

Abordagens de Particionamento de Dados para Redes Sociais

Bruno S. Velasco¹, Carmem S. Hara¹

¹Departamento de Informática – Universidade Federal do Paraná (UFPR)
Caixa Postal 19.081 – 81.531-980 – Curitiba – PR – Brasil

Abstract. *Online Social Networks (OSN) have never been so popular as it is nowadays. Due to its fast growth, designing a suitable infrastructure demands time and resources. Thus, a common approach adopted by several OSN is to set them on the Cloud. Although data gets spread throughout servers, a good balancing of data is barely achieved. To overcome this, several methods have been developed towards a good partitioning. This paper presents the main techniques for OSNs partitioning: general approaches, such as Full Replications and Distributed Hashing, and more specific methods for OSN such as Partitioning and Replication and Time-dependent Partitioning. Also, a comparative study is presented highlighting which approaches are best for each possible context.*

Resumo. *Redes Sociais Online (RSO) nunca foram tão populares quanto são hoje. Devido ao seu rápido crescimento, especificar uma boa infraestrutura demanda tempo e recursos, por esse motivo é comum hospedá-las na nuvem. Mesmo com os dados distribuídos, um bom balanceamento dificilmente é contemplado. Assim, diversas técnicas tem sido desenvolvidas para alcançar um bom particionamento. Este artigo apresenta algumas técnicas para o particionamento de RSO. Elas são classificadas como de uso geral, como Replicação total e Hash distribuído, e métodos específicos para RSO, como o SPAR e Particionamento dependente de tempo. Ao final, um estudo comparativo é apresentando destacando quais abordagens são mais apropriadas para cada cenário.*

1. Introdução

Uma Rede Social Online (RSO) é um serviço que permite a interação entre usuários através de meios bem conhecidos, como troca de texto ou mídia. Usuários desta rede podem estabelecer ligações entre eles e a quantidade deles pode variar de dezenas até milhares. Cada ligação feita pode ser entendida como uma “amizade” que ocorre entre os participantes. Atualmente, há mais de 200 RSO conhecidas cuja maioria tem mais de 1 milhão de usuários [Wikipedia 2011]. O Facebook lidera o topo desta lista com cerca de 750 milhões de usuários registrados [Facebook 2011]. Esta é uma característica marcante das RSO: os usuários se registram de forma acelerada. O Twitter, por exemplo, cresceu cerca de 1382% entre Fevereiro e Março de 2009 [Cnet 2011, Carrasco et al. 2011]. Esse rápido crescimento compromete o desempenho do sistema caso ele não tenha sido desenvolvido para dar suporte a tal carga de trabalho. Uma solução comum é atualizar o hardware existente. Essa abordagem é conhecida como *escalabilidade vertical*, porque a capacidade é incrementada sobre os mesmos recursos. Esta decisão não só é financeiramente cara como ineficiente, a curto prazo, já que o hardware tende tornar-se obsoleto em pouco tempo. Além do mais, não é uma solução escalável [Pujol et al. 2010, Pujol et al. 2009].

Esse rápido crescimento resulta em um dilema entre os desenvolvedores de RSO: “É preferível alocar recursos humanos no desenvolvimento de novas funcionalidades ou

focar na construção de um sistema altamente escalável, capaz de dar suporte a um grande volume de dados?” [Pujol et al. 2010, Pujol et al. 2009]. A primeira escolha é a mais comum e claramente a mais fácil. Por outro lado, tal escolha pode resultar em uma situação conhecida como “morte-pelo-sucesso”, no qual usuários são atraídos pelas funcionalidades do sistema e este, por sua vez, não é capaz de comportar tal demanda. Uma situação semelhante ocorreu com a RSO *Friendstar* [Scalability 2011].

