigada pela sua amizade. Ao Fabrício (Funcesi) pela paciência e flexibilidade.

Professores: Agradeço à professora Cristiane Neri da UFSJ/PucMinas que foi a pessoa que me apresentou o meio acadêmico e ao professor José Wilson do CEFET que me mostrou a área da computação ainda no curso técnico.

Amigos, Professores e Funcionários do DCC: Obrigada pelo apoio que encontrei sempre que precisei. Em especial, o Giovanni e a Glivia.

PPGCC e InWeb: Obrigada por contribuírem no financiamento das publicações deste trabalho.

“Imagination is more important than knowledge.

Knowledge is limited.

Imagination encircles the world.”

(Albert Einstein)

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. Nesta dissertação é apresentada uma caracterização das propriedades dinâmicas da estrutura de uma rede sobreposta formada por uma aplicação de transmissão de vídeo ao vivo. A caracterização foi realizada sob dois pontos de vista: características individuais dos nós (visão local) e características da rede como um todo (visão global). 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. 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.

Palavras-chave: Redes P2P, Caracterização.

Abstract

Peer-to-Peer (P2P) live video streaming systems are becoming increasingly popular. Nevertheless, in spite of various studies of client behavior aspects and system optimizations, the current knowledge about the dynamic properties of the system, particularly how the P2P overlay network changes over time during a live transmission, is still superficial. In this work, we provide a characterization of the dynamic properties of a popular P2P live streaming media application. We use complex network metrics to analyze how the structure of the network evolves over time from the perspective of individual nodes (local view) and of the whole network (global view). We find that peers may be clustered into three profiles based on their centrality properties in the network. Moreover, in spite of peers changing their partners over time, they tend to remain with the same centrality profile. Also, the global network structure tends to remain roughly stable over time, except for a decaying clustering coefficient. Our findings can be used to generate more realistic synthetic P2P workloads and to drive future system designs and simulations.

Keywords: P2P Networks, Characterization.

Lista de Figuras

2.1	Representação da rede sobreposta de sistemas P2P	10
2.2	Representação da rede sobreposta de um sistema de vídeo ao vivo em duas fotografias	12
3.1	SopCast - Sistema de Transmissão de Vídeo ao Vivo em Redes P2P	22
4.1	Algoritmo de Agrupamento <i>k-means</i>	29
4.2	Distribuição das métricas de centralidade por perfil (Experimento 1, Cenário C_1)	31
4.3	Exemplo de um CBMG	33
4.4	CBMG - <i>Customer Behavior Model Graph</i> (Cenário C_1)	33
4.5	Distribuição da diferença de parceiros de um nó em fotografias consecutivas (Cenário C_1)	35
4.6	Distribuição da diferença de parceiros de um nó em fotografias não-consecutivas (Experimento 1, Cenário C_1)	35
4.7	<i>Pagerank</i> dos nós por perfil de centralidade (Cenário C_1)	36
4.8	Porcentagem de variação do <i>pagerank</i> dos nós por perfil de centralidade entre fotografias (Experimento 1, Cenário C_1)	37
4.9	Volume de <i>download</i> e <i>upload</i> por perfil de centralidade (Cenário C_1)	38
4.10	Métricas de rede ao longo do tempo (Experimento 1, Cenário C_1)	41
A.1	Métricas dos nós por perfil de centralidade (Experimento 1, Cenário C_2)	49
A.2	CBMG - <i>Customer Behavior Model Graph</i> (Cenário C_2)	51
A.3	Distribuição da diferença de parceiros de um nó em fotografias consecutivas (Cenário C_2)	51
A.4	Distribuição da diferença de parceiros de um nó em fotografias não-consecutivas (Experimento1, Cenário C_2)	52
A.5	<i>Pagerank</i> dos nós por perfil de centralidade (Cenário C_2)	52
A.6	Métricas de rede ao longo do tempo (Experimento 1, Cenário C_2)	54

B.1	Métricas dos nós por perfil de centralidade (Experimento 1, Cenário C_3)	55
B.2	CBMG - <i>Customer Behavior Model Graph</i> (Cenário C_3)	57
B.3	Distribuição da diferença de parceiros de um nó em fotografias consecutivas (Cenário C_3)	57
B.4	Distribuição da diferença de parceiros de um nó em fotografias não-consecutivas (Experimento 1, Cenário C_3)	58
B.5	<i>Pagerank</i> dos nós por perfil de centralidade (Cenário C_3)	58
B.6	Métricas de rede ao longo do tempo (Experimento 1, Cenário C_3)	60

Lista de Tabelas

3.1	Sumário dos Experimentos Realizados	26
4.1	Perfil de centralidade dos nós em cada experimento (Cenário C_1)	31
4.2	<i>Pagerank</i> dos nós por perfil de centralidade em cada experimento do Cenário C_1	37
4.3	Medidas das métricas da rede (Cenário C_1)	40
A.1	Perfil de centralidade dos nós em cada experimento (Cenário C_2)	50
A.2	<i>Pagerank</i> dos nós por perfil de centralidade em cada experimento (Cenário C_2)	53
A.3	Medidas das métricas da rede (Cenário C_2)	53
B.1	Perfil de centralidade dos nós em cada experimento (Cenário C_3)	56
B.2	<i>Pagerank</i> dos nós por perfil de centralidade em cada experimento (Cenário C_3)	59
B.3	Medidas das métricas da rede (Cenário C_3)	59

Sumário

Agradecimentos	vii
Resumo	xi
Abstract	xiii
Lista de Figuras	xv
Lista de Tabelas	xvii
1 Introdução	1
1.1 Motivação	1
1.2 Objetivos	4
1.3 Contribuições	4
1.4 Organização	5
2 Fundamentos	7
2.1 Sistemas em Redes Par-a-Par	7
2.2 Caracterizações de Sistemas de Transmissão de Vídeo ao Vivo	12
2.3 Métricas de Redes Complexas	16
2.3.1 Métricas dos Nós	16
2.3.2 Métricas da Rede	18
3 Metodologia de Coleta de Dados	21
3.1 SopCast	21
3.2 Monitoramento e Coleta de Dados	23
3.3 Sumário dos Experimentos	25
4 Caracterização da Evolução da Rede Sobreposta do SopCast	27
4.1 Overview da Metodologia	27

4.2	Caracterização dos Nós	28
4.2.1	Perfis de Centralidade dos Nós	29
4.2.2	Mudanças da Centralidade dos Nós	32
4.2.3	Mudança na Lista de Parceiros	34
4.2.4	Pagerank	36
4.2.5	<i>Download</i> e <i>Upload</i>	38
4.2.6	Validação	38
4.3	Caracterização da Rede Sobreposta	39
4.3.1	Validação	41
4.4	Sumário dos Resultados	42
5	Conclusões e Trabalhos Futuros	45
	Apêndice A Resultados em Canal Fechado Com <i>Churn</i> - Cenário C_2	49
	Apêndice B Resultados em Canal Fechado Sem <i>Churn</i> - Cenário C_3	55
	Referências Bibliográficas	61

Capítulo 1

Introdução

Existem várias aplicações populares na Internet atualmente que têm como foco a distribuição de vídeos em larga escala. Algumas dessas aplicações são baseadas em arquiteturas descentralizadas do tipo Par-a-Par (P2P), na qual os clientes se organizam em uma rede sobreposta para disseminar o conteúdo. Visando prover um conhecimento importante sobre o funcionamento dessas aplicações, esta dissertação apresenta uma caracterização das propriedades dinâmicas da rede sobreposta em uma aplicação popular deste tipo. A próxima seção apresenta as principais motivações do trabalho desenvolvido. Os principais objetivos e contribuições desta dissertação são apresentados nas Seções 1.2 e 1.3, respectivamente. Por fim, a Seção 1.4 detalha a organização dessa dissertação.

1.1 Motivação

Vídeos podem ser transmitidos na Internet de diversas maneiras, duas delas são: sob demanda e ao vivo. A transmissão sob demanda de um vídeo é iniciada a partir de uma requisição para o servidor onde o vídeo está armazenado. O Youtube¹ é um dos sistemas mais populares para o compartilhamento e transmissão sob demanda de vídeos, ocupando o terceiro lugar no *ranking* dos sites mais acessados no mundo de acordo com o Alexa.com [Alexa.com, 2012], perdendo apenas para o Google² e para o Facebook³.

Já a transmissão de vídeo ao vivo ocorre em tempo real e de forma sincronizada entre todos os usuários. No modelo tradicional de transmissão pelo sistema televisivo,

¹<http://www.youtube.com>

²<http://www.google.com>

³<http://www.facebook.com>

os usuários precisam estar próximos à base de operações das redes de TVs para que consigam receber o conteúdo com qualidade [Xiao & Ye, 2008]. O uso da Internet não apenas flexibiliza este tipo de transmissão como também contribui para a disseminação da informação de maneira geral. Por exemplo, um chinês que mora no Brasil pode assistir a programação local da China apenas com o acesso à Internet. São exemplos de aplicações de transmissão de vídeo ao vivo pela Internet populares atualmente: SopCast⁴, PPLive⁵ e PPStream⁶. Pesquisas recentes mostram que em 2013 o número de usuários desse tipo de aplicação pode chegar a 83 milhões [Sentinelli et al., 2007].

As tecnologias utilizadas para a transmissão de vídeo ao vivo também vêm avançando ao longo dos últimos anos. O modelo tradicional de transmissão de vídeo baseado em uma arquitetura centralizada do tipo cliente-servidor impõe sérios desafios para a escalabilidade da aplicação à medida em que o número de clientes simultâneos aumenta. Isto ocorre porque, nessa arquitetura, o servidor tipicamente cria um fluxo de transmissão independente para cada cliente (transmissão *unicast*), gerando assim grandes demandas por recursos do servidor, particularmente largura de banda de rede, e da própria rede conectando servidor e clientes [Pourebrahimi et al., 2005]. Esses desafios são ainda maiores para aplicações de transmissão de vídeo ao vivo, pois existem restrições rígidas para o tempo de entrega dos pacotes durante a transmissão do vídeo para garantir que ela ocorra em tempo real [Borges et al., 2008].

Uma arquitetura alternativa para aplicações de transmissão de vídeo ao vivo é baseada em redes Par-a-Par (P2P). Nessas redes, os clientes (ou nós) atuam também como servidores repassando o conteúdo recebido para os outros clientes (seus pares). Assim, a demanda por recursos é compartilhada entre os nós, garantindo uma melhor escalabilidade do sistema. A transmissão de conteúdo em um sistema P2P é feita a partir do estabelecimento de uma rede sobreposta entre os clientes, ou seja, uma rede lógica de transmissão de dados, em que os nós desta rede (clientes) trocam dados entre si. A estrutura da rede sobreposta pode ser organizada em árvore ou em malha. Na organização em árvore cada cliente recebe dados de apenas um outro cliente da rede (seu pai na árvore). Já na organização em malha, cada cliente estabelece parcerias com outros clientes realizando pedidos explícitos de dados. 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. Várias aplicações populares atualmente, tais como SopCast e PPLive, têm uma arquitetura baseada em malha. De fato, vários trabalhos apontam que arquiteturas baseadas em

⁴<http://www.sopcast.org>

⁵<http://www.pptv.com>

⁶<http://www.pps.tv>

malha são mais robustas que as baseadas em árvore por serem mais resistentes à entrada e à saída dinâmica dos nós (*churn*) [Borges, 2010].

Conforme descrito, a rede sobreposta de uma aplicação Par-a-Par pode se alterar ao longo do tempo à medida em que as parcerias são feitas e desfeitas. Essas mudanças podem indiretamente afetar a qualidade do vídeo recebido pelos clientes, ocasionando, por exemplo, atrasos na entrega de conteúdo. Vários estudos sobre aplicações P2P de transmissão de vídeo ao vivo analisaram diferentes aspectos destes sistemas, incluindo aspectos relacionados ao comportamento dos usuários (p.ex.: duração de sessão) [Huang et al., 2008; Hei et al., 2007; Qiu et al., 2009; Borges et al., 2012], aos padrões de tráfego gerado (p.ex.: taxas de *upload* e *download*) [Silverston et al., 2007b,a, 2006, 2009; Hei et al., 2006; Bermudez et al., 2011] e à qualidade de serviço provido (p.ex.: atrasos no vídeo recebido) [Fallica et al., 2008; Liang et al., 2009]. Em comum, estes trabalhos analisam fotografias estáticas da rede que representam uma visão agregada do sistema durante toda a transmissão. Logo, eles não analisam como esses aspectos variam ao longo do tempo durante uma transmissão. Alguns poucos trabalhos analisaram o comportamento dessas aplicações *ao longo do tempo*. Porém, estes trabalhos focaram primordialmente em um único aspecto, a saber, o número de parcerias estabelecidas, ou seja, o grau dos nós na rede sobreposta [Stutzbach et al., 2008; Tang et al., 2009; Wu & Li, 2007].

Entretanto, existem outros aspectos também importantes. A centralidade de um nó, por exemplo, estimada por diversas métricas (incluindo grau) [Easley & Kleinberg, 2010] indica a posição do nó na rede e a sua conectividade [Freeman, 1979]. Logo, ela expressa, ainda que indiretamente, o potencial de comunicação do nó, o que influencia na qualidade de conteúdo transmitido e recebido. Mais amplamente, a estrutura da rede como um todo, que pode ser representada por várias métricas, tais como diâmetro e menor caminho [Watts, 1999; Watts & Strogatz, 1998], também pode impactar na qualidade do vídeo recebido e na escalabilidade da rede. Até onde sabemos, nenhum trabalho anterior analisou como as propriedades estruturais da rede sobreposta criada em aplicações P2P de transmissão de vídeo ao vivo evoluem durante a transmissão.

A caracterização dessas propriedades pode fornecer subsídios importantes para o projeto futuro de aplicações mais robustas. Mais ainda, os resultados de tal caracterização podem ser explorados por pesquisas futuras na simulação, em cenários mais realistas, de novos mecanismos e estratégias otimizadas. De fato, ela pode levar à identificação de fontes de sobrecarga no sistema com potencial impacto negativo no desempenho do mesmo. Ela pode também levantar indícios das implicações do uso de certos mecanismos comumente usados em aplicações P2P de transmissão ao vivo, tais como estratégias para seleção de parceiros e para requisição de dados de parceiros.

1.2 Objetivos

O principal objetivo desta dissertação é caracterizar como as propriedades estruturais da rede sobreposta formada por uma aplicação P2P de transmissão de um vídeo ao vivo evoluem com o tempo durante uma transmissão. Especificamente, pretende-se caracterizar essas propriedades sob duas perspectivas: visão local de cada nó e visão global da rede como um todo. Para a visão local de cada nó, tem-se por objetivo analisar primordialmente métricas que expressam a centralidade do nó na rede. Já para a visão global, pretende-se utilizar métricas de redes complexas para capturar as propriedades importantes da estrutura da rede.

1.3 Contribuições

As principais contribuições desta dissertação são [Gonçalves et al., 2012, 2011]:

- Foi realizada uma extensa caracterização da evolução temporal das propriedades estruturais da rede sobreposta criada pelo SopCast, uma das aplicações de transmissão ao vivo mais populares atualmente. Os experimentos foram realizados utilizando centenas de máquinas do PlanetLab [Peterson et al., 2006].
- Ao avaliar a centralidade dos nós ao longo do tempo (visão local), foi possível identificar três perfis de centralidade ao longo do tempo. Entretanto, observou-se que um nó tende a permanecer com o mesmo perfil de centralidade ao longo de uma transmissão, apresentando, portanto, pouco dinamismo na rede, no que tange a sua centralidade.
- Observando as parcerias dos nós ao longo de uma transmissão, foi possível identificar que, a despeito de um nó tender a manter seu perfil de centralidade, ele tende a mudar frequentemente de parceiros.
- A estrutura topológica da rede (visão global), assim como os nós individuais, tendem a apresentar uma estabilidade ao longo do tempo, principalmente no que tange o diâmetro, os caminhos mínimos médios, o coeficiente de agrupamento e o grau máximo da rede.
- Esses resultados foram validados em vários experimentos em um canal aberto bastante popular do SopCast assim como em experimentos com um canal fechado (com uma população de clientes controlada).

1.4 Organização

O restante desta dissertação está organizado como se segue. No Capítulo 2 são abordados os fundamentos teóricos e os trabalhos relacionados a essa dissertação. O Capítulo 3 apresenta a metodologia de coleta e a modelagem de dados. Os principais resultados da caracterização realizada são discutidos no Capítulo 4. Por fim, no Capítulo 5 são discutidas as conclusões obtidas bem como algumas direções para trabalhos futuros.

Capítulo 2

Fundamentos

Este capítulo apresenta os fundamentos teóricos principais sobre sistemas Par-a-Par, particularmente sistemas de transmissão de vídeo ao vivo (Seção 2.1), assim como os trabalhos relacionados em caracterizações de sistemas de transmissão de vídeo ao vivo (Seção 2.2). Ele ainda faz uma revisão das principais métricas de caracterização de propriedades estruturais de diferentes tipos de redes (Seção 2.3).

2.1 Sistemas em Redes Par-a-Par

No modelo de transmissão tradicional do tipo cliente-servidor, o servidor é responsável por prover todo serviço e conteúdo para os clientes. No entanto, este modelo traz grandes desafios de escalabilidade à medida em que o número de clientes simultâneos aumenta. Isto ocorre pois, tipicamente, o servidor cria um canal de comunicação independente para cada cliente para que seja feita a transmissão de conteúdo. Assim, o consumo de recursos de processamento e de largura de banda do servidor cresce linearmente com o número de clientes.

Um modelo de transmissão alternativo é baseado em redes Par-a-Par (P2P). Nestas redes, cada nó pode atuar tanto como cliente quanto como servidor. Os nós em redes P2P são capazes, por exemplo, de se auto-organizar, de se adaptar às populações transientes e de compartilhar recursos e serviços entre si. Os nós em uma rede P2P formam uma rede sobreposta (rede lógica) sobre a rede física para a troca de dados entre eles. Existem várias formas de se categorizar as redes P2P existentes. Elas podem ser, por exemplo, classificadas em relação à centralização da rede, à estrutura topológica da rede ou quanto ao mecanismo de busca utilizado.

Quando classificadas de acordo com a centralização, as redes P2P podem ser *puramente descentralizadas*, *parcialmente centralizadas* e *híbridas descentralizadas*

[Pourebrahimi et al., 2005]. Na arquitetura *puramente descentralizada*, os nós atuam como servidor/cliente sem um controle central. Portanto, eles se auto-organizam e possuem a mesma importância na rede. Essa arquitetura permite uma grande escalabilidade do sistema e tolerância a falhas, pois não existe nenhuma unidade centralizadora, e a entrada e saída (*churn*) de qualquer nó pode ser compensada por outro nó da rede. Em contrapartida, esta arquitetura não permite uma visão completa do sistema, acarretando problemas na descoberta de conteúdos e serviços.

Já a arquitetura *parcialmente centralizada* possui nós na rede com papéis diferentes. Por exemplo, alguns nós podem atuar como unidades centralizadoras da rede, facilitando a busca por conteúdo e serviços entre os demais nós da rede. Estes nós geralmente são chamados de “super nós” e são dinamicamente criados [Yang & Garcia-Molina, 2003]. Neste tipo de arquitetura, o tempo gasto na busca de informações é reduzido, porém a escalabilidade do sistema diminui.

Por fim, na arquitetura *híbrida descentralizada*, servidores centrais são responsáveis por manter as informações da rede (*metadados*) enquanto que a troca de conteúdo/serviço é feita diretamente entre os nós da rede. Esta é uma solução que busca resolver o problema de localização de recursos na rede. Entretanto, com a centralização das informações em alguns servidores, a rede fica mais susceptível a ataques e falhas, o que possivelmente diminui a escalabilidade da rede [Borges, 2010].

As redes P2P também podem ser classificadas quanto à sua organização topológica. Elas podem ser classificadas em redes *estruturadas* e *não estruturadas* [Pourebrahimi et al., 2005]. A rede é dita *estruturada* quando a topologia da rede é controlada e os conteúdos/serviços disponíveis estão em nós específicos. A topologia é controlada pois existe um mapeamento dos nós que estão na rede e dos recursos disponíveis em cada um deles. A vantagem desse tipo de rede é a facilidade de localização de recursos. Entretanto, o custo de manutenção do mapeamento das informações existentes na rede pode ser alto, uma vez que a entrada e saída dinâmica dos nós da rede exigem o remapeamento dos recursos disponíveis na rede [Stutzbach & Rejaie, 2006].

Na rede *não estruturada* a localização de recursos é desassociada da topologia da rede. Para se obter informações dos recursos da rede, os nós fazem pesquisas aleatórias (por exemplo, inundação de mensagens). O uso de mecanismos de busca na rede afeta o desempenho do sistema como um todo, uma vez que a quantidade de pesquisas aumenta com o número de nós na rede. A vantagem da rede *não estruturada* é a capacidade de adaptação à entrada e saída de nós, pois mecanismos de busca aplicados a este tipo de rede não exigem o mapeamento dos nós que estão na rede [Lua et al., 2005].

