

Sobre dois fenômenos em redes P2P do tipo BitTorrent

Fabício Murai Ferreira¹
Orientadores: Daniel R. Figueiredo¹
Edmundo A. de Souza e Silva¹

¹Programa de Engenharia de Sistemas e Computação – COPPE
Universidade Federal do Rio de Janeiro – Rio de Janeiro, Brasil

{fabricio,daniel,edmundo}@land.ufrj.br

Palavras-chave: avaliação de desempenho, aplicações peer-to-peer, modelagem

Resumo. *Este artigo resume a dissertação de mestrado de Fabrício Murai¹, aprovada pelo PESC - COPPE/UFRJ em abril de 2011.*

1. Motivação e relevância do tema de dissertação

O paradigma P2P possibilitou o surgimento de aplicativos de larga escala para a Internet, pois oferece diversas vantagens incluindo alta escalabilidade e disponibilidade a baixo custo (número reduzido de servidores e baixos requisitos de banda, por exemplo). Esta escalabilidade advém do fato de que a arquitetura P2P aproveita os recursos dos peers (usuários) para o funcionamento do sistema. Desta forma, usuários que chegam na rede trazem consigo, além da demanda pelo serviço oferecido, recursos que serão utilizados pelo sistema para servir outros peers. Por exemplo, em aplicações P2P para compartilhamento de arquivo, um peer obtém pedaços do arquivo de outros usuários ao mesmo tempo que oferece os pedaços já recuperados àqueles que ainda não os possuem.

Em contraposição aos sistemas P2P, os serviços baseados em arquiteturas cliente-servidor necessitam de forte investimento em infraestrutura para alcançar escalabilidade e disponibilidade. Como exemplo de serviço baseado nessa arquitetura temos o YouTube, que a fim de suprir a demanda de mais de 4 bilhões de exibições de vídeo por dia², faz uso de uma CDN (Content Distribution Network) composta por milhares de servidores espalhados pelo mundo. Devido ao elevado custo dessa solução, grandes provedores de conteúdo como CNN têm interesse em alternativas baseadas no paradigma P2P [arstechnica 2009].

Uma das aplicações P2P mais populares e também responsável por grande parte do tráfego nas redes é o BitTorrent (BT), devendo grande parte desse sucesso ao seu bom desempenho no compartilhamento e recuperação de arquivos grandes, como filmes. Usuários do BT interessados em obter ou disseminar um mesmo conteúdo (um ou mais arquivos) formam uma rede, denominada swarm. O conteúdo a ser compartilhado é dividido em vários pedaços, que são disseminados entre os peers conectados ao swarm. Os peers podem ser de dois tipos: seeds, que são usuários que possuem uma cópia completa do conteúdo e fazem apenas upload para outros peers; e leechers, usuários que ainda não recuperaram todo o conteúdo e estão fazendo download e upload de pedaços para outros peers. Cada peer interage somente com sua vizinhança, isto é, um subconjunto dos peers

¹Atualmente, Fabrício Murai é aluno de doutorado no Departamento de Ciência da Computação na Universidade de Massachusetts em Amherst, USA.

²Youtube statistics: www.youtube.com/t/press_statistics

do swarm. Em particular, um peer indica que pode fazer upload para um dado vizinho enviando uma mensagem de *unchoke*. Enquanto a maior parte dos unchokes são concedidos baseados em tit-for-tat (isto é, são dados aos vizinhos que mais contribuíram com aquele peer), uma pequena fração conhecida como unchokes otimistas, é concedida de maneira aleatória.

Embora o BT seja baseado em mecanismos distribuídos e que cada peer só possua informação local (i.e., sobre sua vizinhança), ele é conhecido por ser muito eficiente. Entretanto, o BT foi desenvolvido sem um estudo aprimorado acerca da influência de seus diversos mecanismos na sua dinâmica global. Isso deixa margem para questionamentos, tais como “qual o impacto dos mecanismos específicos no desempenho global do sistema?”, ou ainda, “existem cenários em que os mecanismos distribuídos induzam um desempenho insatisfatório?”. O desempenho aqui pode ser medido em termos de tempo médio de download, vazão, quantidade de dados transmitida pelo servidor, disponibilidade do conteúdo, justiça entre usuários, etc.