No entanto, a segunda opção, que consiste em criar uma base sólida para acomodar uma carga de trabalho que pode nunca vir a acontecer, dispense um grande investimento em termos de recursos humanos e financeiros. Tais esforços podem não ser recompensados justamente por não se saber se esta RSO terá o retorno esperado. Porém, expansões futuras tornam-se mais simples e baratas pois novos servidores de baixo desempenho podem ser utilizados. Esta abordagem é conhecida como *escalabilidade horizontal* [Pujol et al. 2010, Pujol et al. 2009].

Tornar uma RSO escalável não é uma tarefa trivial. Consequentemente, o seu particionamento torna-se complexo [Carrasco et al. 2011] [Viswanath et al. 2009] [Curino et al. 2010] [Pujol et al. 2010]. Como ela é uma comunidade, uma possível representação computacional de uma RSO é um grafo cujos nodos representam os usuários, e as arestas, as ligações (amizades) ocorrendo entre os participantes. Esse grafo é conhecido como *grafo de amizades*. Entretanto, tal grafo contém nenhuma informações sobre a coesão das ligações. Por isso, para um determinado intervalo de tempo, a *rede de atividade* [Viswanath et al. 2009] pode ser considerada, que consiste no subgrafo originado da interação entre os pares com peso nas arestas, que representa o volume de interações entre o par.

Apesar dos desafios que uma solução escalável apresenta, ela é necessária para comportar um alto tráfego de dados. Um destes desafios é o particionamento de dados. O problema em questão é o quão eficiente e “bom” um particionamento pode ser, ou seja, fazer com que “dados afins” (usuários) sejam armazenados nos mesmos servidores, evitando sobrecarga da rede. A definição de “dados afins” não é precisa e pode variar de meramente ser uma ligação de amizade, para a frequência em que um par de usuários interage em um dado período de tempo. O objetivo do particionamento é explorar a localidade espacial entre os dados de um mesmo nodo de rede, evitando a comunicação desnecessária entre os servidores.

Um bom particionamento de uma RSO depende diretamente de como os desenvolvedores definem “dados afins”, ou seja, sob quais critérios os dados serão particionados. Atualmente, existem as seguintes técnicas de distribuição de dados para RSO:

1. **Abordagens gerais:** são técnicas agnósticas do conceito de dados afins. Dentre elas podem ser citadas as Tabelas de Dispersão Distribuídas (*Distributed Hash Tables - DHT*) e Replicação Total.
2. **Particionamento baseado no grafo de amizade:** são técnicas que utilizam a estrutura do grafo de amizades para determinar o particionamento de usuários e replicação dos dados. Dentre elas, existe o SPAR [Pujol et al. 2010], que é uma plataforma de particionamento e replicação capaz de escalar, transparentemente, a estrutura de grafo para atingir localidade de dados, ao mesmo tempo que minimiza as replicações.
3. **Particionamento dependente de tempo:** são técnicas que também utilizam o

grafo de amizades, porém consideram em conjunto com o espaço temporal para definir o conceito de dados afins, que constitui a rede de atividade. Uma das técnicas que segue essa abordagem foi proposta em [Carrasco et al. 2011], que busca agrupar pares de usuários com maior taxa de interação nos mesmos servidores.

O resto desse trabalho está organizado como segue. A Seção 2 descreve o comportamento de uma RSO de acordo com Viswanath et al. A Seção 3 detalha cada uma das abordagens principais de particionamento. A Seção 4 contém um breve estudo comparativo entre as abordagens de particionamento e a Seção 5 traz o fechamento desse trabalho.

2. Características de uma RSO

Alguns estudos tem sido desenvolvidos para melhor compreender o comportamento e particularidades de uma RSO [Viswanath et al. 2009, Benevenuto et al. 2009]. Embora diversos estudos sejam baseados no grafo de amizades e na rede de atividades, poucos consideram sua modificação ao longo do tempo, ou seja, se as ligações entre os usuários tornam-se mais fortes ou fracas. Viswanath et al. propuseram analisar essas modificações sob dois aspectos: microscópico e macroscópico. O primeiro nível atem-se em estudar o comportamento para cada par de usuários, enquanto o nível macro avalia o quanto usuários individuais alteram o comportamento e propriedades da rede.