Por fim, os sistemas P2P podem ainda ser classificados quanto ao modelo de

busca por recursos/serviços utilizados. A busca pode ser feita, por exemplo, por *índices centralizados*, tipicamente em arquiteturas parcialmente centralizadas ou híbridas ou por *inundação de consultas* em redes descentralizadas ou não estruturadas. A busca pode ainda ser baseada em *roteamento de consultas* utilizando tabelas “hash” em redes estruturadas [Liu et al., 2008].

Redes P2P, que aplicam o mecanismo de busca baseado em *índices centralizados*, possuem um serviço centralizado com as localizações dos recursos disponíveis na rede. Quando um cliente deseja um determinado recurso, ele utiliza esse serviço central para encontrar a lista de clientes na rede que possuem este recurso. Os clientes passam, então, a realizar a troca de recursos entre si diretamente [Liu et al., 2008].

Já na busca por *inundação de consultas*, os nós não possuem qualquer serviço/servidor central para a localização de recursos. Cada cliente disponibiliza na rede informações sobre seus próprios recursos. Assim, quando um cliente deseja determinado recurso, ele envia consultas aos seus parceiros da rede na tentativa de localizar o recurso desejado. Estes parceiros por sua vez, enviam para os seus parceiros, e assim por diante até que um número máximo de reenvios de consultas seja atingido ou satisfaça uma condição de parada (por exemplo, a localização do recurso) [Liu et al., 2008].

Por fim, os mecanismos de busca baseados em *roteamento de consultas* são organizados utilizando, geralmente, tabelas “hash”. Quando um cliente busca por um determinado recurso, ele utiliza a tabela “hash” para obter um mapeamento entre o recurso desejado e a sua localização. Assim, é criada uma rota entre o cliente que busca pelo recurso e os clientes que dispõem do recurso [Borges, 2010].

Existem várias aplicações que utilizam redes P2P, em suas várias arquiteturas e mecanismos de busca, para a transmissão de conteúdo para os clientes da rede, incluindo aplicações de compartilhamento de arquivos tais como Gnutella¹ e BitTorrent², e aplicações de transmissão de vídeos, seja sob demanda (e.g., Tribler³) ou ao vivo. Exemplos de aplicações de transmissão de vídeo ao vivo baseadas em redes P2P bastante populares atualmente são SopCast, PPLive e PPStream. Essas aplicações tipicamente mantêm múltiplos canais, cada um transmitindo um conteúdo ao vivo. Além disso, uma rede P2P independente é criada para a transmissão de cada canal. Assim, um cliente interessado em receber o conteúdo ao vivo sendo transmitido em determinado canal deve participar da rede P2P (rede sobreposta) criada para aquela transmissão.

¹<http://gnutella.wego.com>

²<http://www.bittorrent.com>

³<http://www.tribler.org/trac>

A organização da rede sobreposta em aplicações de transmissão de vídeo ao vivo pode ser classificada em duas categorias: baseadas em *árvore* e em *malha*. Nas redes baseadas em *árvore*, cada nó repassa o conteúdo recebido para os nós filhos conforme mostrado na Figura 2.1a. Note que em uma rede baseada em *árvore*, cada cliente recebe o conteúdo de apenas um outro nó da rede (seu pai na árvore). A raiz da *árvore* é o servidor de transmissão de conteúdo. Assim, o sistema é susceptível a interrupções de transmissão e queda na qualidade de serviço quando ocorre a saída dinâmica de pares de rede, já que seus descendentes ficarão sem fonte de conteúdo até que a estrutura da *árvore* seja reconstruída. Em outras palavras, o *churn*, entrada e saída dinâmica de pares da rede, pode ter um impacto grande na qualidade de serviço de sistemas P2P de transmissão ao vivo baseados em *árvores* [Liu et al., 2008].

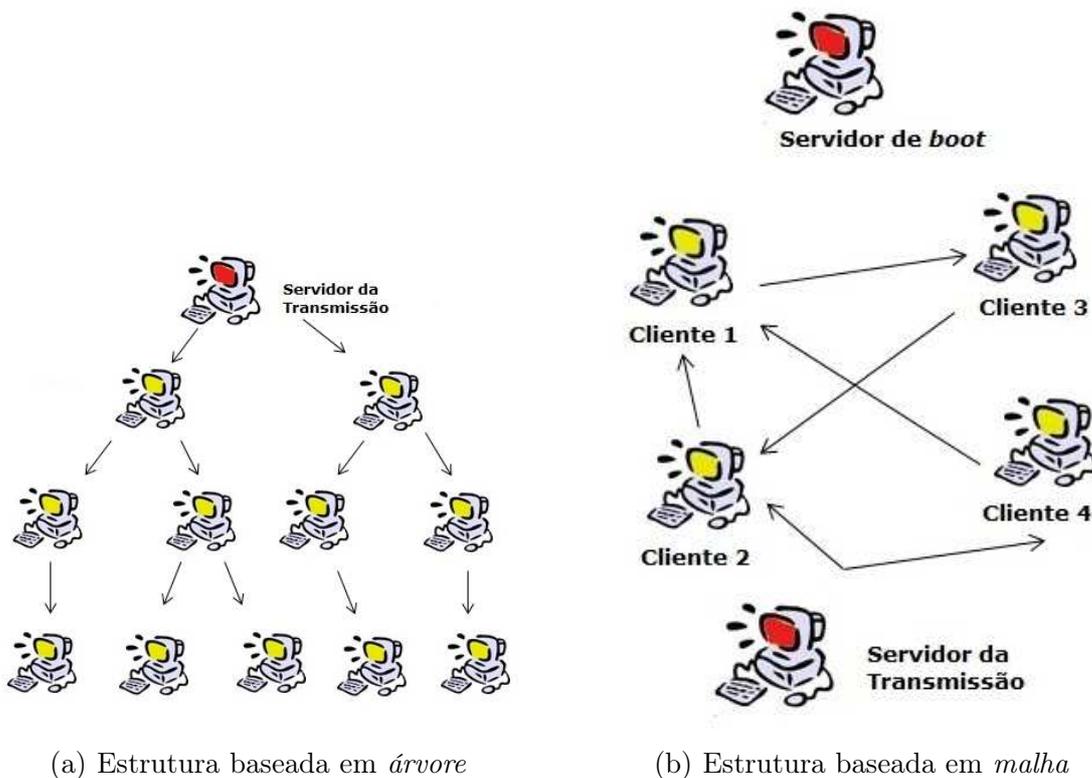


Figura 2.1: Representação da rede sobreposta de sistemas P2P

Já as redes P2P baseadas em *malha* não apresentam uma topologia fixa, e cada participante recebe conteúdo de vários outros nós da rede. Isto faz com que a rede seja mais robusta ao *churn*, porém o fluxo de transmissão do vídeo entre os nós fica praticamente imprevisível com a topologia dinâmica da rede [Liu et al., 2008]. A Figura 2.1b mostra uma representação de uma rede sobreposta baseada em *malha*,

que é tipicamente composta por servidor de transmissão, servidor de *bootstrap* e nós (clientes).

O servidor de transmissão é a fonte geradora de vídeo para todos os participantes da rede. Ele é responsável por codificar e fazer a divisão do vídeo em pequenos pedaços (*chunks*) identificados temporalmente. Ele atua na rede repassando os *chunks* aos clientes da rede conectados diretamente a ele.

O servidor de *bootstrap* é um servidor centralizado que mantém uma lista com nós que atualmente participam da rede e o identificador do *chunk* atual que o servidor de transmissão está transmitindo. Assim quando um nó entra na rede, ele se cadastra neste servidor, recebe uma lista de parceiros candidatos para a transmissão do vídeo e o identificador do *chunk* mais recente. Um nó que já está na rede também pode, a qualquer momento, solicitar uma nova lista de parceiros.

Cada nó da rede mantém um mapa de *chunks* para identificar aqueles que ele já tem e os que necessita. Constantemente este mapa é atualizado funcionando como um *buffer circular*, ou seja, a posição inicial contém o próximo *chunk* a ser consumido pela aplicação e a posição final deve ser atualizada com *chunks* mais recentes. Este mapa representa um trecho contínuo da mídia que será reproduzido pelo cliente do sistema de transmissão de vídeo ao vivo.

Um nó deve estabelecer uma conexão com outros nós da rede para a troca dos *chunks* necessários. Uma conexão entre dois nós da rede para a troca de conteúdo entre eles é chamada de *parceria*. Um novo nó ao se conectar na rede, recebe do *bootstrap* uma lista de parceiros candidatos dos quais ele pode tentar estabelecer parcerias. Além disso, sempre que necessário um nó qualquer pode solicitar ao *bootstrap* uma nova lista de parceiros. Os nós da rede estabelecem e desfazem dinamicamente suas parcerias na rede dependendo da disponibilidade de recursos em seus parceiros. Para representar a dinamicidade das parcerias, a Figura 2.2 mostra duas fotografias consecutivas (*A* ocorre antes de *B*) de uma rede sobreposta hipotética formada para a transmissão de um vídeo ao vivo. Na Figura 2.2 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*.

Para que ocorra a transmissão de vídeo entre os parceiros, primeiramente o servidor de transmissão repassa os *chunks* do vídeo aos seus nós parceiros. Estes nós, por sua vez, repassam aos seus parceiros e assim sucessivamente. Essa transmissão de dados entre parceiros pode acontecer de duas maneiras: encaminhamento automático (*push*) ou pedidos explícitos (*pull*).

Na transmissão em redes baseadas em malha por encaminhamento automático (*mesh-push*), quando um nó recebe um *chunk*, ele transmite a todos os seus parceiros.

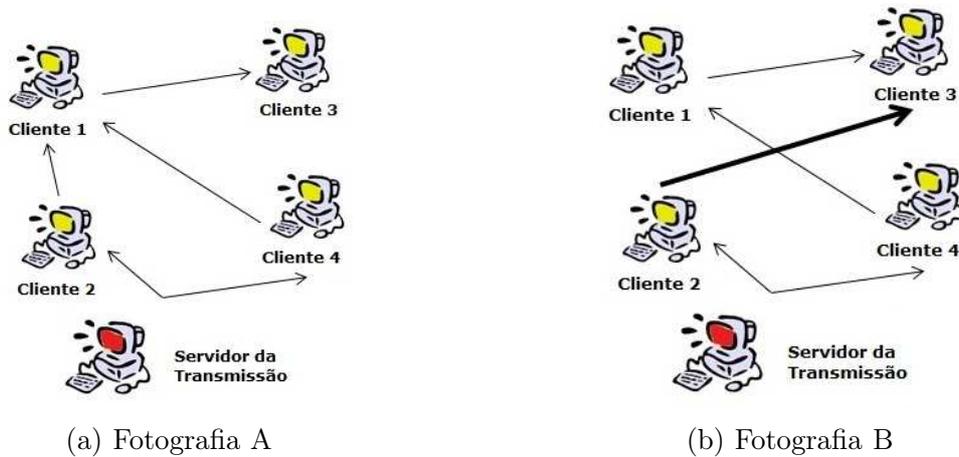


Figura 2.2: Representação da rede sobreposta de um sistema de vídeo ao vivo em duas fotografias

Neste tipo de transmissão pode ocorrer a redundância de *chunks*, pois, por exemplo, se dois nós possuem um parceiro em comum, eles receberão o mesmo *chunk* dos dois parceiros. Já na transmissão por pedido explícito (*mesh-pull*), primeiramente os nós trocam as informações dos *chunks* existentes em cada um, para depois fazerem as requisições dos *chunks* pendentes aos seus parceiros. Neste caso, a redundância de *chunks* pode ser evitada, pois os nós solicitam apenas os *chunks* que ainda estão pendentes. No entanto, podem ocorrer atrasos na transmissão devido a necessidade de pedido explícito e pela necessidade de se encontrar os *chunks* pendentes entre os parceiros [Borges, 2010].

Atualmente a maioria das aplicações de transmissão de vídeo ao vivo usam redes sobrepostas baseadas em malha e com pedido explícito por dados, também conhecido como “*data-driven mesh-pull overlay*” [Hei et al., 2008].

2.2 Caracterizações de Sistemas de Transmissão de Vídeo ao Vivo

Sistemas P2P de transmissão de vídeo ao vivo vêm sendo 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 estabelecimento de rede sobreposta de forma a reduzir o atraso de difusão de conteúdo em uma transmissão ou aumentar a escalabilidade do sistema sem perda de qualidade do

conteúdo [Liu et al., 2008; Silverston et al., 2009; Pourebrahimi et al., 2005]. Em particular, as caracterizações anteriores de sistemas P2P de transmissão de vídeo ao vivo abordaram, principalmente, aspectos relacionados ao comportamentos dos usuários destas aplicações [Huang et al., 2008; Hei et al., 2007; Qiu et al., 2009; Borges et al., 2012], ao tráfego gerado [Silverston et al., 2007b,a, 2006, 2009; Hei et al., 2006; Bermudez et al., 2011; Silverston et al., 2010], à topologia da rede sobreposta [Vu et al., 2010; Tang et al., 2009; Oliveira et al., 2010; Yang & Garcia-Molina, 2003] e à qualidade do vídeo recebido [Fallica et al., 2008; Liang et al., 2009].

No que tange as caracterizações de comportamento de usuários, Huang et al. [2008] analisaram o comportamento dos clientes do PPLive, focando em aspectos como a chegada de clientes e por quanto tempo um cliente permanece assistindo um mesmo canal (comumente chamado de *sessão*). Hei et al. [2007] mostraram que os usuários da aplicação PPLive apresentam um comportamento similar ao de usuários das tradicionais TVs. Eles analisaram o comportamento dinâmico dos usuários no que tange a entrada e a saída do sistema e a distribuição geográfica dos usuários. Mais ainda, eles também caracterizaram o número e a duração das parcerias realizadas, bem como o volume de tráfego trocado entre os parceiros.

Qiu et al. [2009] caracterizaram o comportamento dos usuários em um sistema de transmissão de vídeo ao vivo, focando na duração das sessões, trocas de canais, assim como no comportamento dinâmico (*churn*) dos pares ao longo do tempo. Os autores modelaram a duração das sessões dos usuários utilizando um modelo exponencial misto (*mixture exponential model*). Eles também modelaram a popularidade dos canais deste tipo de aplicação utilizando uma distribuição *Zipf-like* [Tang et al., 2003] para os 10% dos canais mais populares e uma distribuição exponencial para os demais canais. Já Borges et al. [2012] analisaram as semelhanças e diferenças entre os padrões de comportamento dos usuários do SopCast em diferentes tipos de transmissões. Os autores categorizaram as transmissões ao vivo em *transmissões regulares* e *transmissões de eventos*. Uma *transmissão regular* transmite a programação regular do canal, enquanto uma *transmissão de evento* traz conteúdo excepcional, por exemplo a transmissão de jogos da Copa do Mundo. Os autores caracterizaram os números de parceiros dos nós, as durações das parcerias e o tempo que os nós permanecem na rede realizando trocas de dados com seus parceiros. Esta dissertação busca complementar os resultados anteriores sobre o comportamento dos usuários em sistemas P2P de transmissão de vídeo ao vivo, mostrando como as propriedades da rede evoluem ao longo do tempo.

Alguns outros trabalhos analisaram as características do tráfego de aplicações P2P de transmissão de vídeo ao vivo. Por exemplo, Silverston et al. [2007b,a]

analisaram as características do tráfego gerado pelas aplicações SopCast, PPLive, TVAnts⁴ e PPStream, tais como as taxas de *upload* e *download*, bem como o tamanho médio dos pacotes trocados, utilizando um conjunto de dados coletados durante a Copa do Mundo de 2006. Em outro trabalho, os mesmos autores mostraram resultados de uma caracterização mais ampla considerando o uso de diferentes protocolos (TCP e UDP) e portas, assim como as durações das sessões dos clientes e as parcerias estabelecidas ao longo de uma transmissão [Silverston et al., 2009]. Em um outro trabalho [Silverston et al., 2006], os autores analisaram o sistema TVAnts caracterizando as taxas de *upload* e *download* dos clientes, bem como o número de nós ativos ao longo do tempo.

Ainda sobre características de tráfego, Hei et al. [2006] analisaram as taxas de *upload* e *download* de clientes do PPLive, bem como o número e a localização geográfica dos seus parceiros. Em outro trabalho [Silverston et al., 2010], os mesmos autores mostraram uma caracterização considerando os sistemas PPStream, TVUPlayer⁵, SopCast e TVAnts com dados coletados em nós da rede na França e no Japão. Os autores mostraram que, especificamente no SopCast, os nós enviaram mais dados do que receberam em comparação com as outras aplicações. Além disso, os autores correlacionaram a localização geográfica dos nós com suas participações na transmissão de conteúdo na rede. A localização geográfica dos nós no SopCast também foi analisada por Bermudez et al. [2011]. Porém, neste trabalho os autores analisaram se o descobrimento de novos parceiros durante uma transmissão está relacionado com os ISPs dos nós, ou seja, se nós tendem a realizar parcerias com nós do mesmo ISP que eles estão localizados. Além disso, eles analisaram a evolução do tráfego gerado com a entrada e saída dos nós na rede ao longo da duração de uma transmissão.

Existem trabalhos ainda que buscam também analisar as características da rede sobreposta formada em aplicações de transmissão de vídeo ao vivo. Por exemplo, em [Vu et al., 2010] os autores apresentaram modelos para a distribuição de grau dos nós, duração de sessão dos clientes e ainda analisaram a participação simultânea de um nó em múltiplas redes sobrepostas. Entretanto, os autores não analisaram essas propriedades da rede ao longo de uma transmissão. Propriedades da rede sobreposta do SopCast, particularmente o grau de saída dos nós, também foram analisados por Tang et al. [2009]. Em particular, os autores identificaram uma correlação alta entre o grau de saída dos nós com a taxa de *upload* dos clientes. Eles também investigaram os tamanhos dos pacotes de controle e de dados trocados entre os parceiros na rede do SopCast.

⁴tvants.en.softonic.com

⁵<http://www.tvunetworks.com>

Especificamente sobre a centralidade dos nós na rede sobreposta, Oliveira et al. [2010] utilizaram o conceito de centralidade para identificar nós especiais (denominados no trabalho como “super nós”) dentro da rede sobreposta do SopCast. Os autores correlacionaram diferentes métricas de centralidade como grau e *closeness* (veja definição na Seção 2.3.1) com a taxa de *upload* dos nós. Yang & Garcia-Molina [2003] também analisaram os graus de saída do nó para identificar “super nós” em redes P2P de compartilhamento de arquivos (por ex.: Gnutella e KaZaA⁶).

Vários outros trabalhos analisaram sistemas P2P de transmissão de vídeo ao vivo sob o ponto de vista da qualidade de serviço provido aos clientes. Por exemplo, Liang et al. [2009] mostraram os resultados de uma caracterização dos dados coletados do sistema PPStream de transmissão de vídeo ao vivo. A coleta de dados foi feita em clientes PPStream durante a transmissão da 29^a Olimpíada. Eles analisaram a latência dos pacotes de dados e verificaram se a seleção de parceiros está correlacionada com a localização dos mesmos em ISPs próximos. Fallica et al. [2008] investigaram se existe uma relação entre propriedades estruturais da rede em particular o número de parceiros e propriedades do tráfego (taxa de *download* e *upload*) com a qualidade de serviço fornecido ao cliente, estimada de forma qualitativa com experimentos com usuários.

Em comum, nenhum destes trabalhos focou primordialmente na análise de como as propriedades estruturais da rede sobreposta evoluem com o tempo durante uma transmissão, objetivo principal desta dissertação. As análises feitas neste sentido enfatizaram apenas o grau dos nós, abordando outros aspectos relacionados ao comportamento dos usuários e ao tráfego gerado [Stutzbach et al., 2008; Tang et al., 2009; Wu & Li, 2007]. Entretanto, outras propriedades dos nós e da rede como um todo ao longo de uma transmissão podem contribuir para o entendimento do real dinamismo da rede sobreposta formada em aplicações P2P de transmissões de vídeo ao vivo.

O trabalho que mais se aproxima do desenvolvido nesta dissertação é o de Wu & Li [2007]. Nele, os autores realizaram uma caracterização do sistema comercial de transmissão de vídeo ao vivo, UUSee⁷, analisando as propriedades topológicas da rede, tais como distribuição do grau, coeficiente de agrupamento e reciprocidade, e como elas evoluem ao longo do tempo. O nosso trabalho complementa o de Wu & Li [2007] pois, além de caracterizar como as propriedades dos nós evoluem com o tempo, ele também identifica perfis de centralidade, analisa como os nós mudam de perfil ao longo do tempo, além de analisar várias propriedades da rede, tais como diâmetro, coeficiente de agrupamento, caminho mínimo médio e assortatividade (veja definição na Seção 2.3.2), que não foram analisadas em [Wu & Li, 2007]. No Capítulo 4, nós

⁶<http://www.kazaa.com>

⁷<http://www.uusee.com>

iremos ressaltar as diferenças entre nossos resultados e os obtidos por Wu & Li [2007].

Um outro trabalho que se aproxima do desenvolvido nessa dissertação é o de Stutzbach et al. [2008]. Nele, os autores analisaram como a rede sobreposta da aplicação Gnutella de compartilhamento de arquivos evolui com o tempo. Eles identificaram que mesmo com o crescimento da rede, ela permanece apresentando características da rede “Small World” [Watts, 1999], isto é a maioria dos nós da rede podem ser alcançados a partir de um número pequeno de outros nós. Eles mostraram ainda que quanto maior o tempo de um nó na rede, mais estável a conectividade entre os parceiros deste nó. Diferentemente desse trabalho, essa dissertação foca em uma aplicação de transmissão de vídeo ao vivo que pode exibir propriedades da rede sobreposta bem diferentes considerando que a transmissão ao vivo exibe restrições rígidas no tempo de envio dos pacotes.