Entender a dinâmica dos mecanismos do BT é fundamental para projetar um sistema que possa oferecer um melhor desempenho. Nesse sentido, diversos modelos na literatura foram propostos para representar sistemas P2P e caracterizar essas medidas em função de fatores externos, como taxa de chegada de peers, tempo de permanência no sistema após o download, etc. No entanto, poucos trabalhos estudam a influência dos mecanismos específicos ao protocolo em sua eficiência. Além disso, a maioria dos modelos nessa área assume que um grande número de usuários coexiste no sistema. Contudo, medições recentes em sistemas reais [Hossfeld et al. 2011] e outras realizadas por nós (pós-tese) mostram que a maior parte das redes reais é bem pequena, tendo menos de 5 peers. Apesar de ser o cenário mais comum, o estudo do desempenho do sistema nessas condições foi por muito tempo relegado. O trabalho aqui resumido é um passo no sentido de suprir essas deficiências, isto é, de entender como os mecanismos distribuídos do BT e a baixa popularidade de conteúdo oferecido podem influenciar a dinâmica global e desempenho do sistema.

2. Objetivos e Contribuições

O trabalho desenvolvido se insere no contexto do estudo de aplicativos par-a-par (P2P) para compartilhamento de arquivos, que devido a sua popularidade e bom desempenho se tornaram um recorrente tópico de pesquisa na última década. Modelar e entender a dinâmica desse tipo de sistema é, de fato, uma questão de fundamental importância devido a sua imensa aplicação prática.

O objetivo deste trabalho é investigar dois fenômenos que emergem em redes P2P do tipo BitTorrent, mas que ocorrem em cenários essencialmente diferentes. O primeiro fenômeno se trata da formação de clusters em redes heterogêneas, isto é, onde os usuários do sistema possuem largura de banda diferente. O segundo consiste na ocorrência de tempos de downloads heterogêneos em redes homogêneas, ou seja, onde os usuários possuem mesma largura de banda. Propomos e avaliamos modelos analíticos, realizamos simulações e experimentos em sistemas reais. Embora tenham impacto direto no desempenho do sistema, esses fenômenos eram até então pouco explorados na literatura. Motivações e detalhes específicos de cada uma das contribuições são descritos a seguir.

2.1. Formação de clusters em swarms BT heterogêneos

Recentemente, observou-se a formação de clusters entre os peers (usuários) de uma rede BitTorrent por largura de banda [Legout et al. 2007], ou seja, peers tendem a se conectar e trocar informação com peers que possuem largura de banda semelhante a sua própria. Este fenômeno está diretamente relacionado à escalabilidade e robustez do sistema. Em particular, um sistema P2P com alta clusterização é mais eficiente e oferece melhor serviço a seus usuários [Bharambe et al. 2006]. Entretanto, os mecanismos empregados pelo BT não fazem nenhuma preferência explícita pela troca de dados com peers com largura de banda similar. Esta propriedade simplesmente emerge da dinâmica dos seus mecanismos, como foi observado na literatura. Considerando a importância da clusterização no desempenho do sistema, realizamos um estudo sobre a dinâmica do BT para entender quais mecanismos são responsáveis por induzir este fenômeno.

Neste trabalho propomos um modelo simplificado que captura a essência da dinâmica do BitTorrent. Através de simulações, mostramos que o processo dinâmico, que se baseia apenas em regras locais, leva o sistema de um estado inicial aleatório a um estado com alta clusterização. Estudamos ainda o impacto dos parâmetros do sistema, representados no modelo, neste processo dinâmico. Entre outras observações, nossos resultados indicam que um mecanismo do BT, o *unchoke otimista*, é essencial para que o sistema atinja alta clusterização. A seguir apresentamos uma descrição do modelo em alto nível, acompanhado da intuição por trás da modelagem e alguns de seus resultados.