Os dados empregados para a avaliação foram baseados no Facebook. A forma de interação escolhida para o estudo foi postagens no mural de recados que cada usuário tem em seu perfil. Esse método de interação pode apresentar algumas limitações se comparado às outras maneiras de interatividade providas pelo Facebook como mensagens, fotos, vídeos, e conversa [Viswanath et al. 2009]. Porém, o mural foi escolhido por conter o *timestamp* que é utilizado para monitorar as ações dos usuários ao longo do tempo.

Uma característica de alto nível observada foi que a rede de atividade representa cerca de 12,2% do grafo de amizades. Isto confirma um estudo anterior que afirma que o grau de um nodo de uma rede de atividade (ou seja, o número de amigos com os quais usuários interage na média) é significamente menor que o grau de um nodo da rede de amigos [Viswanath et al. 2009].

2.1. Interação de um par de usuários

O primeiro comportamento interessante que foi observado remete à distribuição do número de postagens feitas no mural dos usuários. Esse comportamento é mostrado na Figura 1, que tem no eixo *y* a distribuição cumulativa e no eixo *x* o número de postagens para cada par. A mediana do número de postagens por par de usuários é 2, sendo que 81% dos pares não trocaram mais que 5 postagens [Viswanath et al. 2009].

Essa distribuição sugere dois níveis de usuários: com alta e baixa taxa de interação. Os autores consideraram um usuário de baixa interatividade se esse mandou não mais que 5 postagens no mural, caso contrário, tal usuário é considerado de alta interatividade. Para os de baixa interatividade, foi constatado que as postagens ocorreram, ou por causa da primeira mensagem, ou devido ao aplicativo do Facebook que lembra do aniversário de seus amigos [Viswanath et al. 2009]. Na verdade, 54% dos pares do grupo de baixa interatividade tiveram ao menos uma postagem relacionada à felicitações de aniversário.

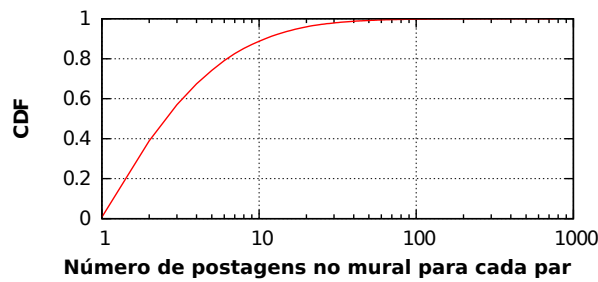


Figura 1. Distribuição do número de postagens no mural entre cada par [Viswanath et al. 2009].

O grupo de alta interatividade teve um comportamento diferente. A maioria destes usuários trocou a sua primeira mensagem tão logo a ligação foi efetuada. Além disso, a interação decaiu ao longo do tempo. A Figura 2 ilustra a taxa de interatividade no tempo, baseada nas postagens enviadas ao mural desde a criação da ligação. As postagens são agrupadas por mês.

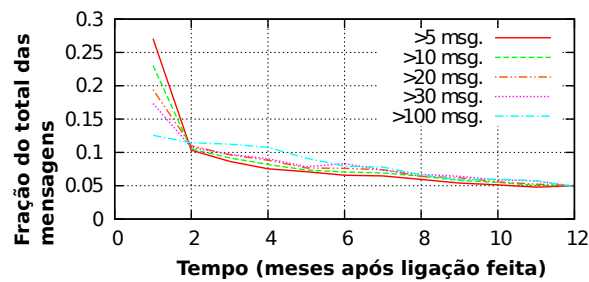


Figura 2. Fração das mensagens enviadas ao longo do tempo [Viswanath et al. 2009].