2.3 Métricas de Redes Complexas

Existem vários trabalhos que buscam analisar as propriedades de diferentes tipos de redes tais como redes de email [Gomes et al., 2009], redes P2P de compartilhamento de arquivos [Iliofotou et al., 2009; Stutzbach et al., 2008], redes sociais [Kitsak et al., 2010; Weng et al., 2010], redes de sensores sem fio [Scherrer et al., 2008; Lederer et al., 2009], redes de organizações criminosas [Xu et al., 2004], entre outras.

Nesta dissertação, foi analisada a rede sobreposta formada pelos clientes conectados em um sistema P2P de transmissão de vídeo ao vivo. A rede sobreposta é representada por um grafo direcionado $G = (V, E)$, onde V é o conjunto de vértices (nós da rede) e E é o conjunto de arestas que conectam dois elementos de V . Uma aresta direcionada (v_i, v_j) representa uma parceria estabelecida entre os clientes v_i e $v_j \in V$ para a transmissão de dados de v_i para v_j . Para analisar como as propriedades estruturais deste grafo evoluem com o tempo, objetivo principal desta dissertação, foram utilizadas várias métricas que são apresentadas a seguir.

2.3.1 Métricas dos Nós

Para caracterizar as propriedades topológicas dos nós, foram utilizadas métricas que avaliam a centralidade do nó na estrutura da rede sobreposta. Centralidade é um conceito antigo que possui duas interpretações diferentes [Freeman, 1979]: uma se refere à localização dos nós na estrutura da rede, e a outra se refere à conectividade do nó. A primeira interpretação de centralidade de um nó baseia-se na frequência com que ele se situa entre outros pares de nós considerando os caminhos mais curtos que os

ligam. Esta visão de centralidade indica quão estratégica é a localização de um nó na rede. A ideia é que um nó com centralidade mais alta pode ter uma influência maior na distribuição de conteúdo. A segunda interpretação de centralidade refere-se ao número de conexões de um nó, isto é, o número de arestas adjacentes a ele, indicando, portanto, o potencial de comunicação de um nó na rede [Freeman, 1979].

Bonacich & Lloyd [2001] apontam a importância de levar em consideração múltiplas métricas ao se analisar a centralidade de nós em redes. Três métricas comumente utilizadas para capturar a centralidade de um nó em uma rede são grau, *betweenness* e *closeness*, definidas a seguir.

O grau de um vértice v_i , $g(v_i)$, é definido pela soma do grau de entrada e grau de saída de um vértice v_i . O grau de entrada de v_i , $g_{in}(v_i)$, é o número de arestas que apontam para v_i no grafo, enquanto que o grau de saída, $g_{out}(v_i)$, é o número de arestas que saem de v_i . O grau de um vértice é uma das propriedades estruturais mais básicas que avalia o número total de arestas adjacentes [Easley & Kleinberg, 2010], isto é:

$$g(v_i) = g_{in}(v_i) + g_{out}(v_i) \quad (2.1)$$

O *betweenness* de um vértice v_i , $b(v_i)$, é a soma das frações dos números de caminhos mínimos entre pares de vértices do grafo que passam por v_i , computadas para todos os pares v_s e v_w de V ($v_s \neq v_w \neq v_i$). O *betweenness* de um vértice, portanto, representa a probabilidade dele estar em um caminho mínimo entre dois outros vértices quaisquer. Ele é formalmente definido como:

$$b(v_i) = \sum_{s \neq w \neq i \in V} \frac{\sigma_{sw}(v_i)}{\sigma_{sw}} \quad (2.2)$$

O *closeness* de um vértice v_i , $c(v_i)$, é o inverso do caminho mínimo médio, l , definido entre o vértice v_i e todo vértice v_w alcançável a partir de v_i no grafo, ou seja:

$$c(v_i) = \frac{|V| - 1}{\sum_{i \neq w} l(v_i, v_w)} \quad (2.3)$$

Com essa métrica é possível identificar o quão próximo um vértice está dos demais vértices da rede. Considerando o contexto específico de uma rede sobreposta de transmissão de vídeo ao vivo, um valor de *closeness* alto de um nó indica que uma informação disseminada por ele possivelmente poderá atingir os demais nós rapidamente.

Além de métricas de centralidade, uma outra métrica comumente usada para analisar a importância de vértices em diferentes redes é o *pagerank*. O *pagerank*

foi proposto para analisar a importância de páginas da Web no grafo formado por hiperlinks [Brin & Page, 1998]. Desde então, ele já foi usado em outros contextos, como exemplo para identificar usuários mais influentes em redes sociais [Tang et al., 2012; Weng et al., 2010], para detecção de *spammers* em redes de email [Chirita et al., 2005] e para melhorar a busca por arquivos em sistemas P2P de compartilhamento de arquivos [Sankaralingam et al., 2003; Shi et al., 2003].

O *pagerank* de um vértice v_i , $pr(v_i)$, é definido recursivamente. Inicialmente todos os vértices do grafo são inicializados com $pr(v_{v \in V}) = 1/|V|$. A seguir, o *pagerank* de cada vértice passa a ser atualizado conforme a equação abaixo:

$$pr(v_i) = (1 - s) + s \left(\sum_{j \in V} \frac{pr(v_j)}{g_{out}(v_j)} \right) \quad (2.4)$$

Na Equação 2.4, s é um parâmetro tipicamente utilizado com valor igual a 0.85, o mesmo adotado nesta dissertação [Brin & Page, 1998]. A recursividade termina quando os valores de *pagerank* encontrados em iterações sucessivas estabilizam. A intuição por trás do *pagerank* é que existe um “fluido” que circula na rede pelas arestas existentes entre os vértices, acumulando-se nos vértices mais importantes da rede [Easley & Kleinberg, 2010].

2.3.2 Métricas da Rede

A estrutura da rede sobreposta como um todo pode ser caracterizada utilizando várias métricas de redes complexas, tais como caminho mínimo médio, diâmetro, coeficiente de agrupamento, grau máximo da rede, assortatividade e coeficiente de reciprocidade. A seguir, nós definimos formalmente cada uma destas métricas.

O diâmetro de um grafo, $d(G)$, 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 $m(v_i, v_j)$ definida pelo número de arestas entre dois vértices v_i e v_j , considerando todos os pares de vértices no grafo G , ou seja:

$$d(G) = \max_{v_i, v_j \in V} m(v_i, v_j) \quad (2.5)$$

Esta métrica captura então a dispersão do grafo. No contexto específico de redes sobrepostas de transmissão de vídeo ao vivo, ela pode ser utilizada para fazer avaliações aproximadas sobre latência, impactos na rede, entre outros.

O caminho mínimo médio de um grafo G , $cmm(G)$, é definido como a média do menor caminho entre qualquer par de vértices, isto é:

$$cmm(G) = \frac{\sum_{i \neq w \in V} l(v_i, v_w)}{|V|} \quad (2.6)$$

O coeficiente de agrupamento de um grafo G , $cc(G)$, é dado pela média do coeficiente de agrupamento de todos os vértices do grafo. Dados três vértices v_i , v_j e v_w com arestas definidas entre (v_i, v_j) e entre (v_j, v_w) . O coeficiente de agrupamento de um vértice, $c(v_i)$, é definido como a razão entre o número de arestas entre os vizinhos de v_i e o seu maior valor possível definido por $g(v_i)(g(v_i) - 1)$, para grafos direcionados. Assim o coeficiente de agrupamento de um grafo é dado por:

$$cc(G) = \frac{\sum_{v_i \in V} c(v_i)}{|V|} \quad (2.7)$$

O grau máximo de um grafo G é dado pelo maior grau de um vértice, computado sobre todos os vértices de G .

A reciprocidade da rede captura a reciprocidade das interações da rede como um todo. O coeficiente de reciprocidade, $\rho(G)$ é definido como:

$$\rho(G) = \frac{\sum_{i \neq j \in V} (I(v_i, v_j) - \bar{a})(I(v_j, v_i) - \bar{a})}{\sum_{i \neq j \in V} (I(v_i, v_j) - \bar{a})^2} \quad (2.8)$$

onde $\bar{a} = \frac{\sum_{i \neq j \in V} I(v_i, v_j)}{V(V-1)}$, V o número de vértices na rede, e $I(v_i, v_j)$ é igual a 1 quando houver uma aresta direcionada de v_i para v_j e igual a 0 caso contrário. O coeficiente de reciprocidade $\rho(G)$ varia de -1 a 1. Valores de $\rho(G)$ maiores que 0 indicam que a rede é recíproca, caso contrário ela é anti-recíproca Benevenuto [2010].

A assortatividade da rede captura se vértices com muitas arestas tendem a se conectar com outros vértices com muitas ou poucas arestas. A assortatividade de um grafo G com $|E|$ arestas, $r(G)$, é definida como Newman [2002]:

$$r(G) = \frac{\frac{\sum_{i \in E} k_i * j_i}{|E|} - \left[\frac{\sum_{i \in E} k_i + j_i}{|E| * 2} \right]^2}{\frac{\sum_{i \in E} k_i^2 + j_i^2}{|E| * 2} - \left[\frac{\sum_{i \in E} k_i + j_i}{|E| * 2} \right]^2} \quad (2.9)$$

onde k_i e j_i são os graus dos dois vértices nas extremidades da aresta $i \in E$. A assortatividade captura a correlação entre os graus de nós adjacentes. Ela varia entre -1 e +1. Quando positiva, a rede é dita assortativa, o que implica que vértices de grau alto tendem a se conectar com vértices também de grau alto. Quando negativa, a rede é dita disassortativa. Isso implica que vértices de grau alto tendem a se conectar com vértices com graus mais baixos [Newman, 2002].

Nessa dissertação, nós aplicamos as métricas descritas acima para caracterizar

a evolução da topologia da rede sobreposta do sistema SopCast de transmissão de vídeo ao vivo. Para tal, nós seguimos a proposta de Acer et al. [2011], e analisamos as propriedades do grafo que modela essa rede ao longo do tempo. Em outras palavras, nós tiramos várias “fotografias” da rede ao longo da transmissão, caracterizando as propriedades de cada fotografia utilizando as métricas descritas aqui. No próximo capítulo, nós discutimos os dados utilizados nessa caracterização, focando particularmente na metodologia de coleta desses dados e nas características do sistema alvo de estudo, o SopCast.

Capítulo 3

Metodologia de Coleta de Dados

A base de dados utilizada nesta dissertação foi obtida através de experimentos realizados no PlanetLab [Peterson et al., 2006; Chun et al., 2003] com clientes da aplicação SopCast. Na Seção 3.1 é apresentada uma breve descrição desta aplicação e de seu funcionamento. A Seção 3.2 apresenta a descrição dos experimentos feitos. Por fim, a sumarização dos experimentos é apresentada na Seção 3.3.

3.1 SopCast

O SopCast é uma das aplicações Par-a-Par de transmissão de vídeo ao vivo mais populares atualmente (quando esta dissertação foi elaborada), tendo superado outras aplicações, como PPLive e PPStream, em termos de tráfego em 2010 [Google, 2010]. A aplicação transmite vídeo ao vivo em vários *canais*, sendo que uma rede sobreposta é criada para a transmissão do conteúdo de cada canal. Os clientes SopCast interessados em receber um determinado conteúdo se conectam ao canal, passando a fazer parte da sua rede sobreposta. Os canais do SopCast podem ser *abertos* ou *fechados*. Um *canal aberto* fica visível na página da aplicação, disponível para o acesso de qualquer cliente. O SopCast permite ainda a criação de *canais fechados* que não são divulgados pela aplicação. Neste caso, a população de clientes é mais restrita e controlada. A Figura 3.1 é um screenshot da aplicação, em que a esquerda são listados os canais disponíveis para o usuário logado e a direita é exibido o conteúdo que está sendo transmitido no canal selecionado pelo usuário.

O SopCast implementa redes P2P com estrutura em malha. A rede sobreposta criada para a transmissão do conteúdo de um canal é composta por um servidor de transmissão ao vivo, um servidor de *boot* (*bootstrap*) e os clientes (nós) da rede. O servidor de transmissão ao vivo é a fonte inicial do conteúdo e participa da rede

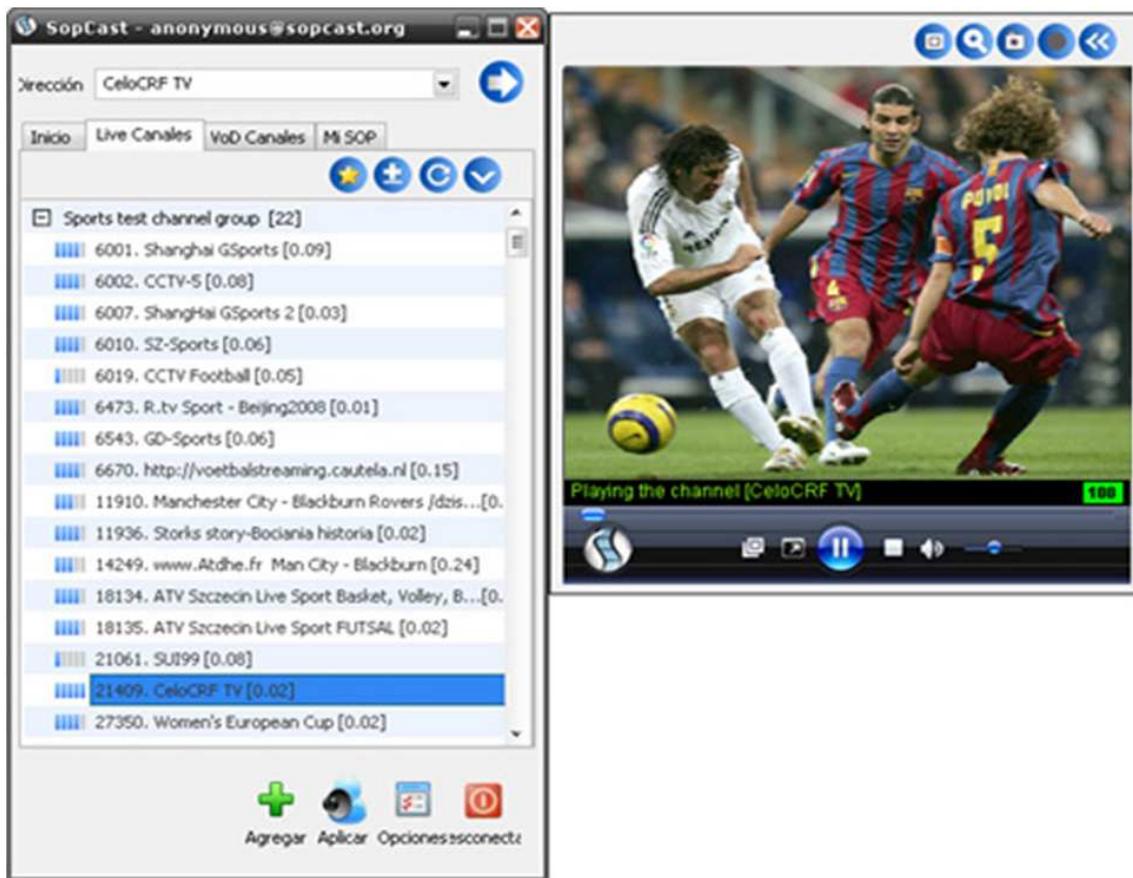


Figura 3.1: SopCast - Sistema de Transmissão de Vídeo ao Vivo em Redes P2P

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 para cada 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. Esses potenciais parceiros são clientes que já estão na rede sobreposta do canal desejado e que podem estabelecer parcerias para a troca de dados. Ao longo de uma transmissão, os clientes podem tentar estabelecer parcerias com os clientes conhecidos ou voltar ao servidor de *boot* para solicitar uma nova lista de potenciais parceiros. 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.

3.2 Monitoramento e Coleta de Dados

Foram realizados um conjunto de experimentos com o SopCast, utilizando um número N de computadores do PlanetLab [Peterson et al., 2006; Chun et al., 2003]. Cada computador foi configurado para executar um cliente SopCast e armazenar todas as trocas de dados realizadas ao longo de uma transmissão. A metodologia para a realização dos experimentos desta dissertação é composta basicamente por duas etapas: configuração dos computadores do PlanetLab e monitoramento dos mesmos durante uma transmissão de vídeo ao vivo no SopCast.

Durante a etapa de configuração, não foi imposta nenhuma restrição de armazenamento ou de taxa de *upload* e *download* nos computadores do PlanetLab. As versões de todos os softwares utilizados foram atualizadas em todos os computadores. Além do próprio cliente SopCast, foi também utilizado o Tshark¹ para capturar o tráfego de rede observado durante o monitoramento de uma transmissão de vídeo. O Tshark 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). Em particular, foram capturados e armazenados em arquivos de *log* locais em cada cliente com todos os cabeçalhos dos pacotes enviados e recebidos na porta configurada, contendo data, hora, IP de origem, IP de destino e tamanho do pacote. Para que todos os computadores estivessem sincronizados nos experimentos foi utilizado o *Network Time Protocol*² (NTP) com a configuração de um mesmo servidor de horário entre eles [Pathak et al., 2008]. Essa sincronização garante que a diferença de tempo entre os computadores seja muito pequena (menos de um segundo).

Após a etapa de configuração, os computadores passaram a executar o SopCast e a ferramenta de monitoramento Tshark. Como clientes comuns do SopCast, eles passaram a monitorar e registrar nos seus respectivos arquivos de *log* todas as parcerias realizadas, ou seja, todas as trocas de dados entre eles e outros clientes da rede sobreposta.

As conexões dos clientes SopCast ao canal desejado foram realizadas durante um período inicial de T minutos, de forma que, passando esse intervalo inicial, todos os clientes faziam parte da rede sobreposta correspondente. A partir de então, cada computador armazenou localmente todos os cabeçalhos dos pacotes enviados e recebidos durante os próximos D minutos.

Finalizados os D minutos de monitoramento, todos os arquivos de *logs* dos clientes SopCast foram combinados em um único arquivo de *log* usado para reconstruir a rede

¹<http://www.wireshark.org/docs/man-pages/tshark.html>

²<http://www.ntp.org>

sobreposta do SopCast durante a transmissão de vídeo. Baseando-se nas informações da data e hora de cada pacote, foram retiradas fotografias (*snapshots*) da rede em intervalos de I segundos. Fotografias da rede capturam, portanto, todas as trocas de pacotes de dados realizadas de I em I segundos. Cada fotografia é modelada como um grafo direcionado em que um vértice representa um nó (identificado pelo IP) e uma aresta entre vértices i e j representa o envio de dados de i para j .

A etapa de monitoramento foi realizada em três cenários diferentes. No cenário identificado por C_1 , os clientes SopCast se conectam ao canal aberto CCTV-1 da China, durante horário de pico local (20 horas) [Borges et al., 2012]. 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. Neste cenário, os computadores do PlanetLab passam a atuar como coletores de informações (*crawlers*) na rede sobreposta do SopCast. Os *crawlers* capturam apenas os pacotes que passam por eles, sejam eles trocados com outros *crawlers* ou com clientes reais da rede. Os *crawlers* foram conectados ao canal durante um intervalo inicial T igual a 10 minutos, permanecendo na rede ao longo de uma transmissão de D igual a 40 minutos. O instante de conexão de um nó no canal foi determinado segundo uma distribuição uniforme com valores entre 0 e 10 minutos.

Note que a rede sobreposta reconstruída no cenário C_1 é uma visão parcial da rede em que os *crawlers* foram conectados, pois os *crawlers* não registram as trocas de pacotes realizadas entre clientes reais³ da rede, uma vez que eles enxergam apenas os clientes com quem eles estabelecem parcerias durante a transmissão. Essa visão parcial da rede é uma limitação da escolha de monitoração de um canal aberto. Entretanto, abordagem semelhante foi adotada em várias análises anteriores de sistemas P2P de transmissão de vídeo ao vivo [Vu et al., 2007; Fallica et al., 2008]. Uma das principais contribuições desta dissertação na tentativa de reduzir a visão parcial da rede está no número de *crawlers* utilizados: enquanto os trabalhos da literatura utilizaram no máximo 70 clientes [Fallica et al., 2008], nesta dissertação foram utilizados entre 200 a 465 *crawlers*, o que pode levar a uma visão muito mais completa da população de clientes reais [Borges et al., 2012].

Já nos cenários C_2 e C_3 , os clientes SopCast são conectados a um canal fechado, criado especificamente para a coleta de dados. Para tal, foi utilizado um servidor de transmissão de vídeo à taxa de 280 kbps. Esses cenários são ambientes controlados, com uma visão completa da rede. Eles foram projetados a fim de permitir a validação dos resultados encontrados com a visão parcial da rede do cenário C_1 . Tanto em C_2

³O termo cliente real é usado em referência aos clientes do mundo real para distingui-los dos *crawlers* desta dissertação

quanto em C_3 , existe um tempo inicial T de 10 minutos seguido por um tempo D de 40 minutos de monitoramento.