Modelo da dinâmica de conexões

Um peer que participa de um swarm BT conhece apenas uma fração dos usuários presentes na rede. Uma razão para isso é que o número de vizinhos de um peer é limitado (p.ex. 80 peers), embora o swarm possa ter mais de 1000 peers. Esse fato certamente limita a capacidade do swarm de formar clusters, pois um peer só pode se conectar e trocar dados com seus vizinhos. Por isso, modelamos o conhecimento sobre os vizinhos como um grafo regular (todos os vértices tem o mesmo grau) não-direcionado (se A está conectado a B, então B está conectado a A). Assumimos que este grafo é estático, o que é razoável num cenário onde não ocorrem muitas chegadas ou saídas de peer.

O upload de pedaços do conteúdo entre peers, entretanto, determina uma relação que muda ao longo do tempo. Esta relação não é bidirecional, sendo por nós representada através de um grafo direcionado. Este grafo está definido sobre o grafo de conhecimento, pois um peer não pode transmitir dados para outro que não conhece. Enquanto no BT a seleção de vizinhos para os quais um peer vai fazer upload é função de um algoritmo relativamente complexo, nosso modelo de simulação alterna entre o tit-for-tat e o unchoke otimista, capturando a essência do mecanismo originalmente empregado no sistema (para mais detalhes, ver Seção 3.2 da dissertação). Uma das vantagens dessa abordagem é permitir estudar a influência dos diferentes tipos de unchoke, apenas alterando a probabilidade com que cada um deles ocorre. A Figura 1, mostra a evolução do grau de clusterização da rede medido em termos do *Assortative Coefficient* (Eq. (3.1) da tese) ao longo do tempo, para diferentes probabilidades de se realizar um unchoke otimista (que correspondem às diferentes curvas). Claramente, observamos que o unchoke otimista tem um papel fundamental sobre a formação de clusters no BT. Em particular, este estudo revelou um compromisso entre a velocidade na formação de clusters e o grau

de clusterização alcançado depois de muito tempo. Repare que para $p = 1.0$ o assortative coefficient cresce de forma acentuada inicialmente, mas atinge valores menores, enquanto para $p = 0.01$ observamos o comportamento oposto.

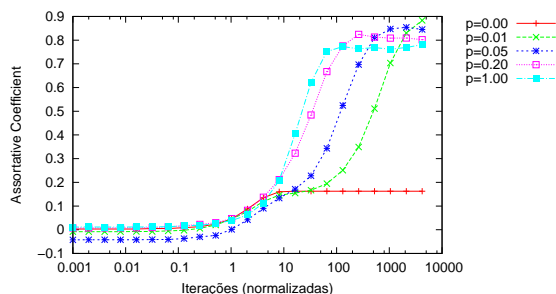


Figura 1. Evolução do grau de clusterização para diversos valores de p .

Além do unchoke otimista, estudamos a influência de outros parâmetros do sistema na clusterização por largura de banda, observando algumas invariâncias e outros compromissos (ver tese). Como complemento, propomos um modelo Markoviano simplificado de tempo contínuo passível de ser resolvido analiticamente. Este modelo fornece uma aproximação para o regime estacionário do sistema e mostrou capturar as características principais em estudo nesta seção.

2.2. Tempos de download heterogêneos em swarms BT homogêneos

A segunda contribuição deste trabalho trata da identificação e estudo de um fenômeno, até então inexplorado na literatura, que ocorre predominantemente em redes pequenas (que são a maioria). Mais precisamente, observa-se que os usuários do sistema podem obter taxas de download diferentes mesmo quando a população é homogênea com relação à banda. Isto tem diversas implicações, como alta variabilidade dos tempos de download, injustiça em relação à ordem de chegada dos peers e sincronização dos pedaços do conteúdo detidos pelos peers. Embora este comportamento seja mais prevalente em swarms pouco populares, este é o cenário mais comum hoje em dia e impacta negativamente o desempenho do sistema.

Para estudar a heterogeneidade das taxas de download em populações do BT onde todos os usuários possuem a mesma largura de banda, desenvolvemos um modelo de simulação detalhado que captura fielmente os diversos mecanismos do BT. Além disso, realizamos experimentos reais no PlanetLab. Finalmente, propomos um modelo analítico simplificado capaz de explicar esse fenômeno e suas consequências. Esse modelo prevê taxas de download heterogêneas para peers homogêneos, como função do conteúdo de cada peer. A seguir, apresentamos uma descrição resumida do modelo analítico proposto.