Após observar a redução constante, os autores voltaram suas atenções para quão longa uma ligação entre os pares permanece na rede de atividade. No total, apenas 23% dos pares do grupo de alta interatividade se mantiveram durante o intervalo de 12 meses. Viswanath et al. concluíram que para ambos os grupos, de alta e baixa interatividade, usuários tendem a interagir menos com o passar do tempo.

2.2. Evolução da rede ao longo do tempo

Após ter estudado o comportamento entre os pares de usuários, os autores pesquisaram o impacto na rede que esse comportamento ocasiona. Para tanto, 9 *snapshots* foram realizadas na rede, com 90 dias de intervalo cada. Eles estavam interessados em quantos dos nodos permaneciam na rede ao longo do tempo. Para isso, o conceito de *semelhança* foi empregado, que consiste em tomar 2 *snapshots* consecutivas e ver quantos dos nodos residem em ambas as amostras. A Figura 3 apresenta a análise para todas as amostras. É possível observar que as duas primeiras *snapshots* possuem semelhança de 60% aproximadamente, e essa taxa reduz levemente ao longo do tempo. Apesar desta mudança, cerca de 45% das ligações permaneceram ativas durante o período.

Acerca das propriedades do grafo resultante da rede de atividade, os autores utilizaram algumas métricas para estudo de grafos e, interessantemente, todos os testes

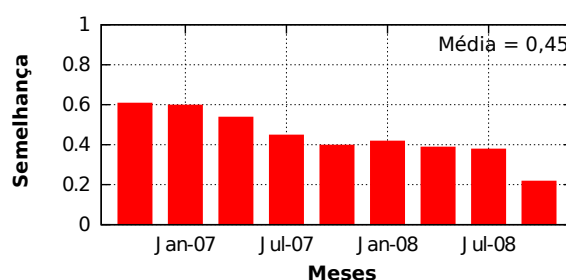


Figura 3. Semelhança da rede de atividade [Viswanath et al. 2009].

mantiveram estáveis para todas as amostras, evidenciando que grandes variações entre as ligações individuais não afetam a rede como um todo.

3. Abordagens para distribuição de dados de uma RSO

Nesta seção são apresentadas algumas técnicas propostas na literatura para o particionamento e replicação de dados de uma RSO. Dentre as abordagens gerais são descritas as técnicas de Tabela de Dispersão Distribuída (**DHT**) e **Replicação total**. Essas estratégias não consideram a afinidade entre os usuários para distribuir os dados como suas ligações ou interações, assim sendo, não garantem a localidade dos dados. Como representante da abordagem baseada no grafo de amizade é apresentado o **SPAR** [Pujol et al. 2010]. Além disso, é apresentada a proposta de [Carrasco et al. 2011], que segue a abordagem de particionamento dependente de tempo. Ambas consideram a afinidade dos usuários.

3.1. Abordagens gerais

As técnicas de DHT e Replicação total são métodos de abordagem geral também aplicados para RSO. A Replicação total, largamente utilizada por sua simplicidade [Pujol et al. 2010], consiste simplesmente em replicar todos os nodos (usuários) do grafo por todos os servidores. Essa operação requer grande capacidade de armazenamento e possui um alto custo computacional já que cada réplica precisa estar sincronizada. Sua vantagem, porém, é que a localidade espacial é total, uma vez que todo nodo tem sua réplica em cada servidor. A DHT, por sua vez, agrupa usuários mediante uma chave. Essa chave pode ser o nome de usuário ou um valor aleatório. Opcionalmente, replicação pode ser empregada para oferecer uma melhor localidade espacial. A DHT é amplamente utilizada para a distribuição de dados por ser um método escalável [Pujol et al. 2010] sem controle centralizado e tolerante a falhas.