A diferença entre os dois cenários está no fato de que os clientes SopCast do cenário C_2 realizam *churn* segundo a distribuição apresentada na caracterização do comportamento dos usuários de um popular canal aberto do SopCast [Borges et al., 2012]. Neste trabalho, os autores caracterizaram o comportamento dos clientes utilizando modelo de dois estados: *ON* e *OFF* descritos em [Borges et al., 2012]. O estado *ON* representa o tempo que os clientes estão na rede e trocam dados com seus parceiros. Os tempos do estado *ON* foram modelados segundo uma distribuição *Weibull* com parâmetros $\alpha = 2.032$ e $\beta = 0.233$. Já o estado *OFF* representa o tempo em que os clientes não trocam dados com parceiros, permanecendo inativos. No estado *OFF*, os tempos foram modelados segundo uma distribuição *Exponencial* com média $\lambda = 0.054^2$ [Borges et al., 2012]. No cenário C_3 os clientes SopCast permanecem na rede ao longo de toda a transmissão, ou seja, os clientes não exibem *churn*.

3.3 Sumário dos Experimentos

Em suma foram realizados 7 experimentos no cenário C_1 (canal aberto) entre os dias 28/10/2010 e 17/11/2010 com duração de transmissão de vídeo de 50 minutos em cada dia. Foram utilizados em cada experimento entre 200 a 465 clientes SopCast conectados ao CCTV-1.

Foram também realizados 6 experimentos tanto no cenário C_2 quanto no cenário C_3 entre os dias 19/11/2011 e 24/02/2012, cada um com duração de 50 minutos. Foram utilizados em média 450 clientes do SopCast, conectados ao canal fechado criado em um servidor particular de transmissão disponibilizado na rede do Departamento de Ciência da Computação da Universidade Federal de Minas Gerais.

Em todos os cenários, os *logs* de cada cliente SopCast foram agrupados por experimento e retiradas 40 fotografias (*snapshots*) de 60 segundos cada fotografia. Foi considerado o tempo inicial T de 10 minutos e a duração D de monitoramento igual a 40 minutos. Todos os dados dos cenários estão sumarizados na Tabela 3.1.

Tabela 3.1: Sumário dos Experimentos Realizados

	C_1	C_2	C_3
Número de experimentos realizados	7	6	6
Período dos experimentos	28/10/2010 a 17/11/2010	19/11/2011 a 22/11/2011	02/02/2012 a 24/02/2012
Número de computadores por experimento (N)	200 a 465	450	450
Canal utilizado	CCTV-1	Fechado(Com <i>Churn</i>)	Fechado(Sem <i>Churn</i>)
Horário	20:00 hrs	Entre 16 e 20 horas	Entre 18 e 03 horas
Tempo inicial de conexão (T)	10 minutos	10 minutos	10 minutos
Duração da transmissão (D)	40 minutos	40 minutos	40 minutos
Número de fotografias da rede por experimento	40	40	40
Duração de cada fotografia	60 segundos	60 segundos	60 segundos

Capítulo 4

Caracterização da Evolução da Rede Sobreposta do SopCast

Este capítulo apresenta a metodologia e os resultados da caracterização da rede sobreposta de um sistema P2P de transmissão ao vivo. Na Seção 4.1 é apresentada uma visão de geral das etapas de caracterização. Estas etapas são detalhadas e seguidas pelos resultados encontrados nas Seções 4.2 (visão local do nó) e 4.3 (visão global da rede).

4.1 Overview da Metodologia

A caracterização da evolução da rede sobreposta do SopCast foi realizada sob duas perspectivas: visão local de cada nó e visão global da rede. Para tal, foram realizados experimentos com a aplicação SopCast de transmissão de vídeo ao vivo e coletados os dados para a reconstrução da rede sobreposta de cada experimento, conforme descrito no Capítulo 3.

Na visão local do nó, a primeira questão foi verificar qual importância de um nó em termos de sua centralidade na rede sobreposta. Para tal, foram medidos valores de grau, *betweenness* e *closeness* de cada nó por fotografia de rede. Os valores medidos por nó foram submetidos ao algoritmo de agrupamento *k-means* para que fossem identificados perfis de centralidade dos nós. Os resultados desta etapa são apresentados na Seção 4.2.1.

Com os perfis identificados, a próxima questão foi avaliar como um nó muda de perfil de centralidade ao longo de uma mesma transmissão. Neste caso, foi utilizado a técnica de modelagem *Customer Behavior Model Graph* (CBMG) para que se

representasse a probabilidade transição de um nó de um perfil de centralidade para outro ao longo de uma transmissão (Seção 4.2.2).

Com os resultados destas duas questões é possível analisar um nó em termos de sua centralidade e como ele muda sua centralidade ao longo de uma transmissão, porém estas informações não avaliam como sua lista de parceiros se altera ao longo dessa mesma transmissão. Sendo esta, portanto, a próxima questão desta caracterização em que o objetivo é identificar se os nós tendem a variar seus parceiros ao longo do tempo. Para tal, avaliou-se a diferença da lista de parceiros dos nós por perfil de centralidade entre fotografias consecutivas e não consecutivas da rede sobreposta. Esta avaliação da lista de parceiros de um nó é apresentada na Seção 4.2.3.

Contudo, outra questão a ser avaliada foi qual a importância de um nó na transmissão de conteúdo dado seu perfil de centralidade na rede sobreposta. Nesta dissertação foi utilizada a métrica *pagerank* para que se avaliasse se a importância de um nó está relacionada com seu perfil de centralidade. Os resultados dessa etapa de caracterização são apresentados na Seção 4.2.4. Além disso, avaliou-se o volume de *download* e *upload* realizados pelos nós por perfil de centralidade (Seção 4.2.5).

Esta visão local do nó fornece informações dos nós dentro da rede sobreposta que podem impactar na estrutura da rede sobreposta como um todo. Na caracterização da visão global da rede sobreposta buscou-se avaliar a evolução das propriedades estruturais dessa rede ao longo do tempo. Para tal, foram utilizadas métricas de rede (p.ex.: diâmetro, coeficiente de agrupamento, recipricidade, entre outras) que buscam caracterizar essas propriedades.

4.2 Caracterização dos Nós

A caracterização dos nós (visão local dos nós) foi possível através da reconstrução da rede sobreposta para cada fotografia coletada em cada experimento durante a coleta de dados. Inicialmente, nesta dissertação são apresentados os resultados encontrados a partir da caracterização das redes sobrepostas coletadas nos experimentos do cenário C_1 , realizados em um canal aberto do SopCast (vide Capítulo 3). Apesar de ser um ambiente real de transmissão, este cenário fornece apenas uma visão parcial da rede, já que os *crawlers* utilizados na coleta de dados não conseguem capturar as trocas de dados entre usuários reais. Os resultados encontrados neste cenário (C_1) foram validados com os resultados da caracterização das redes sobrepostas coletadas nos experimentos dos cenários C_2 e C_3 , discutidos no final desta seção. Em cada experimento foram retiradas fotografias da rede sobreposta. Cada fotografia é representada por um grafo direcionado

em que os vértices (nós da rede) deste grafo são identificados pelo endereço IP e as arestas representam as trocas de pacotes realizadas entre os vértices (parcerias).

4.2.1 Perfis de Centralidade dos Nós

Para cada fotografia da rede foram extraídas as medidas de grau (Equação 2.1), *betweenness* (Equação 2.2) e *closeness* (Equação 2.3) de cada um dos nós com o objetivo de se identificar perfis de centralidade dos nós ao longo do tempo (visão local dos nós). Para tal, cada nó v_i presente em uma fotografia t da rede foi representado por um vetor $\vec{v}_i^t = (g^t(v_i), b^t(v_i), c^t(v_i))$, onde $g^t(v_i)$, $b^t(v_i)$ e $c^t(v_i)$ são o grau, *betweenness* e *closeness* respectivamente de v_i na rede sobreposta capturada pela fotografia t . Um experimento com n fotografias e, em média $|V|$ vértices possui, portanto, $n|V|$ vetores (em média) com as medidas de centralidade dos nós por fotografia da rede.

Para que perfis distintos de centralidade dos nós fossem identificados, foi utilizado o algoritmo de agrupamento *k-means* [Wan et al., 1988]. O *k-means* é um algoritmo não supervisionado que deve descobrir, sem influências externas, possíveis relações, padrões, classificações ou categorias nos dados recebidos como entrada. Ele agrupa os dados, identificando grupos que apresentam atributos semelhantes. Especificamente, nesta dissertação o *k-means* foi executado, recebendo como entrada todos os vetores \vec{v}_i^t de todas as fotografias retiradas em um experimento. Logo, os atributos considerados foram o grau, *betweenness* e *closeness* de cada vértice em cada fotografia da rede. O *k-means* foi aplicado separadamente para os dados coletados de cada experimento para verificar a consistência dos resultados em diferentes experimentos.

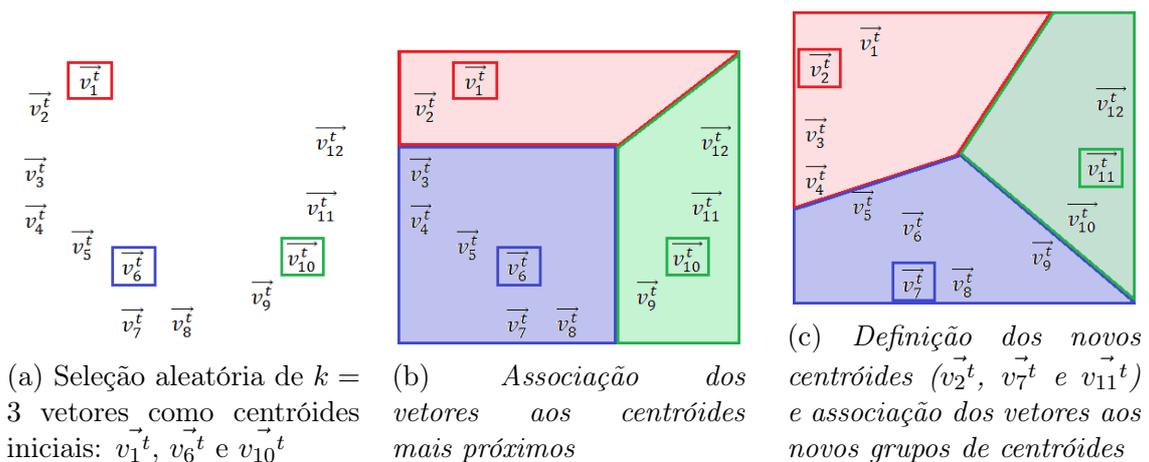


Figura 4.1: Algoritmo de Agrupamento *k-means*

O *k-means* funciona da seguinte maneira. Inicialmente são selecionados k vetores

iniciais como centros de grupos ou centróides (Figura 4.1a), onde k é um parâmetro de entrada do algoritmo e representa o número de grupos a serem criados. Os outros vetores são associados aos centróides mais próximos, passando a fazer parte do grupo do centróide em que eles foram associados (Figura 4.1b). A definição de proximidade (distância) utilizada no algoritmo pode ser medida de diferentes formas e captura a similaridade entre um vetor e cada centróide. Nesta dissertação foi utilizada a distância Euclidiana, de , que é calculada para dois vetores \vec{v}_i^t, \vec{v}_j^t como:

$$de(\vec{v}_i^t, \vec{v}_j^t) = \sqrt{[g^t(v_i) - g^t(v_j)]^2 + [b^t(v_i) - b^t(v_j)]^2 + [c^t(v_i) - c^t(v_j)]^2} \quad (4.1)$$

onde $g^t(v_i)$, $b^t(v_i)$ e $c^t(v_i)$ são o grau, *betweenness* e *closeness* respectivamente de v_i na rede sobreposta capturada pela fotografia t . Na próxima etapa do algoritmo, os centróides são substituídos por vetores mais centrais de cada grupo (Figura 4.1c), e então, os outros vetores são novamente associados aos centróides mais próximos e assim sucessivamente até que ocorra uma convergência dos resultados em uma melhor divisão dos vetores em grupos de acordo com o critério de similaridade dos dados utilizado. Isto é, a distância total entre os vetores de um grupo e seu respectivo centróide para todos os grupos deve ser minimizada.

O número k de grupos do *k-means* pode ser definido utilizando a relação β_{cv} que é a razão do coeficiente de variação (CV) das distâncias entre os vetores e os centróides dos seus respectivos grupos pelo coeficiente de variação das distâncias entre os próprios centróides. A ideia é que o número de grupos k ideal deve minimizar o coeficiente de variação intra-grupos enquanto maximiza o coeficiente de variação entre os grupos. A estratégia para determinar k consiste em executar o *k-means* para valores crescentes de k , calculando o β_{cv} para cada agrupamento criado. Em seguida, deve-se avaliar como o valor do β_{cv} varia à medida que k aumenta. O melhor valor de k deve ser quando o β_{cv} se torna relativamente estável [Menasce & Almeida, 2007].

Os grupos identificados no *k-means* representam perfis de centralidade dos nós nas fotografias dos experimentos. Os valores medidos de grau, *betweenness* e *closeness* dos nós nos experimentos do cenário C_1 foram aplicados ao algoritmo de agrupamento *k-means* que levou à identificação de três perfis de centralidade em todos os experimentos realizados. Estes perfis foram denominados nesta dissertação como *Centralidade Alta* (CA), *Centralidade Intermediária* (CI) e *Centralidade Baixa* (CB). A Tabela 4.1 sumariza os valores médios e os coeficientes de variação (CV) das medidas de centralidade dos nós em cada perfil 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.

Tabela 4.1: Perfil de centralidade dos nós em cada experimento (Cenário C_1)

	# Crawlers	Perfil	% de Nós	Grau		<i>Betweenness</i>		<i>Closeness</i>		
				Média	CV	Média	CV	Média	CV	
Experimento	1	315	CA	4.76%	282.83	0.17	3312.52	0.45	0.005	1.25
			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	229	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	393	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%	77.06	1.14	135.47	2.30	0.003	0.90
	4	308	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	266	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%	89.74	1.14	120.73	2.24	0.003	1.04
	6	288	CA	2.69%	241.23	0.20	11865.11	0.63	0.004	0.12
			CI	8.99%	252.24	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	293	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

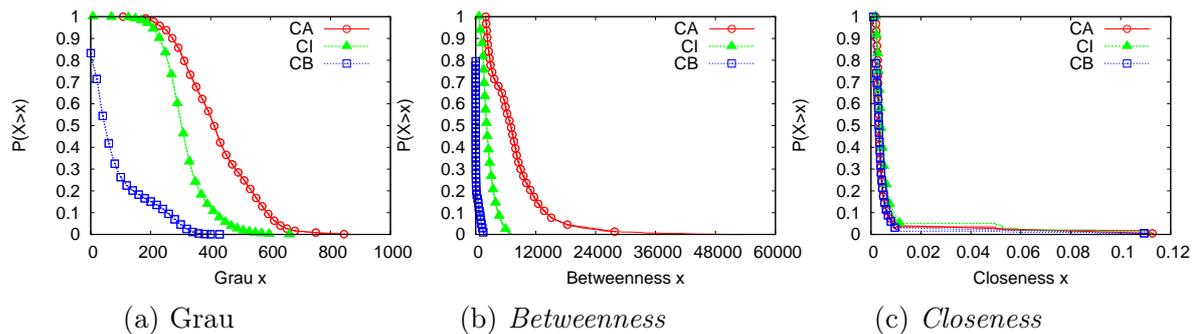


Figura 4.2: Distribuição das métricas de centralidade por perfil (Experimento 1, Cenário C_1)

O perfil *CA* representa nós que possuem grau e *betweenness* mais elevados do que os demais nós. Eles correspondem a uma pequena fração da população de nós da rede, variando entre 2.5% a 7% dos nós (ocasionalmente 20%). Os nós do perfil *CB* são aqueles que apresentam baixos valores de grau e *betweenness*, podendo ser

intuitivamente entendidos como nós da “periferia” da rede. Eles são, também, os mais comuns na rede contendo em torno de 47% a 82% dos nós. Já o perfil *CI* tem características intermediárias aos perfis *CA* e *CB*.

As diferenças entre os perfis de centralidade são apresentadas na Figura 4.2, que mostra as distribuições acumuladas complementares (CCDFs) de cada métrica para o Experimento 1 (ver Tabela 4.1). Resultados semelhantes foram obtidos para os demais experimentos e, por isso, foram omitidos. A Figura 4.2a mostra que nós com perfil *CA* tendem a ter maiores graus: apenas 4.8% dos nós possuem grau inferior a 200. Em contrapartida, em torno de 62.5% dos nós do perfil *CB* possuem grau abaixo de 2. A distribuição do *betweenness* (Figura 4.2b) também apresenta uma clara diferença entre os perfis: nós do perfil *CA* possuem valores mais altos que nós dos perfis *CI* e *CB*. No entanto, os valores medidos do *closeness* (Figura 4.2c) não exibem variações significativas entre os perfis de centralidade.

4.2.2 Mudanças da Centralidade dos Nós

Com os perfis de centralidade definidos, foi analisado como os nós mudam de perfil de centralidade durante uma transmissão ao vivo. Ou seja, um nó que possui um determinado perfil no início de uma transmissão pode ter outro perfil de centralidade em outro momento da mesma transmissão. Essa alteração dinâmica da centralidade dos nós pode ocasionar mudanças na estrutura da rede como um todo. Assim é possível analisar se um nó tende a permanecer no mesmo perfil ou se ele tende a mudar frequentemente de perfil durante uma mesma transmissão.

Para caracterizar estas alterações das centralidades dos nós ao longo do tempo, foi utilizado um modelo de grafo chamado *Customer Behavior Model Graph* (CBMG) [Menasce & Almeida, 2007]. O CBMG é uma técnica de modelagem proposta para capturar o comportamento de usuários em termos de padrões de navegação entre estados, que podem ser páginas Web, sistemas, perfis, entre outros. O CBMG pode ser utilizado para predição de carga de trabalho e da navegação de usuários. Um CBMG é composto por estados e transições entre estados, sendo que a cada transição está associada uma probabilidade dela ocorrer.

A Figura 4.3 apresenta um exemplo de CBMG. A seta identificada por $P(BA)$ representa a probabilidade de transição do *Estado B* para o *Estado A*, e uma transição no sentido contrário pode ocorrer com probabilidade $P(AB)$. As probabilidades $PL(A)$ e $PL(B)$ estão associadas à permanência do usuário no mesmo *Estado A* ou *B*, respectivamente. As setas que saem do círculo representam as probabilidades iniciais dos estados, e as setas que apontam para nenhum estado as probabilidades finais. Nesta

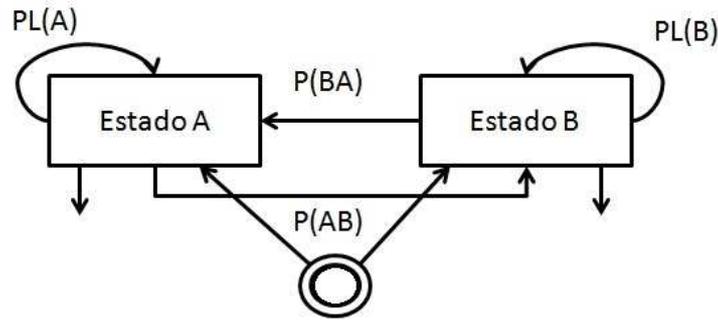
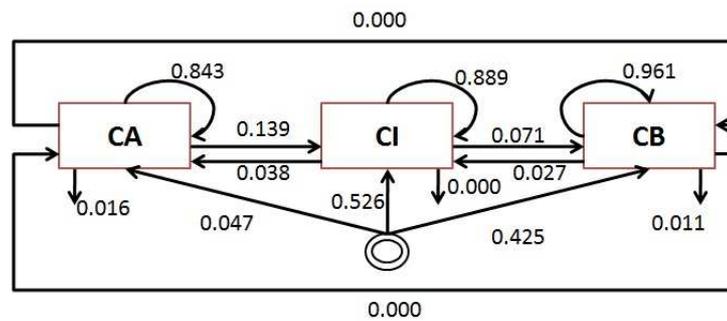
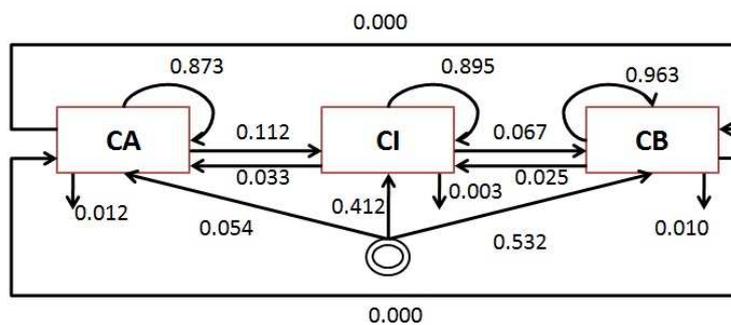


Figura 4.3: Exemplo de um CBMG

dissertação, cada estado (retângulo) do CBMG representa um perfil de centralidade do nó, e as arestas entre estados representam a probabilidade de um nó ter sua centralidade alterada de um perfil para o outro entre fotografias sucessivas.