Modelo de fluido da dinâmica de troca de dados

Considere um swarm pouco popular, onde um seed está sempre presente. Suponha que os leechers têm largura de banda idênticas, que eles chegam de forma sequencial e saem do sistema assim que concluem o download. Neste cenário, observamos a ocorrência de saídas em rajada (vários peers deixam o sistema em um intervalo curto de tempo), o que implica que os peers mais velhos (isto é, que chegam antes dos outros no swarm) obtêm menores taxas de download durante algum tempo. Através de uma análise cuidadosa do progresso individual do download dos peers, concluímos que esse fato ocorre porque os

pedaços do arquivo que os peers mais novos têm e que são de interesse dos mais velhos rapidamente se esgotam.

Ao invés de modelar a troca de pedaços de forma discreta, utilizamos um modelo onde os peers trocam “fluido” a uma dada taxa. Essas taxas são determinadas considerando as relações de interesse entre os peers, isto é, peer i está interessado em j se i possui um pedaço que j não possui. Note que essa relação é assimétrica e muda ao longo do tempo. Uma observação chave para determinar tais relações é que um peer mais velho i sempre possui pedaços interessantes a um peer mais novo j . Por outro lado, mesmo que o peer mais novo j não possua pedaços interessantes a i em um determinado instante, ele irá receber pedaços do seed e de outros peers possivelmente mais velhos que i , que serão interessantes a i com alta probabilidade. Após calcularmos a taxa máxima de dados que pode ser transmitida entre cada par de peers baseado nas suas relações de interesse, estimamos a taxa efetiva de transmissão considerando as limitações de banda dos peers (ver Seção 4.5 da dissertação para mais detalhes).

Validamos o modelo analítico anterior comparando suas previsões de taxas de download com resultados do nosso modelo de simulação detalhado. Considere, por exemplo, um swarm homogêneo tal que o uplink dos seeds e leechers é 64 kbps, onde ocorrem 5 chegadas. A Figura 2a mostra a evolução do número de pedaços obtidos por cada leecher, que correspondem às diferentes curvas. O primeiro leecher chega no instante $t = 0$ e os quatro outros leechers, nos instantes $t = 30, 40, 50, 60$, todos com zero blocos. Escolhemos pontos sobre as curvas em instantes de tempo onde ocorrem eventos que podem alterar as taxas de download dos peers. Mais precisamente, marcamos com identificadores numéricos (de 1 a 25) os pontos em que ocorrem chegadas ou em que dois peers passam a ter o mesmo número de pedaços. A Figura 2b mostra as taxas de download obtidas por simulação e pelo modelo proposto para os pontos numerados na Figura 2a. Observamos que o modelo é bastante acurado, obtendo erro relativo menor que 1% para todos os pontos analisados, exceto para o 5º (7%) e o 24º (3%).