No topo da Figura 4 é ilustrada uma RSO com 10 usuários e 15 ligações (amizade bidirecional). Imediatamente abaixo, são ilustradas as abordagens de Replicação total (a), DHT (b) e DHT com replicação (c), juntamente com as métricas para comparação, que são o *tráfego gerado na leitura* de um usuário, *tráfego gerado na escrita* de um usuário e o espaço em memória utilizado. Em (a), Replicação total atinge 100% de localidade espacial (*tráfego na leitura* = 0), porém requer alta capacidade de armazenamento (*memória* = 10) e sincronização constante (*tráfego na escrita* = 10). Em (b), a DHT é empregada para distribuir os usuários sem considerar a ligação entre os nodos. Apesar de exigir menos espaço de armazenamento e dispensar sincronização, não é oferecida localidade de dados. Assim, para amenizar o custo da localidade, em (c) é empregada replicação dos dados, que exige armazenamento e sincronização extras.

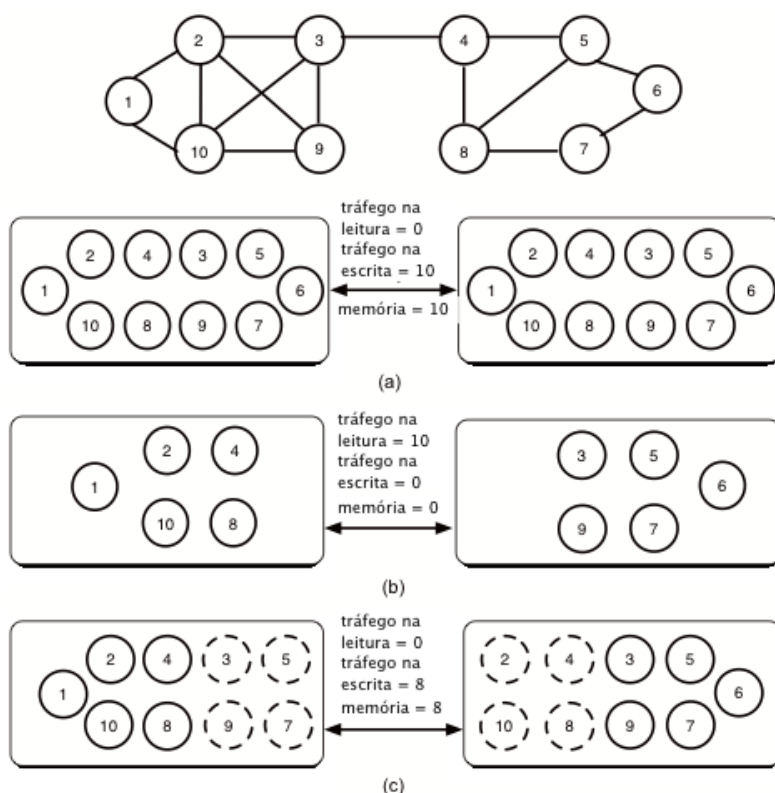


Figura 4. Parte de uma RSO sendo particionando em 2 servidores usando (a) Replicação total, (b) DHT, (c) DHT com replicação dos vizinhos [Pujol et al. 2010].

3.2. Particionamento e replicação (SPAR)

Essa abordagem foi introduzida por Pujol et al., sendo um dos primeiros métodos de particionamento para RSO a considerar a ligação entre os usuários, ao contrário das abordagens gerais. O SPAR atua como um *middleware*, permitindo que o usuário escolha entre um sistema de armazenamento relacional ou chave-valor. A distribuição dos dados é feita de forma automática, particionando e replicando os dados conforme necessário durante a execução, sem a necessidade de interrupção do serviço. Quando alguma alteração no grafo é realizada, seja pela adição ou remoção de nodos e ligações, ele é capaz de decidir quais ações tomar [Pujol et al. 2010].



Figura 5. Particionamento usando SPAR [Pujol et al. 2010].