(a) Experimento 1



(b) Todos os experimentos

Figura 4.4: CBMG - *Customer Behavior Model Graph* (Cenário C_1)

As probabilidades do CBMG foram calculadas analisando os perfis de um mesmo nó em fotografias consecutivas para todos os nós e todas as fotografias de cada experimento. A Figura 4.4a apresenta o CBMG encontrado para o Experimento 1 (Tabela 4.1) do cenário C_1 . O mesmo padrão de transição entre os perfis pode

ser observado para os outros experimentos. A Figura 4.4b mostra o CBMG médio computado para todos os sete experimentos realizados.

Note que, com altas probabilidades, os nós tendem a permanecer no mesmo perfil de centralidade em fotografias consecutivas da rede. Por exemplo, a probabilidade de um nó com baixa centralidade permanecer no mesmo perfil é 0.961 no Experimento 1 (Figura 4.4a). Já a probabilidade de um nó que tem alta centralidade permanecer no mesmo perfil é também bem alto (0.843). O CBMG mostra também que a probabilidade de um nó reduzir sua centralidade é maior do que de aumentá-la. Por exemplo, ainda no CBMG do Experimento 1, a transição entre os perfis *CA* e *CI* ocorre com probabilidade de 0.139 enquanto que o contrário ocorre com probabilidade de 0.038. A Figura 4.4b mostra que os resultados são bastantes similares em todos os experimentos.

4.2.3 Mudança na Lista de Parceiros

Outra questão que foi considerada na caracterização da visão local de um nó é se os nós tendem a variar seus parceiros ao longo do tempo. Pois, apesar dos nós tenderem a se manter no mesmo perfil de centralidade, eles podem variar a suas listas de parceiros ao longo do tempo. Isto pode acontecer sem que as características de centralidade dos nós sejam alteradas. Para tal, foi analisada a porcentagem de parceiros diferentes de um nó, $p(v_i)$, em duas fotografias diferentes (t_1 e t_2), definida pelo Coeficiente de Jaccard [Manning et al., 2008]:

$$p(v_i) = \frac{p^{t_1}(v_i) \cap p^{t_2}(v_i)}{p^{t_1}(v_i) \cup p^{t_2}(v_i)} \quad (4.2)$$

onde $p^{t_1}(v_i)$ são os parceiros do nó v_i na fotografia t_1 e $p^{t_2}(v_i)$ os parceiros do nó v_i na fotografia t_2 . Essa porcentagem de parceiros diferentes de um nó entre fotografias permite analisar se um nó tende a variar muito a lista de parceiros com quem realiza as trocas de dados.

A Figura 4.5 apresenta, para cada perfil de centralidade, a CCDF da porcentagem de parceiros diferentes de um mesmo nó em duas fotografias consecutivas da rede, tanto para o Experimento 1 quanto para todos os experimentos.

A porcentagem de parceiros diferentes entre fotografias consecutivas (Equação 4.2) tende a ser relativamente baixa nos três perfis. Em particular, para os perfis *CA* e *CI* a porcentagem máxima observada foi de 20%. A Figura 4.5b mostra, para cada perfil, a CCDF correspondente considerando todos os experimentos. Assim como o Experimento 1, a porcentagem de parceiros diferentes entre fotografias consecutivas

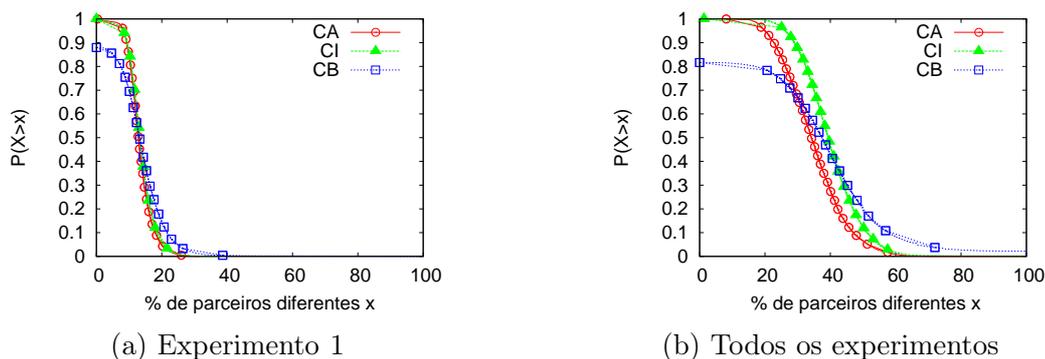


Figura 4.5: Distribuição da diferença de parceiros de um nó em fotografias consecutivas (Cenário C_1)

para todos os experimentos também tende a ser baixa.

Entretanto, ao se verificar a diferença entre parceiros de fotografias não consecutivas da rede é possível observar que a lista de parceiros sofre mais mudanças. A Figura 4.6 mostra os resultados da diferença entre parceiros de um mesmo nó em fotografias não consecutivas, para o Experimento 1. Resultados semelhantes foram encontrados nos demais experimentos, por isso foram omitidos. Na Figura 4.6a, é apresentada a diferença entre a fotografia t e a fotografia $t + 3$, ou seja, um salto de 3 fotografias da rede. Já a Figura 4.6b mostra os resultados considerando um salto de 5 fotografias, isto é, as diferenças entre as fotografias t e $t + 5$.

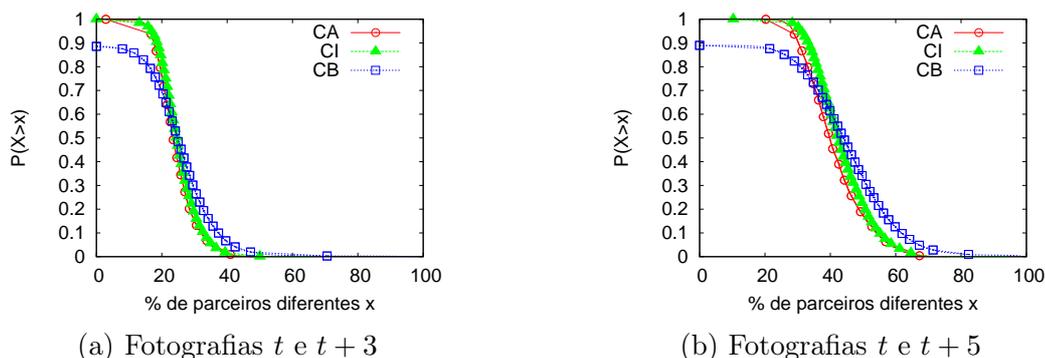


Figura 4.6: Distribuição da diferença de parceiros de um nó em fotografias não-consecutivas (Experimento 1, Cenário C_1)

Estes resultados indicam que existe uma variação na lista de parceiros ao longo de uma transmissão, apesar da estabilidade dos nós em seus perfis de centralidade. Por exemplo, em torno de 20% dos nós com perfil CA possuem pelo menos 50% de parceiros diferentes depois de um intervalo de tempo correspondente a 5 fotografias da

rede (p.ex.: 300 segundos). Essa porcentagem também é alta para os nós com perfis *CI* e *CB*. Logo, embora os nós tendem a se manter no mesmo perfil de centralidade ao longo de uma transmissão, eles tendem a mudar seus parceiros à medida que o tempo passa.

4.2.4 Pagerank

O *pagerank* (Equação 2.4) foi utilizado para tentar identificar se a importância de um nó na distribuição de conteúdo está relacionada com seu perfil de centralidade. O *pagerank* (Equação 2.4) dos nós por perfil de centralidade foi analisado em cada fotografia da rede sobreposta para avaliar a influência dos nós na transmissão de conteúdo em relação às suas respectivas centralidades. Analisou-se também a variação do valor de *pagerank* de um mesmo nó ao longo do tempo. Essa variação foi medida pela diferença percentual (em módulo) entre os valores de *pagerank* de um nó em fotografias diferentes.

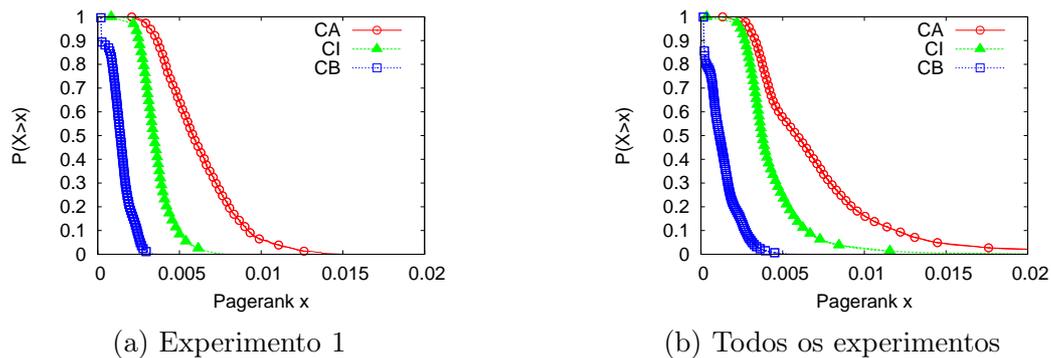


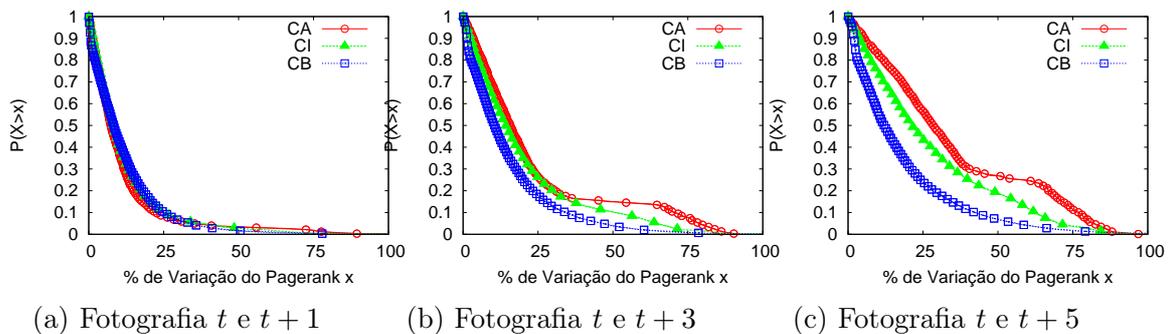
Figura 4.7: *Pagerank* dos nós por perfil de centralidade (Cenário C_1)

A Figura 4.7 mostra as distribuições dos valores de *pagerank* por perfil de centralidade no Experimento 1 (Figura 4.7a) e em todos os experimentos (Figura 4.7b). É possível observar que nós com perfil *CA* possuem valor de *pagerank* mais alto que os demais, indicando sua importância na distribuição de conteúdo entre os nós. Os valores de *pagerank* por perfil de centralidade estão sumarizados na Tabela 4.2.

A variação do *pagerank* entre fotografias consecutivas do Experimento 1, para nós com diferentes perfis, é mostrada na Figura 4.8a. Nota-se que a probabilidade de se ter uma variação nos valores de *pagerank* acima de 25% é de aproximadamente 10% para todos os perfis. Analisando o *pagerank* em fotografias não consecutivas t e $t + 3$ (Figura 4.8b), a probabilidade de variação dos valores de *pagerank* entre os perfis torna-se mais clara: a probabilidade de variação de *pagerank* é superior para nós de perfil *CA*, seguidos por nós de perfis *CI* e *CB*. Observando um intervalo um pouco

Tabela 4.2: *Pagerank* dos nós por perfil de centralidade em cada experimento do Cenário C_1

		Perfil	<i>Pagerank</i>		<i>Upload</i> (Mb)		<i>Download</i> (Mb)	
			Média	CV	Média	CV	Média	CV
Experimento	1	CA	0.0062	0.37	1.72	0.89	0.90	0.73
		CI	0.0035	0.27	1.32	0.91	1.42	0.82
		CB	0.0014	0.49	1.51	0.74	0.96	0.79
	2	CA	0.0073	0.24	1.47	0.85	1.31	0.81
		CI	0.0041	0.22	1.23	0.81	0.87	0.82
		CB	0.0009	0.92	0.98	0.71	1.12	0.87
	3	CA	0.0091	0.30	1.31	0.69	1.1	0.81
		CI	0.0042	0.38	1.27	0.83	0.97	0.89
		CB	0.0010	0.71	1.21	0.87	1.23	0.84
	4	CA	0.0130	0.34	0.99	0.67	1.24	0.72
		CI	0.0061	0.47	1.08	0.66	1.23	0.86
		CB	0.0011	0.87	1.02	0.83	1.01	0.77
	5	CA	0.0070	0.28	1.32	0.85	1.29	0.88
		CI	0.0043	0.29	1.34	0.71	1.26	0.82
		CB	0.0011	0.94	1.01	0.65	1.12	0.86
	6	CA	0.0209	0.43	1.35	0.83	1.26	0.67
		CI	0.0071	0.56	1.27	0.76	1.23	0.66
		CB	0.0017	0.62	1.22	0.81	0.97	0.73
	7	CA	0.0035	0.16	1.32	0.73	1.27	0.89
		CI	0.0028	0.15	1.25	0.88	1.27	0.71
		CB	0.0013	0.56	1.22	0.87	1.03	0.70

Figura 4.8: Porcentagem de variação do *pagerank* dos nós por perfil de centralidade entre fotografias (Experimento 1, Cenário C_1)

maior de 5 fotografias, t e $t+5$, a probabilidade de nós do perfil CB terem acima de 25% de variação no *pagerank* é de aproximadamente 30%. Já nos perfis CI e CA essa mesma probabilidade gera uma variação nos valores de *pagerank* de aproximadamente 40% e 50%, respectivamente (Figura 4.8c). Ou seja, a probabilidade de maiores variações aumenta à medida que analisa-se fotografias mais espaçadas.

Note que, a despeito da estabilidade nos perfis de centralidade, os valores de pagerank de um nó estão variando. Isto ocorre devido a variação na lista de parceiro já identificada na seção anterior. O pagerank é um algoritmo que não apenas considera a quantidade de conexões de um nó, mas também para quem este nó aponta.

4.2.5 Download e Upload

Além dos valores de *pagerank*, os volumes totais de dados que cada nó transferiu (*upload*) e recebeu (*download*) de seus parceiros também foram analisados para os nós com diferentes perfis de centralidade visando identificar se existe alguma correlação entre a centralidade de um nó e o volume de tráfego gerado por ele. As Figuras 4.9a e 4.9b apresentam as CCDFs dos volumes de *download* e de *upload*, respectivamente, para nós com diferentes perfis de centralidade no Experimento 1. Nós do perfil *CA* possuem uma ligeira tendência de terem maior volume de *upload* que os nós dos demais perfis de centralidade. Entretanto, nota-se que a despeito das diferenças de centralidade, os volumes de download e de upload seguem distribuições semelhantes, para nós nos diferentes perfis. Em outras palavras, não existe uma tendência clara de nós com centralidade mais alta terem um volume total de *upload* (ou *download*) maior que nós com centralidade inferior. O mesmo resultado foi observado nos outros experimentos.

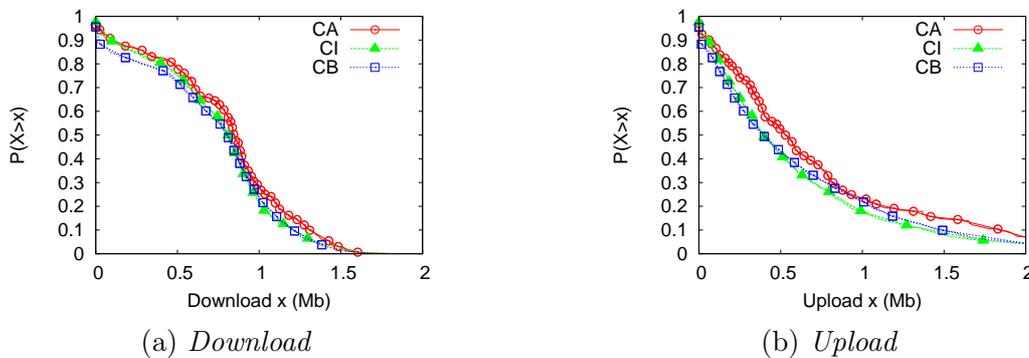


Figura 4.9: Volume de *download* e *upload* por perfil de centralidade (Cenário C_1)

4.2.6 Validação

Discute-se agora os resultados obtidos nos cenários C_2 e C_3 , que correspondem a experimentos com um canal fechado, com e sem *churn*, respectivamente. No geral, os resultados obtidos nesses cenários são muito semelhantes aos obtidos no cenário C_1 , discutidos acima, validando assim os mesmos. Estes resultados são apresentados nos

Apêndices A e B. A seguir, discutimos apenas as distorções em relação aos resultados do cenário com canal aberto (C_1). Estes se concentram nas distribuições das diferenças na lista de parceiros de um mesmo nó em fotografias diferentes.

Nos cenários de experimentos realizados com canal fechado, os nós exibem distribuições de probabilidade das diferenças de parceiros em fotografias consecutivas (Figura A.3 e Figura B.3) bastante semelhantes às distribuições correspondentes para fotografias não consecutivas (Figura A.4 e Figura B.4), diferentemente do encontrado no C_1 . Isto decorre devida a visão parcial da rede dos experimentos no cenário C_1 em que a população dinâmica de nós do mundo real vista pelos *crawlers* varia ao longo de uma mesma transmissão.

Na literatura, Wu & Li [2007] também analisaram as distribuições dos graus de entrada e saída dos nós em um sistema de transmissão de vídeo ao vivo. Entretanto, eles não analisaram os perfis de centralidade, a diferença da lista de parceiros dos nós, os valores de *pagerank* e os volumes de *upload* e *download* como foi feito nesta dissertação.

4.3 Caracterização da Rede Sobreposta

Além de caracterizar as várias métricas de nós, foram caracterizadas as métricas de rede para analisar as propriedades estruturais da rede como um todo ao longo de uma transmissão de vídeo ao vivo (visão global). Para analisar a evolução das propriedades estruturais da rede sobreposta como um todo foram extraídos valores de métricas de rede de cada fotografia dos experimentos. Para tal, foram medidos o diâmetro (Equação 2.5), o caminho mínimo médio (Equação 2.6), o grau máximo, o coeficiente de agrupamento (Equação 2.7), a reciprocidade (Equação 2.8) e a assortatividade (Equação 2.9) de cada fotografia t para cada experimento realizado. Primeiramente, foram analisados os valores médios de cada métrica em todas as fotografias de cada experimento do cenário C_1 . A Tabela 4.3 sumariza as medidas encontradas por experimento, apresentando também os coeficientes de variação (CV) de cada métrica.

A Figura 4.10 mostra as seis métricas — diâmetro, caminho mínimo médio, grau máximo, coeficiente de agrupamento, assortatividade e reciprocidade — da rede ao longo do tempo para o Experimento 1, a título de exemplo dos demais experimentos. Resultados semelhantes foram obtidos nos demais experimentos também, sendo por isso omitidos. Note que, conforme mostrados nas Figuras 4.10a, 4.10b e 4.10f, os valores de diâmetro, caminho mínimo médio e reciprocidade tendem a se manter estáveis ao longo de uma transmissão, o que também é indicado pelos baixos coeficientes de variação

Tabela 4.3: Medidas das métricas da rede (Cenário C_1)

		Experimento						
		1	2	3	4	5	6	7
Diâmetro	Média	4.11	4.13	4.33	4.27	4.22	4.23	4.05
	CV	0.07	0.09	0.13	0.10	0.09	0.10	0.05
Caminho Mínimo Médio	Média	1.98	2.08	2.09	2.16	2.14	2.10	2.00
	CV	0.03	0.03	0.02	0.03	0.02	0.07	0.02
Coeficiente de Agrupamento	Média	0.24	0.15	0.15	0.12	0.17	0.18	0.32
	CV	0.34	0.28	0.46	0.56	0.21	0.64	0.11
Grau Máximo	Média	361.47	431.51	429.71	389.88	398.68	339.83	377.00
	CV	0.08	0.11	0.06	0.09	0.12	0.11	0.05
Reciprocidade	Média	0.868	0.866	0.864	0.854	0.857	0.850	0.878
	CV	0.008	0.01	0.007	0.01	0.01	0.01	0.01
Assortatividade	Média	-0.534	-0.563	-0.518	-0.639	-0.534	-0.591	-0.343
	CV	0.23	0.16	0.35	0.31	0.23	0.43	0.32

para essas métricas (Tabela 4.3). Em particular, o diâmetro da rede permanece igual a 4 durante quase toda a transmissão, com algumas variações para 5 (Figura 4.10a). O caminho mínimo médio é aproximadamente igual a 2 (Figura 4.10b) o que esta consistente com os resultados reportados em [Wu & Li, 2007], onde os autores encontraram um caminho mínimo médio entre 2 e 6 na rede sobreposta coletada do sistema de transmissão de vídeo ao vivo UUsee.

O grau máximo da rede também se mantém estável e alto, variando entre 339.8 e 431.5 (Figura 4.10c), o que indica que durante um intervalo de tempo correspondente a uma fotografia da rede (p.ex.: 1 minuto), um nó pode trocar dados com centenas de outros nós. A troca de dados na rede é aproximadamente 80% das vezes recíproca, permanecendo assim durante toda a transmissão, ou seja, um nó que envia dados para um parceiro também recebe dados dele em 80% dos casos (Figura 4.10f). Essa reciprocidade alta da rede também foi encontrada por Wu & Li [2007]. Note que uma alta reciprocidade da rede pode contribuir para um melhor desempenho da distribuição de vídeo ao vivo em redes P2P em malha, já que não existe uma ordem pré-definida da distribuição de conteúdo, sendo assim necessária a cooperação entre os nós para uma distribuição mais efetiva dos dados [Wu & Li, 2007].