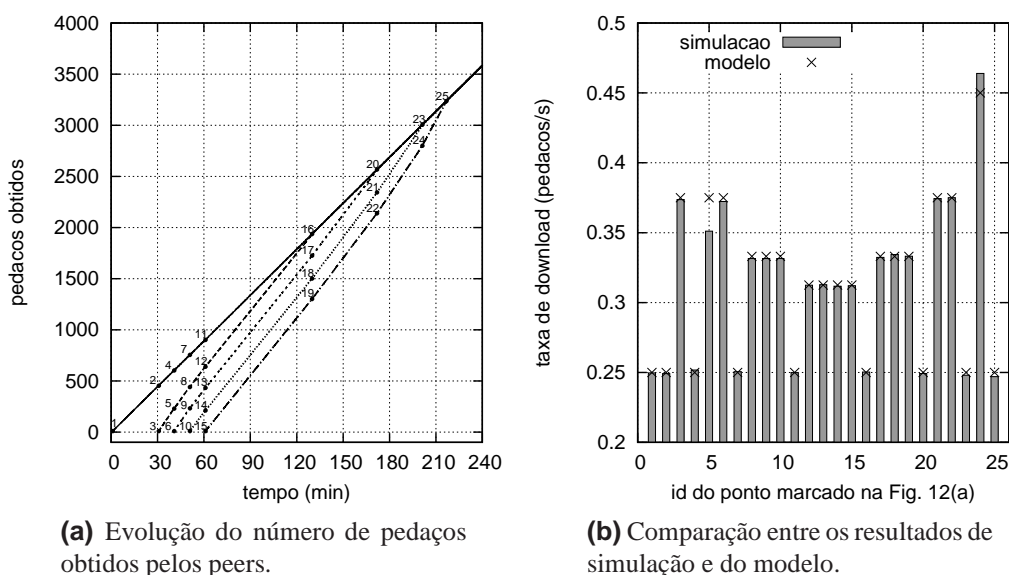


Figura 2. Validação do modelo.

Aplicamos esse modelo na previsão do número médio de saídas em rajada em um

intervalo de tempo. Essa quantidade fornece uma medida indireta da injustiça inerente ao sistema, pois quando peers chegam ao swarm em instantes de tempo diferentes e saem (quase) juntos, podem ocorrer tempos de download significativamente diferentes (duas vezes maiores, por exemplo). Observamos que a fração média de leechers que saem em rajada decresce com a popularidade do conteúdo e com a capacidade do seed. Portanto, esta medida poderia ser usada para ajustar a capacidade do servidor de forma a atender certas restrições de justiça (mais detalhes na dissertação).

3. Publicações

A primeira parte do trabalho, que trata da clusterização no BT onde a largura de banda é heterogênea foi inicialmente publicada como artigo completo em [i]. Em seguida, foi publicado um resumo estendido do trabalho em [ii], quando tive a oportunidade de conversar com pesquisadores que estavam desenvolvendo na época modelos de desempenho para o BT em redes heterogêneas.

Os primeiros resultados empíricos evidenciando a ocorrência de tempos de download heterogêneos em swarms homogêneos, referentes à segunda parte do trabalho, foram publicados como resumo estendido em [iii]. O modelo matemático capaz de explicar esse fenômeno e suas consequências, juntamente com resultados obtidos através de experimentos no PlanetLab e de um modelo de simulação detalhado do BT foram publicados em um artigo completo em uma influente revista internacional na área de redes [iv]. Vale ressaltar que diversos modelos desta dissertação foram desenvolvidos no ambiente de modelagem Tangram-II, fazendo uso de algumas de suas novas funcionalidades que foram desenvolvidas e apresentadas em [v].

Listagem de publicações

- [i] Murai, F. and Figueiredo, D., “Formação de clusters em redes P2P por similaridade entre os nós”, SBC/SBRC 2009.
- [ii] Murai, F. and Figueiredo, D., “Assortative Mixing in BitTorrent-like Networks”, IEEE/INFOCOM 2009 Student Workshop.
- [iii] Murai, F., Rocha, A., Figueiredo, D. and de Souza e Silva, E., “Can identical BitTorrent peers experience different download times?”, IFIP/Performance 2010 Poster Session.
- [iv] Murai, F., Rocha, A., Figueiredo, D. and de Souza e Silva, E., “Heterogeneous download times in a homogeneous BitTorrent swarm”, Elsevier/Computer Networks 2012.
- [v] Rocha, A., Jaime, G., Murai, F., Figueiredo, D. and de Souza e Silva, E., “Novas evoluções integradas à ferramenta Tangram-II v3.1”, SBC/SBRC 2009 (Salão de Ferramentas).

Referências

- arstechnica (2009). Cnn p2p video streaming. arstechnica.com/web/news/2009/02/cnn-p2p-video-streaming-tech-raises-questions.
- Bharambe, A., Herley, C., and Padmanabhan, V. (2006). Analyzing and improving a bittorrent network’s performance mechanisms. In *IEEE INFOCOM 2006*.
- Hossfeld, T., Lehrieder, F., Hock, D., Oechsner, S., and Despotovic, Z. (2011). Characterization of bittorrent swarms and their distribution in the internet. *Computer Networks*.
- Legout, A., Liogkas, N., Kohler, E., and Zhang, L. (2007). Clustering and sharing incentives in bittorrent systems. In *ACM SIGMETRICS 2007*.