O particionamento resultante do SPAR garante que todos os usuários da RSO e seus vizinhos diretos (conhecidos como *vizinho a um-salto*) estão armazenados no mesmo servidor, atingindo melhor localidade espacial para os dados mais importantes (amigos, seguidores, etc.) [Pujol et al. 2010]. A Figura 5 mostra como esse método particionaria a RSO referente à Figura 4; note que somente os vizinhos a um-salto relevantes são

replicados em ambos os servidores (*tráfego na leitura = 0*). Ao final, os custos de armazenamento, leitura e escrita na RSO, são inferiores do que as abordagens de propósito geral.

Para avaliar o desempenho do sistema, os autores realizaram experimentos focados na sobrecarga envolvida para a replicação, considerando que o gargalo está diretamente relacionado com a replicação de usuários. A fonte de dados utilizada envolveu dados reais do Orkut, Twitter e Facebook, coletados em diferentes períodos.

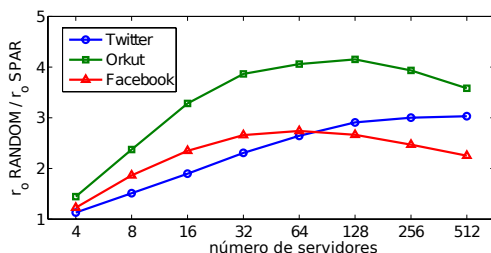


Figura 6. SPAR versus DHT [Pujol et al. 2010].

A Figura 6 ilustra um dos experimentos realizados, a fim de verificar o quão inferior é a sobrecarga para o particionamento do SPAR se comparado com a DHT. A notação (r_o) expressa o número total de réplicas necessárias para manter a localidade semântica. No eixo y do gráfico está plotada a relação da sobrecarga referente ao DHT com do SPAR, enquanto no eixo x é variado número de servidores. Por exemplo, para 128 servidores, o SPAR utilizou 4 vezes menos réplicas que a DHT para distribuir dados provenientes do Orkut.

À medida que novas ligações (arestas) são adicionadas, algumas ações ocorrem para manter o balanceamento do grafo a fim de manter a localidade. A Figura 7 mostra essas ações, evidenciando que na maior parte do tempo o SPAR permanece em espera, apesar de algumas ações ocorrerem.

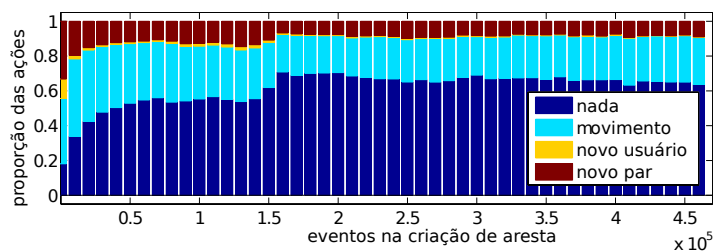


Figura 7. Percentual de cada ação promovida pelo SPAR ao longo da criação de ligações [Pujol et al. 2010].

Pujol et al. concluem seus resultados demonstrando que o SPAR é capaz de alcançar 100% de localidade com uma sobrecarga de operação mínima, apresentando um desempenho muito superior à DHT.

3.3. Particionamento dependente de tempo

O método proposto por [Carrasco et al. 2011] visa particionar uma RSO baseada na interação dos usuários participantes. O conceito de usuários afins envolve a dimensão tempo, ou seja, a frequência com que eles interagem.

Baseado no trabalho de Viswanath et al., apresentado na Seção 2, a rede de atividade é construída sobre o grafo de amizades correspondente, que ao contrário do SPAR, considera a dimensão tempo. Uma vantagem desta abordagem é que a rede de atividade é consideravelmente menor que o grafo de amizades em si, (12,2% como observado na Seção 2).