Como exceção, notamos que os valores de coeficiente de agrupamento e de assortatividade tendem a alterar com o tempo. O coeficiente de agrupamento da rede diminui ao longo do tempo (Figura 4.10d), indicando que os nós estão se conectando a outros nós que não estão conectados entre si. Essa observação é reforçada com a variação dos resultados da assortatividade entre 0 e -0.75 (Figura 4.10e), pois quanto

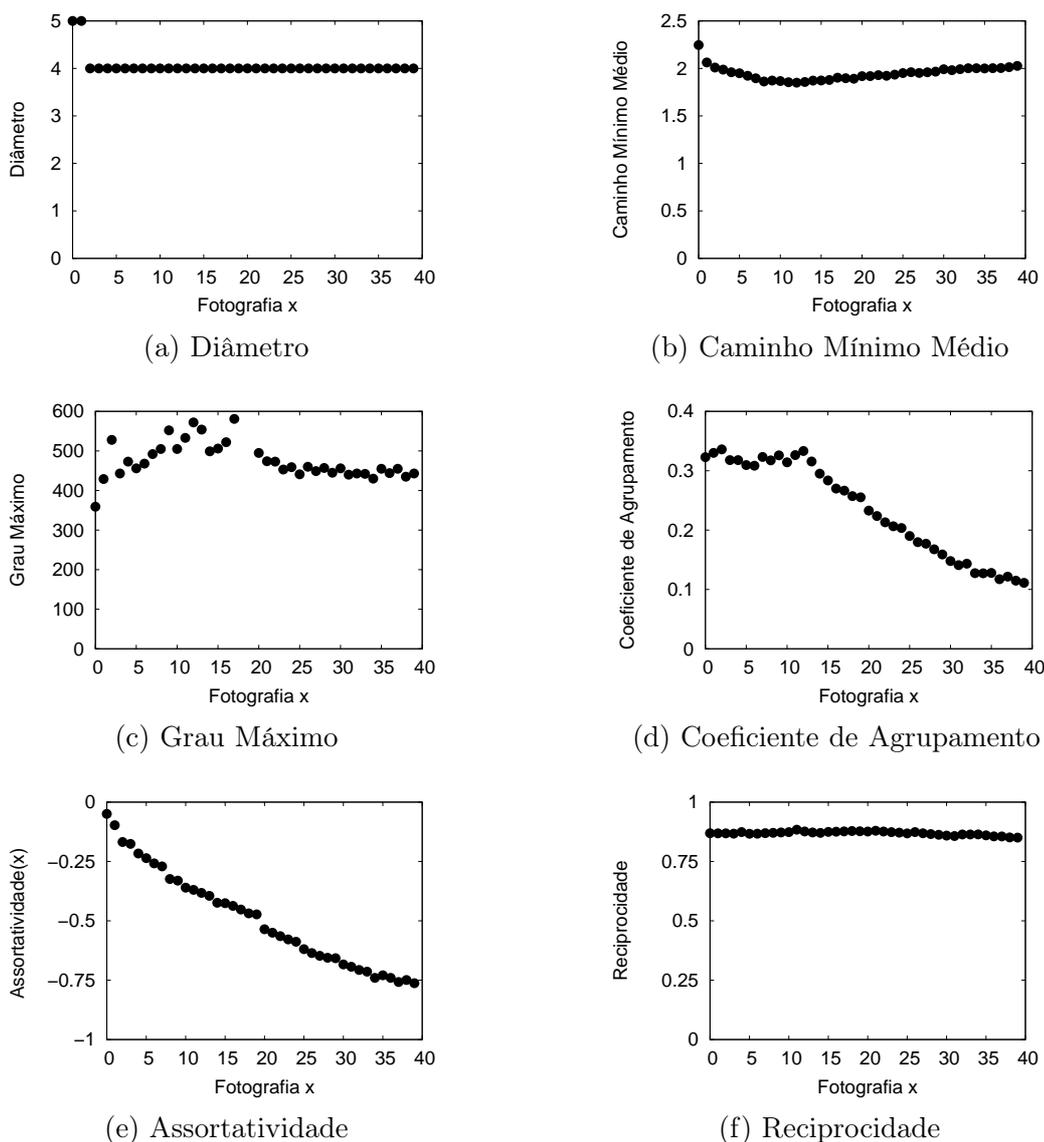


Figura 4.10: Métricas de rede ao longo do tempo (Experimento 1, Cenário C_1)

menor a assortatividade, mais nós de baixo grau tendem a se conectar a nós de alto grau.

4.3.1 Validação

Discutimos agora os resultados obtidos na visão global da rede nos cenários C_2 (Apêndice A) e C_3 (Apêndice B), comparando-os com os resultados do cenário C_1 . Como foi feito na Seção 4.2, apresentamos apenas as divergências entre os cenários que se concentram nos resultados de coeficiente de agrupamento e assortatividade da rede. Os resultados específicos dos cenários C_2 e C_3 podem ser encontrados nos Apêndices

A e B, respectivamente.

Nos cenários C_2 e C_3 , o coeficiente de agrupamento é aproximadamente igual a 0.4 e 0.29, respectivamente, mantendo-se estável durante toda a transmissão, conforme mostrado nas Figuras A.6d e B.6d. Isso difere dos resultados observados no cenário C_1 , mostrados pela Figura 4.10d, em que o coeficiente de agrupamento da rede varia entre 0.35 e 0.1 ao longo do tempo. Essa variação ao longo do tempo ocorre devido à visão parcial da rede existente no cenário C_1 , pois a visão dos nós do mundo real coletada pelos *crawlers* tende a subestimar o grau dos mesmos, e além disso, os *crawlers* mudam de a sua lista de parceiros ao longo de uma transmissão (Figura 4.5).

Já a assortatividade da rede é aproximadamente igual a 0 no cenário C_2 e -0.23 no cenário C_3 também se mantendo estável ao longo da transmissão, conforme pode ser visto nas Figuras A.6e e B.6e. Tal fato, também ocorre devido à visão parcial exibida no cenário C_1 .

As medidas de diâmetro, caminho mínimo médio, grau máximo e reciprocidade nos cenários C_2 e C_3 são qualitativamente semelhantes às do cenário C_1 : elas permanecem estáveis durante toda a transmissão. As diferenças quantitativas refletem as diferenças nas dimensões da rede sobreposta nos diferentes cenários (em média 850 nós no C_1 , 402 nós no C_2 e 450 nós no C_3).

Na literatura, Wu & Li [2007] também analisou a dinamicidade de propriedades estruturais da rede sobreposta do sistema UUSec. Entretanto, diferentemente do trabalho desenvolvido nesta dissertação, eles analisaram apenas o caminho mínimo médio, coeficiente de agrupamento e a reciprocidade da rede sobreposta, conforme apresentado anteriormente.

4.4 Sumário dos Resultados

Nesta dissertação foi possível observar na caracterização realizada sob a visão local dos nós que apesar dos diferentes perfis de centralidade, os nós tendem a permanecer com o mesmo perfil de centralidade ao longo de uma transmissão. Para identificar os perfis, as métricas de grau e *betweenness* apresentaram maior representatividade do que o *closeness*. As probabilidade de um nó permanecer em um mesmo perfil de centralidade são superiores as probabilidade de transição entre os perfis, e quando ocorre o nó possui maior probabilidade de rezudir sua centralidade do aumentá-la.

Entretanto, os nós mudam suas lista de parceiros analisando fotografias da rede não consecutivas. Apesar de um nó manter o mesmo perfil de centralidade, que faz referência ao número de conexões que ele possui, ele está variando com quem ele se

conecta. Isto de fato, é também visualizado pela variação do *pagerank* dos nós em fotografias não consecutivas. Já que o *pagerank* é uma métrica que não apenas avalia o número de conexões de um nó, mas também com quem ele se conecta. Essa variação na lista de parceiros pode ocorrer devida a definições do protocolo que pode, por exemplo, forçar a mudança de parcerias dos nós ao longo de uma transmissão. Também não foi encontrado uma variação muito clara de volume de *upload* e *download* entre nós de perfil de centralidade diferente. Um nó que possui Centralidade Baixa pode estar enviando e recebendo dados assim como um nó de outro perfil.

Já na perspectiva da estrutura topológica da rede, de uma maneira geral, ela permanece estável ao longo de uma transmissão de vídeo ao vivo, exibindo, portanto, pouco dinamismo na sua estrutura topológica. Isto combina com o fato dos nós permanecerem no mesmo perfil de centralidade ao longo do tempo, mostrando uma estabilidade em relação ao número de conexões dos nós. No entanto, as métricas de rede que dentro da visão parcial dependiam dos parceiros de um nó apresentaram variação ao longo do tempo, reforçando o resultado da visão local do nó que identifica uma significativa variação na lista de parceiros de um nó em fotografias não consecutivas da rede sobreposta.

Esses resultados não exibem um padrão de configuração do protocolo para este tipo de aplicação, no entanto, eles mostram que a rede sobreposta utilizada na transmissão de vídeo ao vivo é, além de estável, um ambiente controlado. Além disso, a visão parcial dos experimentos realizados no cenário C_1 apresentam uma limitação inerente do ambiente real de transmissão de vídeo ao vivo. Nesta caracterização, foi possível observar que métricas que dependem de informações dos parceiros de um nó apresentam estimativas dos valores medidos que podem subestimar os reais valores. No entanto, como neste dissertação foram realizadas validações dos resultados encontrados no cenário C_1 com os resultados dos cenários C_2 e C_3 que apresentam uma visão completa da rede, a visão parcial não impactou diretamente nas conclusões sobre as propriedades estruturais da rede sobreposta.

Capítulo 5

Conclusões e Trabalhos Futuros

Nesta dissertação foi apresentada uma caracterização de como as propriedades estruturais da rede sobreposta de um sistema de transmissão de vídeo ao vivo em redes P2P evoluem com o tempo. A caracterização foi feita sob duas perspectivas: visão local do nó e visão global da rede. Na visão local do nó, métricas de centralidade foram utilizadas para identificar perfis de centralidade. E na visão global da rede foram utilizadas métricas que capturam a estrutura da rede ao longo do tempo.

A aplicação SopCast foi utilizada para tal caracterização. Foram realizados um conjunto de experimentos com computadores do PlanetLab, que ao serem conectados ao SopCast passaram a atuar na rede de transmissão do canal como clientes comuns. Esses clientes SopCast monitoraram o tráfego de rede durante a transmissão do vídeo e armazenaram as informações em arquivos de *log*. Esses arquivos foram unificados em um único arquivo para reconstruir uma visão da rede sobreposta formada na transmissão de vídeo.

A rede sobreposta reconstruída foi utilizada para se obter medidas dos nós e da rede. Analisando os nós individualmente em cada experimento, os valores das medidas de grau, *betweenness* e *closeness* foram submetidos ao algoritmo de agrupamento *k-means*. Assim, foi possível identificar três perfis distintos de centralidade, chamados na dissertação de *Centralidade Alta* (CA), *Centralidade Intermediária* (CI) e *Centralidade Baixa* (CB).

Para cada perfil de centralidade, foram geradas as distribuições acumuladas complementares dos valores obtidos nas medidas das métricas. Os resultados de grau e *betweenness* apresentaram uma clara distinção de valores entre os perfis, o que não aconteceu com os valores de *closeness*. Para identificar as mudanças de perfil de centralidade de um nó ao longo de uma transmissão, foram gerados CBMGs (*Customer Behavior Model Graph*) para representar as transições de um nó entre perfis. Com os

CBMGs foi possível concluir que os nós possuem baixa probabilidade de transição entre os perfis de centralidade, ou seja, eles tendem a se manter no mesmo perfil de centralidade ao longo de uma transmissão. Além disso, quando eles realizam uma mudança no seu perfil de centralidade, existe uma maior probabilidade de diminuir sua centralidade do que aumentá-la.

Apesar dos nós tenderem a permanecer no mesmo perfil de centralidade, eles podem mudar a sua lista de parceiros ao longo de uma transmissão. Para analisar essa mudança na lista de parceiros, calculou-se a porcentagem de parceiros diferentes dos nós entre fotografias consecutivas e não consecutivas da rede sobreposta. Foi possível identificar que apesar da estabilidade nos perfis de centralidade ao longo de uma transmissão, os nós mudam significativamente sua lista de parceiros.

O pagerank foi utilizado para tentar identificar uma relação entre o perfil de centralidade de um nó e sua importância na transmissão de conteúdo na rede. Observou-se que nós de Centralidade Alta possuem valores de pagerank maiores que nós de Centralidade Intermediária e Baixa. Logo, a centralidade de um nó está correlacionada com sua importância na transmissão de conteúdo. Analisou-se também os volumes de *upload* e *download* gerados pelos nós em diferentes perfis de centralidade. Observou-se que, a despeito das diferenças de perfis, as distribuições dos volumes totais de *upload* e *download* são semelhantes.

Na visão global da estrutura da rede foi possível observar que o diâmetro, o caminho mínimo médio, o grau máximo, a reciprocidade, o coeficiente de agrupamento e assortatividade da rede tendem a permanecer estáveis ao longo de uma transmissão de vídeo ao vivo.

Os resultados encontrados da caracterização dos experimentos do cenário C_1 foram validados com os resultados da caracterização dos experimentos realizados com canal fechado nos cenários C_2 e C_3 . Na visão local de um nó, todos os resultados foram validados, com exceção da mudança na lista de parceiros dos nós entre fotografias consecutivas e não consecutivas. Essa divergência foi encontrada devida a visão parcial da rede dos experimentos no cenário C_1 em que a população dinâmica de nós do mundo real vista pelos *crawlers* varia ao longo de uma mesma transmissão. Na visão global da rede, as divergências encontradas na validação dos resultados se concentraram nas métricas de coeficiente de agrupamento e assortatividade da rede. A variação dos valores medidos dessas métricas nos cenários ocorreu também devido à visão parcial da rede existente no cenário C_1 , pois a visão dos nós do mundo real coletada pelos *crawlers* tende a subestimar o grau dos mesmos, e além disso, os *crawlers* também mudam suas listas de parceiros ao longo de uma transmissão conforme mencionado na visão local dos nós.

Esta dissertação apresenta resultados de caracterização que fornecem *insights* de como a rede sobreposta de um sistema de transmissão de vídeo ao vivo evolui ao longo de uma transmissão. Assim como argumentado por Silverston et al. [2009] e por outros autores [Wu & Li, 2007; Tang et al., 2009; Vu et al., 2010; Qiu et al., 2009; Hei et al., 2007; Silverston et al., 2007b], estes resultados de caracterização contribuem para aumentar o conhecimento sobre como funciona sistemas P2P de transmissão de vídeo ao vivo, principalmente, no que tange a dinâmica da rede sobreposta. Esse conhecimento pode fomentar e direcionar pesquisas futuras nesta área, além de subsidiar o desenvolvimento de simuladores de cargas de trabalho sintéticas reflitam o real dinamismo da estrutura da rede sobreposta para a transmissão de vídeo ao vivo em redes P2P.

Como trabalho futuro, sugere-se analisar a importância de cada atributo na caracterização. Não apenas considerando as três métricas utilizadas neste trabalho, como também o *pagerank* e os volumes de *download* e *upload*.

Ou ainda, analisar a reatividade da rede sobreposta com mudanças estratégicas na sua estrutura topológica. Como por exemplo, quanto tempo a rede leva para se reestruturar com a saída dos nós de perfil *CA* (ou *CI* ou *CB*)? Ou ainda, qual o impacto na transmissão do vídeo ao vivo (qualidade do conteúdo recebido na percepção do usuário) com a retirada dos nós de perfil *CA* (ou *CI* ou *CB*)? Tal análise, permitiria conhecer a atual robustez da rede sobreposta e conseqüentemente ajudaria a construir redes de transmissão de vídeo ao vivo mais robustas, já que o momento em que os *chunks* de um vídeo são recebidos pelos usuários é um dos fatores relevantes na qualidade de transmissão de vídeo ao vivo em redes P2P, pois os *chunks* devem estar sincronizados e devem ser recebidos em tempo real.

Ainda utilizando os conceitos dos perfis de centralidade, pode ser feita uma análise da definição do perfil de centralidade de um nó no momento em que ele se conecta na rede. Isto é, quais os fatores que levaram um nó entrar na rede com um determinado perfil de centralidade. Esta análise pode ser feita considerando as características de um nó durante o que foi chamado nesta dissertação de tempo inicial T . Pode ser investigado, neste caso, se a definição do perfil de centralidade de um nó quando ele passa a participar da rede sobreposta está relacionada com sua localização geográfica, banda disponível, ou ainda por exemplo com a atuação do servidor de boot (*bootstrap*) no fornecimento de informações de parceiros. Pode ser analisada a existência algum critério de seleção de parceiros em que devido a qualidade de um nó ser alta ele acabe sendo mais escolhido como parceiro e sendo conseqüentemente do perfil de Centralidade Alta. Sugere-se também verificar se as mudanças na lista de parceiros identificadas nesta dissertação ocorrem dentro de comunidades de nós. Isto pode indicar a presença

de grupos de distribuição de conteúdo na rede.

Além disso, sugere-se uma validação destes resultados com outros sistemas de transmissão de vídeo ao vivo, como por exemplo: PPLive, PPStream e UUSee. Nesta dissertação, tentamos realizar experimentos com estes sistemas. No entanto, devido ao atual ambiente que eles foram construídos (Windows) não foi possível realizar experimentos no PlanetLab, que é exclusivamente Linux, com proporções semelhantes ao que foi feito com SopCast. Se não com outra aplicação, pode ser por exemplo com transmissão de conteúdos especiais. Assim é possível analisar as características encontradas nesta caracterização em outros tipos de transmissão de vídeo ao vivo, por exemplo, transmissões de eventos mundiais como Copa do Mundo. Ou ainda, em ambiente onde a população tenha um crescimento muito dinâmico. Além disso, a rede sobreposta de transmissão de vídeo ao vivo observada apresenta características de estabilidade e de um ambiente controlado. Nos resultados dessa dissertação não foi apresentado uma proposta de configuração do protocolo de configuração deste tipo de aplicação. Sugere-se uma análise da qualidade de vídeo enviado/recebido por um nó neste ambiente de estabilidade e um ambiente instável.

Outras análises poderiam também ser feitas para entender melhor a relação dos perfis centralidade dos nós com o tráfego de rede gerado por eles. Como por exemplo, podem ser analisada as taxas de *upload* e *download* por fotografia da rede ao longo de uma transmissão de acordo com o perfis de centralidade dos nós, visando determinar se existe algum dinamismo no tráfego já que a estrutura topológica foi identificada nesta dissertação como sendo estável. Analisar uma possível correlação entre as métricas de grau e *betweenness*.

Apêndice A

Resultados em Canal Fechado Com *Churn* - Cenário C_2

Na perspectiva da visão local de um nó, assim como no cenário C_1 , os resultados encontrados no cenário C_2 também mostram três perfis de centralidade para todos os experimentos conforme pode ser visto na Tabela A.1. Além disso, também como em C_1 , o grau e o *betweenness* dos nós no cenário C_2 (Figura A.1) apresentam variações significativas entre os perfis de centralidade: os nós de perfil *CB* possuem os menores valores, os nós de perfil *CA* têm os maiores valores, enquanto que os nós com perfil *CI* possuem valores de grau e *betweenness* intermediários entre *CA* e *CB*. No entanto, o *closeness* também não exibe uma distinção clara entre os perfis de centralidade.

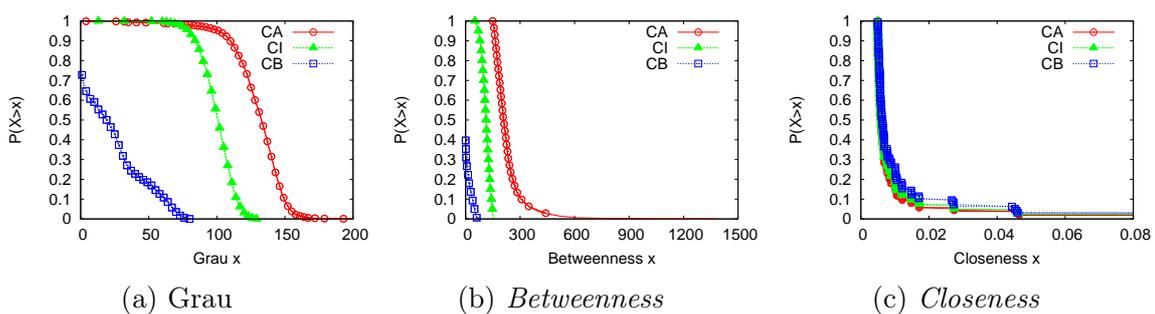


Figura A.1: Métricas dos nós por perfil de centralidade (Experimento 1, Cenário C_2)

Também foram gerados CBMGs dos experimentos realizados neste cenário. A Figura A.2a mostra o CBMG para o Experimento 1 a título de exemplo dos demais experimentos e a Figura A.2b apresenta o CBMG médio considerando todos os experimentos.