A Figura 8 mostra uma pequena RSO com 7 nodos. As mensagens trocadas entre os usuários são representadas por um número entre colchetes sobre a ligação, como pode ser visto na Figura 8(a). Por exemplo, no período 1, os usuários 1 e 2 trocaram 100 mensagens enquanto os usuários 2 e 3 não interagiram; no período 2, usuários 1 e 2 não interagiram, ao contrário dos usuários 2 e 3 que trocaram 100 mensagens. Quando a replicação é utilizada para garantir a localidade de vizinhança a um-salto, tal como SPAR, todas as 200 mensagens são replicadas, como na Figura 8(b). Porém, quando a rede de atividade é particionada em períodos diferentes, por existir menos mensagens entre usuários em um intervalo de tempo, nenhuma replicação é executada, como na Figura 8(c).

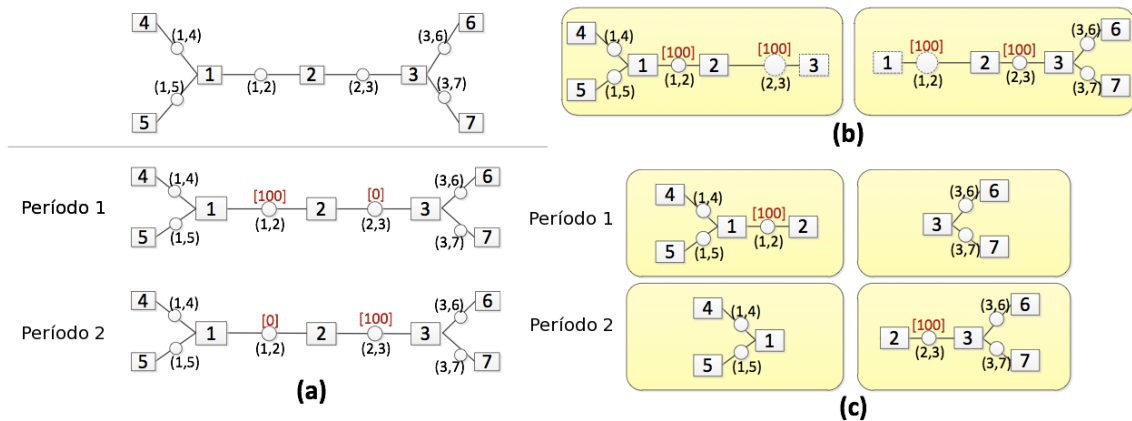


Figura 8. Exemplo de uma pequena RSO sob alguns cenários de particionamento [Carrasco et al. 2011].

A avaliação desse modelo buscou evidenciar quanto da localidade espacial foi corretamente contemplada. Para tal, o mesmo conjunto de dados do Facebook da Seção 2 foi utilizado. A consulta consiste nas 6 mensagens mais recentes dos amigos dos amigos (conhecido como *vizinhança à dois-saltos*). Ao todo, 3 métodos de particionamento foram comparados: o próprio algoritmo, DHT *hash_p1* (todas mensagens geradas pelo mesmo usuário são agrupadas em uma partição) e outra modalidade de DHT *hash_p1p2*, com a diferença que todas as mensagens trocadas entre um par de usuários são agrupadas na mesma partição [Carrasco et al. 2011].

A Figura 9 evidencia que para todos os números de partições, mais de 80% das consultas acessaram no máximo 3 partições com o algoritmo proposto. As autoras concluem que esse método é consideravelmente melhor para a localidade dos dados se comparado com os outros dois algoritmos.

4. Discussão

A importância de um bom particionamento de dados pode ser traduzida através da localidade espacial dos dados, que consiste em agrupar dados afins nas mesmas partições,

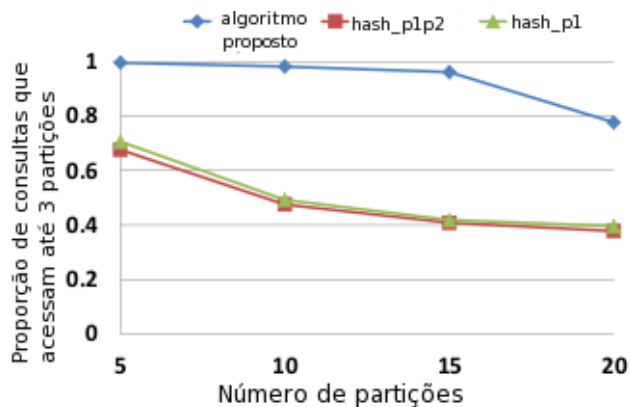


Figura 9. Proporção de consultas que acessam no máximo 3 partições [Carrasco et al. 2011].

evitando tráfego de rede. Quanto mais usuários “afins” pertencem a um mesmo nodo de rede, melhor foi o método de particionamento.