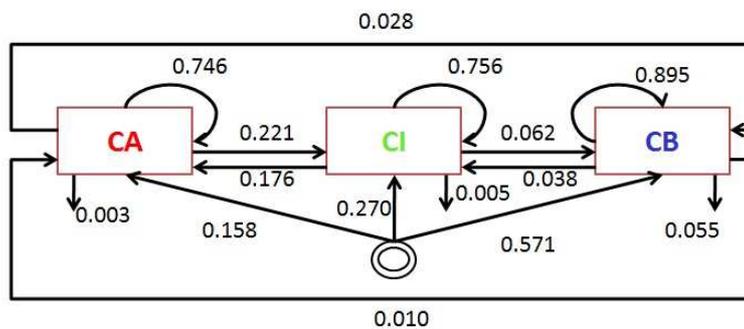
Tabela A.1: Perfil de centralidade dos nós em cada experimento (Cenário C_2)

		# Nós	Perfil	% de Nós	Grau		<i>Betweenness</i>		<i>Closeness</i>	
					Média	CV	Média	CV	Média	CV
Experimento	1	412	CA	21.35%	130.61	0.14	229.15	0.36	0.015	3.41
			CI	49.27%	100.34	0.12	111.33	0.20	0.016	3.26
			CB	29.36%	23.70	0.97	8.59	1.86	0.018	2.93
	2	421	CA	18.09%	133.50	0.15	232.05	0.39	0.014	3.46
			CI	56.19%	102.56	0.11	112.26	0.20	0.016	3.28
			CB	25.95%	24.76	0.92	8.55	1.76	0.018	2.99
	3	385	CA	14.80%	126.57	0.14	233.29	0.39	0.012	3.29
			CI	51.94%	96.84	0.12	111.80	0.21	0.018	3.11
			CB	33.24%	29.03	0.90	14.67	1.47	0.017	2.89
	4	383	CA	17.23%	124.48	0.13	195.89	0.32	0.011	3.48
			CI	52.74%	96.05	0.11	97.63	0.20	0.016	3.33
			CB	30.02%	27.57	0.84	9.78	1.60	0.019	2.78
	5	412	CA	17.23%	134.38	0.13	216.00	0.32	0.014	3.52
			CI	51.25%	104.54	0.12	108.32	0.22	0.017	3.22
			CB	31.55%	29.38	0.85	11.63	1.56	0.018	3.09
	6	404	CA	18.81%	135.45	0.13	227.70	0.35	0.012	3.52
			CI	50.49%	102.90	0.11	109.61	0.20	0.018	3.16
			CB	30.69%	27.82	0.87	11.31	1.63	0.018	2.99

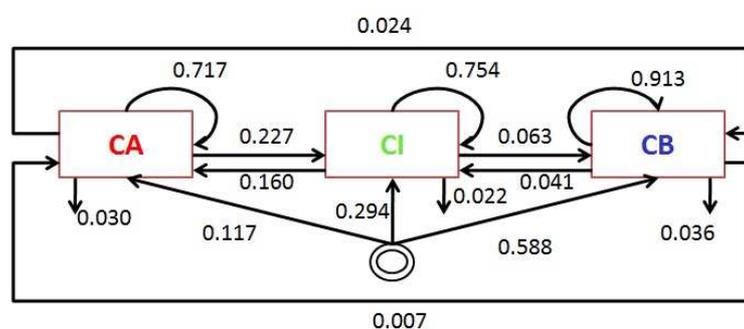
As diferenças nas listas de parceiros dos nós por perfil de centralidade também foi analisada. A distribuição das diferenças de parceiros de um nó em fotografias consecutivas para o Experimento 1 e para os todos experimentos são apresentadas nas Figuras A.3a e A.3b, respectivamente. A distribuição das diferenças na lista de parceiros de um nó em fotografias não consecutivas é mostrada na Figura A.4a considerando um salto de 3 fotografias, ou seja, t e $t + 3$. Já a Figura A.4b apresenta a distribuição das diferenças de parceiros de nó encontrada com um salto entre as fotografias t e $t + 5$.

As Figuras A.5a e A.5b apresentam as distribuições do *pagerank* dos nós por perfil de centralidade do Experimento 1 e para todos os experimentos deste cenário. Os valores de *pagerank* foram sumarizados na Tabela A.2.

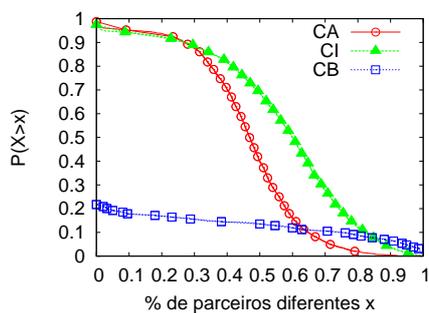
Já na perspectiva da visão global da rede todos os valores medidos das métricas de diâmetro, caminho mínimo médio, grau máximo, coeficiente de agrupamento, assortatividade e reciprocidade estão sumarizados na Tabela A.3. A Figura A.6a mostra o diâmetro da rede ao longo do tempo do Experimento 1. O caminho mínimo médio é mostrado na Figura A.6b, o grau máximo na Figura A.6c e o coeficiente de agrupamento na Figura A.6d. Por fim, a assortatividade e a reciprocidade da rede apresentadas por fotografia do Experimento 1 são mostradas nas Figuras A.6e e A.6f, respectivamente.



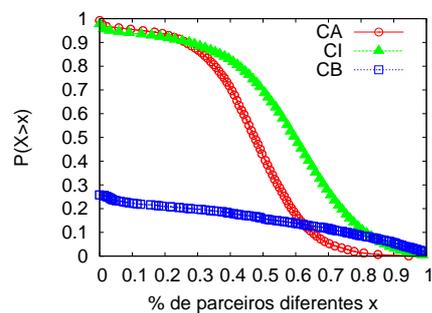
(a) Experimento 1



(b) Todos os experimentos

Figura A.2: CBMG - *Customer Behavior Model Graph* (Cenário C_2)

(a) Experimento 1



(b) Todos os experimentos

Figura A.3: Distribuição da diferença de parceiros de um nó em fotografias consecutivas (Cenário C_2)

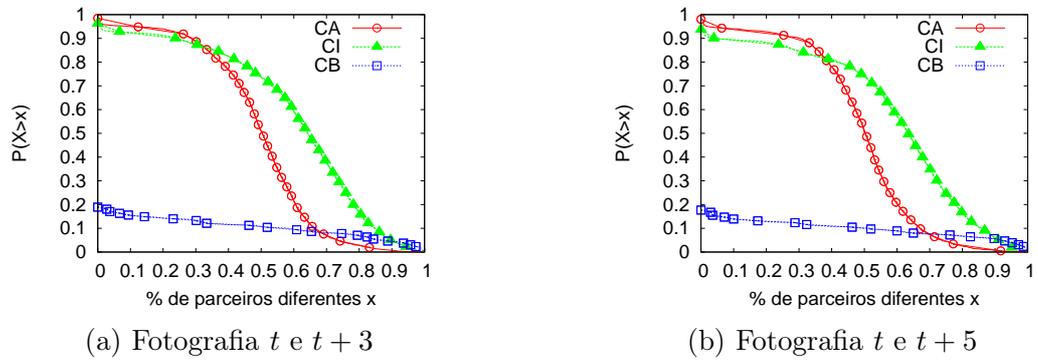


Figura A.4: Distribuição da diferença de parceiros de um nó em fotografias não-consecutivas (Experimento1, Cenário C_2)

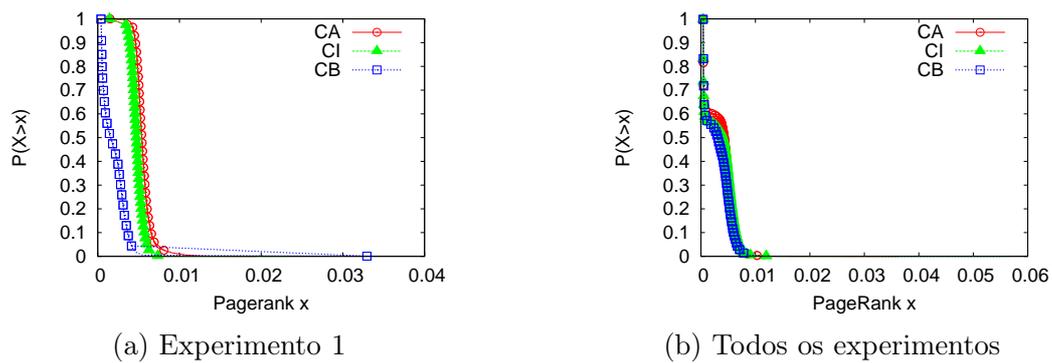


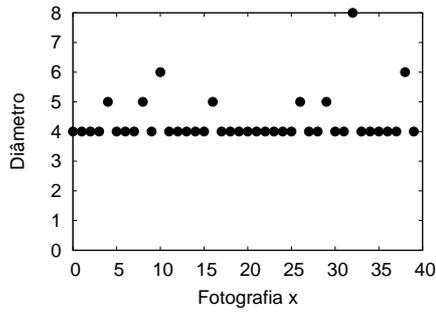
Figura A.5: *PageRank* dos nós por perfil de centralidade (Cenário C_2)

Tabela A.2: *Pagerank* dos nós por perfil de centralidade em cada experimento (Cenário C_2)

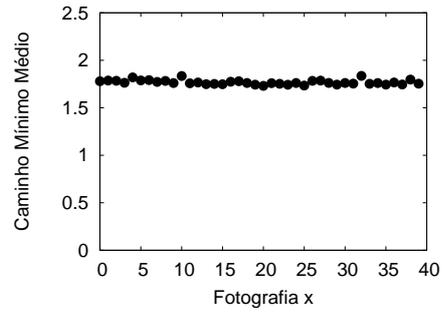
		Perfil	<i>Pagerank</i>	
			Média	CV
Experimento	1	CA	0.0136	0.13
		CI	0.0078	0.35
		CB	0.0046	0.34
	2	CA	0.0055	0.20
		CI	0.0048	0.19
		CB	0.0042	0.38
	3	CA	0.0059	0.21
		CI	0.0052	0.18
		CB	0.0024	0.88
	4	CA	0.0061	0.20
		CI	0.0058	0.23
		CB	0.0026	1.24
	5	CA	0.0056	0.21
		CI	0.0050	0.18
		CB	0.0023	1.05
	6	CA	0.0055	0.22
		CI	0.0049	0.19
		CB	0.0021	0.89

Tabela A.3: Medidas das métricas da rede (Cenário C_2)

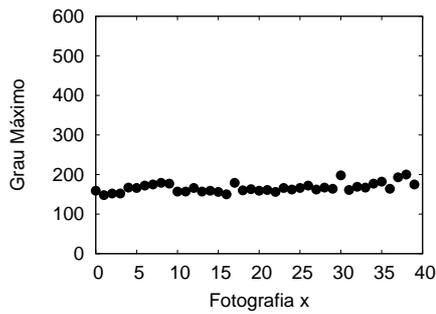
		Experimento					
		1	2	3	4	5	6
Diâmetro	Média	4.32	4.25	4.12	4.17	4.12	4.05
	CV	0.18	0.12	0.08	0.13	0.08	0.05
Caminho Mínimo Médio	Média	1.76	1.75	1.76	1.73	1.74	1.74
	CV	0.013	0.01	0.01	0.01	0.011	0.012
Coeficiente de Agrupamento	Média	0.38	0.39	0.38	0.4	0.4	0.39
	CV	0.02	0.03	0.05	0.06	0.04	0.04
Grau Máximo	Média	166.8	175.65	155.55	153.8	160.3	165.7
	CV	0.07	0.06	0.07	0.05	0.06	0.04
Reciprocidade	Média	0.637	0.629	0.651	0.658	0.632	0.63
	CV	0.04	0.03	0.04	0.04	0.06	0.05
Assortatividade	Média	-0.007	-0.011	-0.014	-0.008	-0.005	-0.014
	CV	2.59	1.33	1.50	2.15	3.33	1.51



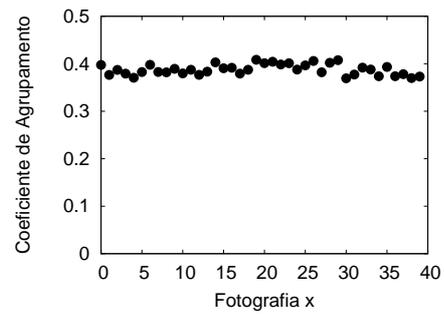
(a) Diâmetro



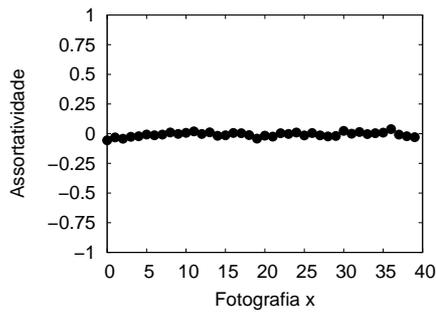
(b) Caminho Mínimo Médio



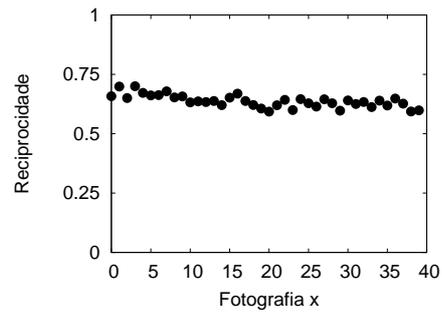
(c) Grau Máximo



(d) Coeficiente de Agrupamento



(e) Assortatividade



(f) Reciprocidade

Figura A.6: Métricas de rede ao longo do tempo (Experimento 1, Cenário C_2)

Apêndice B

Resultados em Canal Fechado Sem Churn - Cenário C_3

Os resultados da caracterização na visão local do nó dos experimentos do cenário C_3 , também mostram três perfis de centralidade para todos os experimentos (Tabela B.1), assim como nos cenários C_1 e C_2 . Ocasionalmente, a porcentagem de nós de perfil CI é maior do que nós com perfil CB .

Assim como em C_1 e C_2 , o grau e *betweenness* dos nós no cenário C_3 (Figura B.1) possuem variações significativas entre os perfis de centralidade: os nós possuem valores do maior para o menor de grau e *betweenness* nos perfis CA , CI e CB , nesta ordem. O *closeness* também não exibe uma distinção clara entre os perfis de centralidade. Além disso, neste cenário sem *churn* (C_3), os valores medidos foram mais altos que os experimentos com *churn* (C_1 e C_2), já que a população de nós da rede permanece estável ao longo de toda a transmissão.

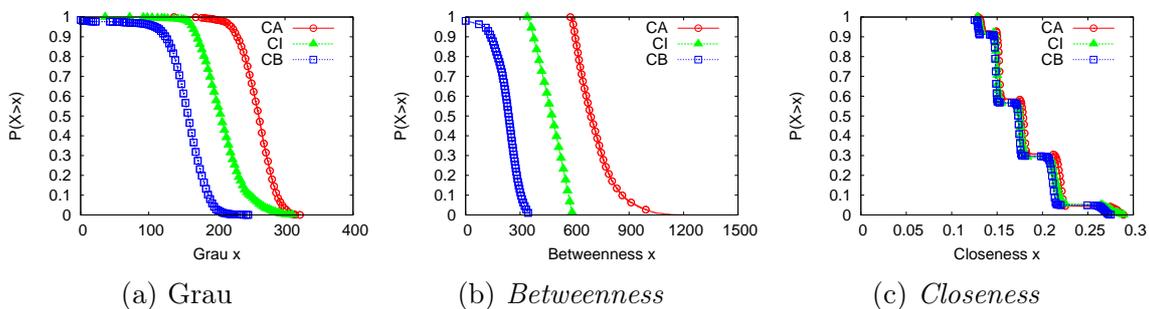


Figura B.1: Métricas dos nós por perfil de centralidade (Experimento 1, Cenário C_3)

Para entender como um nó muda seu perfil de centralidade ao longo de uma transmissão, foram gerados CBMGs assim como foi feito para os experimentos dos

Tabela B.1: Perfil de centralidade dos nós em cada experimento (Cenário C_3)

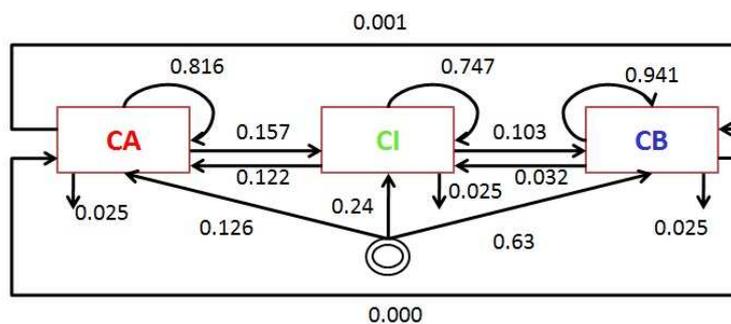
		# Nós	Perfil	% de Nós	Grau		<i>Betweenness</i>		<i>Closeness</i>	
					Média	CV	Média	CV	Média	CV
Experimento	1	450	CA	14.22%	260.32	0.08	715.38	0.15	0.181	0.19
			CI	21.11%	207.91	0.15	472.70	0.14	0.177	0.20
			CB	64.66%	154.77	0.21	227.15	0.30	0.173	0.22
	2	451	CA	29.93%	219.84	0.13	648.85	0.21	0.304	0.18
			CI	57.42%	155.27	0.12	288.87	0.19	0.292	0.17
			CB	12.63%	115.98	0.26	150.00	0.33	0.283	0.21
	3	450	CA	16.44%	240.84	0.13	627.77	0.20	0.231	0.16
			CI	21.77%	174.62	0.11	298.83	0.20	0.225	0.16
			CB	61.77%	126.42	0.28	154.23	0.35	0.213	0.23
	4	452	CA	12.61%	254.93	0.19	732.24	0.21	0.217	0.25
			CI	28.76%	202.73	0.14	465.23	0.13	0.2	0.19
			CB	58.62%	151.51	0.22	228.48	0.33	0.2	0.23
	5	449	CA	26.06%	256.27	0.17	656.01	0.24	0.289	0.18
			CI	58.13%	178.03	0.14	295.59	0.22	0.275	0.15
			CB	15.81%	137.99	0.23	152.48	0.30	0.272	0.19
	6	449	CA	22.94%	262.23	0.12	659.66	0.20	0.229	0.21
			CI	60.36%	184.93	0.11	296.54	0.22	0.224	0.21
			CB	16.7%	133.65	0.24	142.65	0.34	0.212	0.24

cenários C_1 e C_2 . As Figura B.2a e B.2b apresentam os CBMGs gerados dos experimentos realizados neste cenário para o Experimento 1 e o CBMG médio para todos os experimentos, respectivamente.

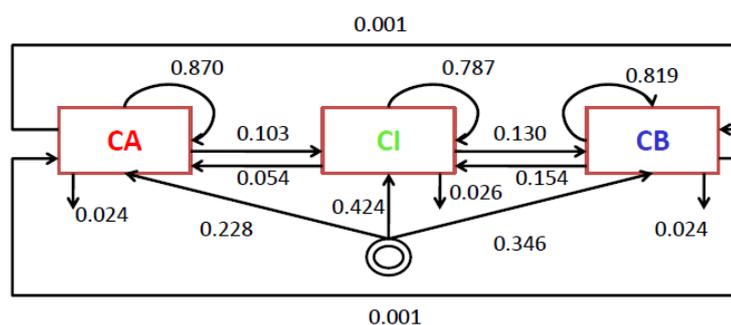
A diferença na lista de parceiros de um nó também foi analisada para os experimentos do cenário C_3 . Na Figura B.3a é mostrada a distribuição da diferença da lista de parceiros de um nó para o Experimento 1 e a distribuição média de todos os experimentos na Figura B.3b considerando fotografias consecutivas da rede. Já na perspectiva de fotografias não consecutivas, a Figura B.4a apresenta a diferença de parceiros encontrada no Experimento 1 considerando um salto de 3 fotografias, ou seja, t e $t + 3$. E a Figura B.4b apresenta a distribuição da diferença de parceiros considerando as fotografias t e $t + 5$.

A distribuição dos valores de *pagerank* dos nós por perfil de centralidade do Experimento 1 deste cenário é apresentada na Figura B.5a. A distribuição dos valores de *pagerank* para todos os experimentos é apresentada na Figura B.5b. Os valores medidos foram sumarizados na Tabela B.2.

Os resultados da caracterização da rede sobreposta dos experimentos realizados no cenário C_3 também foram feitos considerando as métricas de diâmetro, caminho mínimo médio, coeficiente de agrupamento, assortatividade, reciprocidade e grau máximo. A Tabela B.3 sumariza os valores medidos na rede sobreposta dos



(a) Experimento 1

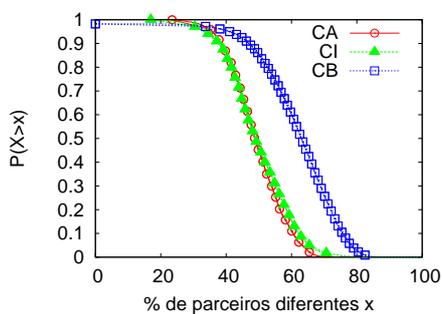


(b) Todos os experimentos

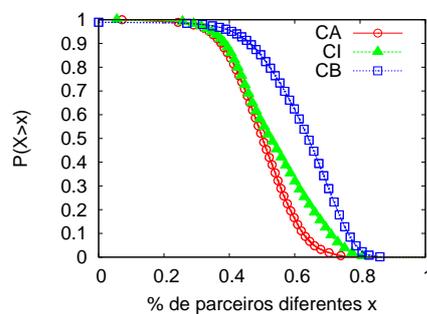
Figura B.2: CBMG - *Customer Behavior Model Graph* (Cenário C_3)

experimentos realizados neste cenário.

Na perspectiva da visão global da rede, neste cenário, as métricas de rede — diâmetro (Figura B.6a), caminho mínimo médio (Figura B.6b), grau máximo (Figura B.6c), coeficiente de agrupamento (Figura B.6d), assortatividade (Figura B.6e) e reciprocidade (Figura B.6f) — também são apresentadas por fotografia da rede sobreposta.



(a) Experimento 1



(b) Todos os experimentos

Figura B.3: Distribuição da diferença de parceiros de um nó em fotografias consecutivas (Cenário C_3)

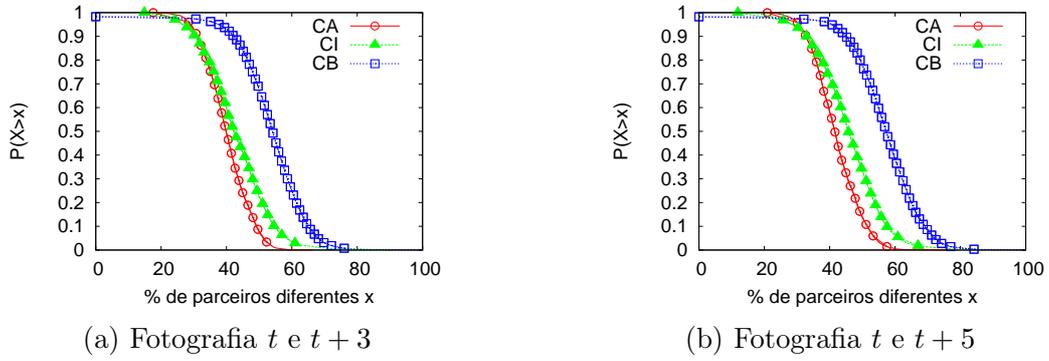


Figura B.4: Distribuição da diferença de parceiros de um nó em fotografias não-consecutivas (Experimento 1, Cenário C_3)

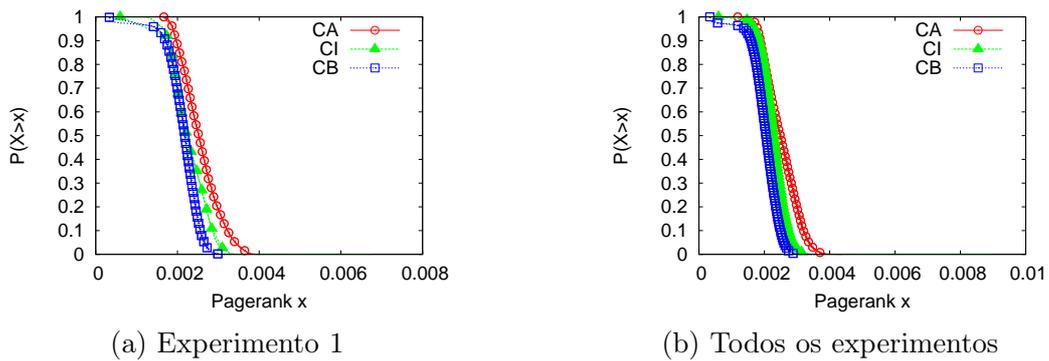


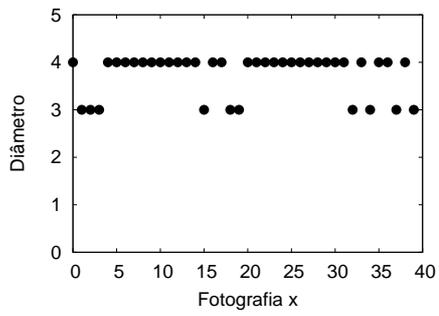
Figura B.5: *Pagerank* dos nós por perfil de centralidade (Cenário C_3)

Tabela B.2: *Pagerank* dos nós por perfil de centralidade em cada experimento (Cenário C_3)

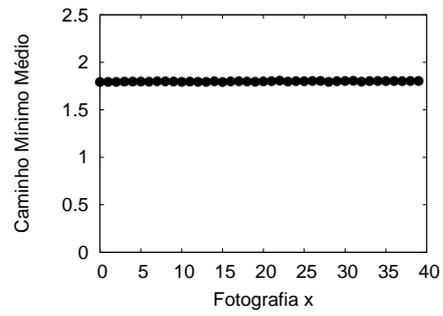
		Perfil	<i>Pagerank</i>	
			Média	CV
Experimento	1	CA	0.0025	0.18
		CI	0.0022	0.19
		CB	0.0021	0.20
	2	CA	0.0024	0.20
		CI	0.0022	0.13
		CB	0.0017	0.21
	3	CA	0.0025	0.18
		CI	0.0022	0.20
		CB	0.0021	0.19
	4	CA	0.0026	0.19
		CI	0.0022	0.19
		CB	0.0020	0.20
	5	CA	0.0025	0.22
		CI	0.0022	0.13
		CB	0.0018	0.18
	6	CA	0.0024	0.20
		CI	0.0022	0.14
		CB	0.0017	0.21

Tabela B.3: Medidas das métricas da rede (Cenário C_3)

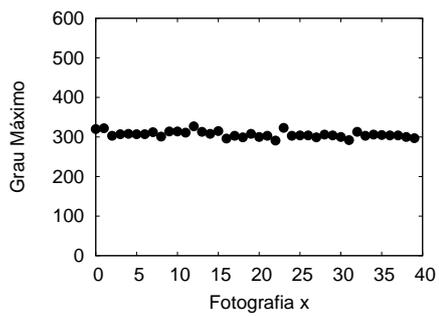
		Experimento					
		1	2	3	4	5	6
Diâmetro	Média	3.75	3.7	3.87	3.82	3.42	3.72
	CV	0.11	0.12	0.08	0.10	0.14	0.12
Caminho Mínimo Médio	Média	1.79	1.81	1.8	1.8	1.79	1.78
	CV	0.002	0.005	0.001	0.010	0.012	0.002
Coeficiente de Agrupamento	Média	0.25	0.24	0.26	0.26	0.26	0.26
	CV	0.02	0.04	0.02	0.22	0.15	0.02
Grau Máximo	Média	306.4	287.17	308.65	351.87	354.7	254.52
	CV	0.02	0.05	0.02	0.14	0.05	0.02
Reciprocidade	Média	0.385	0.432	0.381	0.426	0.411	0.362
	CV	0.02	0.03	0.01	0.12	0.12	0.02
Assortatividade	Média	-0.231	-0.233	-0.195	-0.167	-0.236	-0.241
	CV	0.10	0.09	0.08	0.22	0.13	0.12



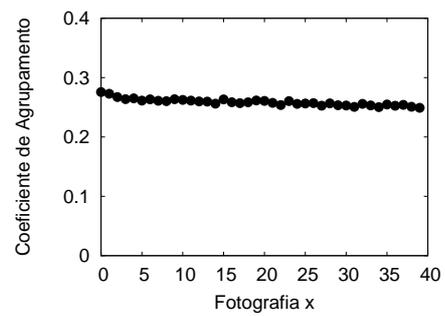
(a) Diâmetro



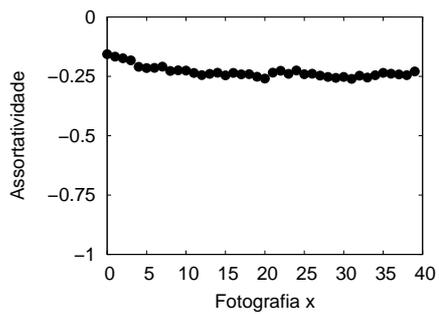
(b) Caminho Mínimo Médio



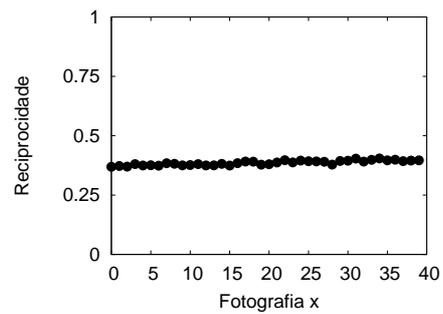
(c) Grau Máximo



(d) Coeficiente de Agrupamento



(e) Assortatividade



(f) Reciprocidade

Figura B.6: Métricas de rede ao longo do tempo (Experimento 1, Cenário C_3)

Referências Bibliográficas

- Acer, U. G.; Drineas, P. & Abouzeid, A. A. (2011). Connectivity in Time-Graphs. *Pervasive and Mobile Computing*, 7(2):160–171.
- Alexa.com (2012). Alexa The Web Information Company.
- Benevenuto, F. (2010). *Redes Sociais Online: Técnicas de Coleta, Abordagens de Medição e Desafios Futuros*, capítulo 2. Simpósio Brasileiro de Sistemas Multimídia e Web.
- Bermudez, I.; Mellia, M. & Meo, M. (2011). Passive Characterization of Sopcast Usage in Residential ISPs. Em *Proceedings of the Peer-to-Peer Computing*.
- Bonacich, P. & Lloyd, P. (2001). Eigenvector-like Measures of Centrality for Asymmetric Relations. *Social Networks*, 23(3):191–201.
- Borges, A. (2010). *Transmissão de Mídia Contínua ao Vivo em P2P: Modelagem, Caracterização e Implementação de Mecanismos de Resiliência a Ataques*. PhD in Ciência da computação, Instituto de Ciências Exatas – Universidade Federal de Minas Gerais, Icx - UFMG, Belo Horizonte, Minas Gerais, Brasil.
- Borges, A.; Almeida, J. & Campos, S. (2008). Combate a Poluição em Sistemas P2P de Mídia Contínua ao Vivo. Em *Proceedings of the Simpósio Brasileiro de Redes de Computadores*.
- Borges, A.; Gomes, P.; Nacif, J.; Mantini, R.; Almeida, J. M. & Campos, S. (2012). Characterizing SopCast Client Behavior. *Computer Communications*, 35(8):1004–1016.
- Brin, S. & Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks ISDN System*, 30(1-7):107–117.

- Chirita, P.-A.; Diederich, J. & Nejdl, W. (2005). MailRank: Using Ranking for Spam Detection. Em *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*.
- Chun, B.; Culler, D.; Roscoe, T.; Bavier, A.; Peterson, L.; Wawrzoniak, M. & Bowman, M. (2003). PlanetLab: An Overlay Testbed for Broad-Coverage Services. *SIGCOMM Computer Communication Review*, 33(3):3–12.
- Easley, D. & Kleinberg, J. (2010). *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press.
- Fallica, B.; Lu, Y.; Kuipers, F.; Kooij, R. & Miegheem, P. V. (2008). On the Quality of Experience of SopCast. Em *Proceedings of the 2008 Second International Conference on Next Generation Mobile Applications, Services and Technologies*.
- Freeman, L. (1979). Centrality in Social Networks: Conceptual Clarification. *Social Networks*, 1(3):215–239.
- Gomes, L. H.; Almeida, V.; Bettencourt, L.; Duarte, F. & Almeida, J. (2009). Quantifying Social and Opportunistic Behavior in Email Networks. Em *Proceedings of the Advances in Complex Systems*.
- Gonçalves, K.; Borges, A.; Almeida, J.; Silva, A.; Marques-Neto, H. & Campos, S. (2011). Caracterização das Propriedades Dinâmicas da Estrutura da Rede Sobreposta do SopCast. Em *Proceedings of the X Workshop em Desempenho de Sistemas Computacionais de Comunicação*.
- Gonçalves, K.; Vieira, A.; Almeida, J.; da Silva, A.; Marques-Neto, H. & Campos, S. (2012). Characterizing Dynamic Properties of the SopCast Overlay Network. Em *Proceedings of the 20th Euromicro International Conference on Parallel, Distributed and Network-Based*.
- Google (2010). Google Trends.
- Hei, X.; Liang, C.; Liang, J.; Liu, Y. & Ross, K. (2007). A Measurement Study of a Large-Scale P2P IPTV System. *IEEE Transactions on Multimedia*, 9(8):1672–1687.
- Hei, X.; Liang, C.; Liang, J.; Liu, Y. & Ross, K. W. (2006). Insights into PPLive: A Measurement Study of a Large-Scale P2P IPTV System. Em *Proceedings of the IPTV Workshop, International World Wide Web Conference*.

- Hei, X.; Liu, Y. & Ross, K. (2008). IPTV over P2P Streaming Networks: the Mesh-pull Approach. *IEEE Communications Magazine*, 46(2):86–92.
- Huang, Y.; Fu, T. Z.; Chiu, D.-M.; Lui, J. C. & Huang, C. (2008). Challenges, Design and Analysis of a Large-Scale P2P-Vod System. Em *Proceedings of the ACM SIGCOMM 2008 Conference on Data Communication*.
- Iliofotou, M.; Faloutsos, M. & Mitzenmacher, M. (2009). Exploiting Dynamicity in Graph-based Traffic Analysis: Techniques and Applications. Em *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*.
- Kitsak, M.; Gallos, L.; Havlin, S.; Liljeros, F.; Muchnik, L.; Stanley, H. & Makse, H. (2010). Identification of Influential Spreaders in Complex Networks. *Nature Physics*, 6(11):888–893.
- Lederer, S.; Wang, Y. & Gao, J. (2009). Connectivity-Based Localization of Large-Scale Sensor Networks with Complex Shape. *ACM Transactions on Sensor Networks*, 5(4):31:1–31:32.
- Liang, W.; Bi, J.; Wu, R.; Li, Z. & Li, C. (2009). On Characterizing PPStream: Measurement and Analysis of P2P IPTV under Large-scale Broadcasting. Em *Proceedings of the 28th IEEE Conference on Global Telecommunications*.
- Liu, Y.; Guo, Y. & Liang, C. (2008). A Survey on Peer-to-Peer Video Streaming Systems. *Peer-to-Peer Networking and Applications*, 1(1):18–28.
- Lua, E. K.; Crowcroft, J.; Pias, M.; Sharma, R. & Lim, S. (2005). A Survey and Comparison of Peer-to-Peer Overlay Network Schemes. *IEEE Communications Surveys and Tutorials*, 7(2):72–93.
- Manning, C. D.; Raghavan, P. & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
- Menasce, D. & Almeida, V. (2007). *Scaling for E-Business: Technologies, Models, Performance, and Capacity Planning*. Prentice Hall.
- Newman, M. E. J. (2002). Assortative Mixing in Networks. *Physical Review Letters*, 89(20):208–701.
- Oliveira, J.; Vieira, A. B.; de Carvalho Gomes, P. & de Aguiar Campos, S. V. (2010). Centralidade em Redes P2P de Transmissão ao Vivo. Em *Proceedings of the VI Workshop de Redes Dinâmicas e Sistemas P2P*.

- Pathak, A.; Pucha, H.; Zhang, Y.; Hu, Y. C. & Mao, Z. M. (2008). A Measurement Study of Internet Delay Asymmetry. Em *Proceedings of the Passive and Active Measurement Conference*.
- Peterson, L.; Bavier, A.; Fiuczynski, M. E. & Muir, S. (2006). Experiences Building PlanetLab. Em *Proceedings of the 7th Symposium on Operating Systems Design and Implementation*.
- Pourebrahimi, B.; Bertels, K. & Vassiliadis, S. (2005). A Survey of Peer-to-Peer Networks. Em *Proceedings of the 16th Annual Workshop on Circuits, Systems and Signal Processing*.
- Qiu, T.; Ge, Z.; Lee, S.; Wang, J.; Xu, J. & Zhao, Q. (2009). Modeling User Activities in a Large IPTV System. Em *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference*.
- Sankaralingam, K.; Sethumadhavan, S. & Browne, J. C. (2003). Distributed Pagerank for P2P Systems. Em *Proceedings of the 12th IEEE International Symposium on High Performance Distributed Computing*.
- Scherrer, A.; Borgnat, P.; Fleury, E.; Guillaume, J. L. & Robardet, C. (2008). Description and Simulation of Dynamic Mobility Networks. *Computer Networks*, 52(15):2842–2858.
- Sentinelli, A.; Marfia, G.; Gerla, M.; Kleinrock, L. & Tewari, S. (2007). Will IPTV Ride The Peer-to-Peer Stream? [Peer-to-Peer Multimedia Streaming]. *IEEE Communications Magazine*, 45(6):86–92.
- Shi, S.; Yu, J.; Yang, G. & Wang, D. (2003). Distributed Page Ranking in Structured P2P Networks. Em *Proceedings of the International Conference on Parallel Processing*.
- Silverston, T.; Fourmaux, O.; Botta, A.; Dainotti, A.; Pescapé, A.; Ventre, G. & Salamatian, K. (2009). Traffic Analysis of Peer-to-Peer IPTV Communities. *Computer Networks*, 53(4):470–484.
- Silverston, T.; Fourmaux, O. & Salamatian, K. (2007a). Characterization of P2P IPTV Traffic: Scaling Analysis. *Analysis*, 67(2):95–172.
- Silverston, T.; Fourmaux, O.; Salamatian, K. & Cho, K. (2010). On Fairness and Locality in P2P-TV Through Large-Scale Measurement Experiment. Em *Proceedings of the IEEE Communications Society*.

- Silverston, T.; Pierre, U.; Paris, M. C. & Fourmaux, O. (2006). P2P IPTV Measurement: A Case Study of TVants. Em *Proceedings of the 2nd Conference on Future Networking Technologies*.
- Silverston, T.; Pierre, U.; Paris, M. C.; Fourmaux, O.; Pierre, U. & Paris, M. C. (2007b). Measuring P2P IPTV Systems. Em *Proceedings of the International Workshop on Network and Operating Systems Support for Digital Audio e Video*.
- Stutzbach, D. & Rejaie, R. (2006). Understanding Churn in Peer-to-Peer Networks. Em *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*.
- Stutzbach, D.; Rejaie, R. & Sen, S. (2008). Characterizing Unstructured Overlay Topologies in Modern P2P File-Sharing Systems. *IEEE/ACM Transactions on Networking*, 16(2):267–280.
- Tang, J.; Lou, T. & Kleinberg, J. (2012). Inferring Social Ties Across Heterogenous Networks. Em *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*.
- Tang, S.; Lu, Y.; Hernández, J. M.; Kuipers, F. & Miegheem, P. (2009). Topology Dynamics in a P2PTV Network. Em *Proceedings of the 8th International IFIP-TC 6 Networking Conference*.
- Tang, W.; Fu, Y.; Cherkasova, L. & Vahdat, A. (2003). MediSyn: A Synthetic Streaming Media Service Workload Generator. Em *Proceedings of the 13th International Workshop on Network and Operating Systems Support for Digital Audio and Video*.
- Vu, L.; Gupta, I.; Liang, J. & Nahrstedt, K. (2007). Measurement and Modeling of a Large-Scale Overlay for Multimedia Streaming. Em *Proceedings of the Fourth International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, Workshops*.
- Vu, L.; Gupta, I.; Nahrstedt, K. & Liang, J. (2010). Understanding Overlay Characteristics of a Large-Scale Peer-to-Peer IPTV System. *ACM Transactions on Multimedia Computing, Communications and Applications*, 6(4):31:1–31:24.
- Wan, S. J.; Wong, S. K. M. & Prusinkiewicz, P. (1988). An Algorithm for Multidimensional Data Clustering. *ACM Transactions on Mathematical Software*, 14(2):153–162.

- Watts, D. J. (1999). Networks, Dynamics, and the Small-World Phenomenon. *The American Journal of Sociology*, 105(2):493–527.
- Watts, D. J. & Strogatz, S. H. (1998). Collective Dynamics of 'Small-World' Networks. *Nature*, 393(6684):440–442.
- Weng, J.; Lim, E.-P.; Jiang, J. & He, Q. (2010). TwitterRank: Finding Topic-Sensitive Influential Twitterers. Em *Proceedings of the Third ACM International Conference on Web Search and Data Mining*.
- Wu, C. & Li, B. (2007). Magellan: Charting Large-Scale Peer-to-Peer Live Streaming Topologies. Em *Proceedings of the 27th International Conference on Distributed Computing Systems (ICDCS)*.
- Xiao, Z. & Ye, F. (2008). New Insights on Internet Streaming and IPTV. Em *Proceedings of the 2008 International Conference on Content-Based Image and Video Retrieval*.
- Xu, J.; Marshall, B.; Kaza, S. & Chen, H. (2004). Analyzing and Visualizing Criminal Network Dynamics: A Case Study. Em *Proceedings of the Intelligence and Security Informatics*.
- Yang, B. & Garcia-Molina, H. (2003). Designing a Super-Peer Network. *International Conference on Data Engineering*, 0(1):49–60.