Existem dois tipos de escolha para particionar uma RSO: abordagens gerais e específicas. A primeira opção traz a vantagem da simplicidade, atingindo bons resultados se comparados às abordagens de escalabilidade vertical. A Replicação total oferece uma melhor localidade dos dados, apesar desse requerer mais espaço de armazenamento para comportar as réplicas. A DHT, por outro lado, exige menos espaço de armazenamento porém oferece menor localidade dos dados, mesmo com réplicas.

A segunda opção oferece algoritmos mais robustos, uma vez que eles consideram a propriedade de semelhança e coesão dos usuários. O SPAR atinge altos níveis de localidade, chegando a 100% quando apenas vizinhança a um-salto é considerada. Essa abordagem mostrou ter baixa latência e maior vazão quando comparado com Cassandra [Pujol et al. 2010], o sistema de armazenamento chave-valor utilizado pelo Facebook que faz distribuição de dados por DHT. Ele também é capaz de servir 3 vezes mais requisições se comparado com algoritmos comerciais baseados em DHT. Sua desvantagem, porém, é que exige muitas réplicas: para 512 partições, no Facebook, uma média de 7 réplicas são usadas enquanto que para alguns usuários suas réplicas existem em todos os servidores. O algoritmo baseado no tempo, por outro lado, possui a vantagem de trabalhar com a rede de atividade, cujo grafo é menor e mais representativo. Apesar de sua localidade ser ligeiramente menor que a solução SPAR, requer muito menos réplicas.

A escolha é necessária entre um método com uma perfeita predição sobre a localidade de dados ao preço de grande espaço de armazenamento, e o uso de um algoritmo com considerável localidade ao preço de armazenamento bastante inferior.

5. Conclusão

O presente trabalho apresentou alguns métodos para particionamento de Redes Sociais Online, incluindo abordagens gerais como **Replicação total** e **DHT**, e métodos específicos tais como **Particionamento e replicação (SPAR)** e **Particionamento dependente de tempo**. Foi observado que ambos os métodos específicos tem um melhor desempenho se comparados aos métodos tradicionais e comerciais utilizados para esse propósito. Apesar do SPAR alcançar 100% de localidade de dados para operações de

vizinhança a um-salto, ele requer muito mais espaço de armazenamento que o dependente de tempo, que se baseia na frequência da interação entre os usuários. A análise desta frequência evidencia uma rede conhecida como rede de atividade, que contém propriedades interessantes que são utilizadas por esse último método, sendo capaz de atingir altos níveis de localidade ao preço de poucas réplicas.

Após destacar as vantagens e desvantagens de cada abordagem, uma breve comparação foi realizada concluindo que existe uma escolha a se fazer entre capacidade de armazenamento e precisão na localidade dos dados.

Referências

- Benevenuto, F., Rodrigues, T., Cha, M., and Almeida, V. (2009). Characterizing user behavior in online social networks. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 49–62. ACM.
- Carrasco, B., Lu, Y., and da Trindade, J. (2011). Partitioning social networks for time-dependent queries. In *Proceedings of the 4th Workshop on Social Network Systems (SNS'11)*. New York, NY, USA: ACM, pages 503–509.
- Cnet (2011). Nielsen: Twitter's growing really, really, really, really fast. http://news.cnet.com/8301-13577_3-10200161-36.html.
- Curino, C., Jones, E., and Madden, S. (2010). Schism : a workload-driven approach to database replication and partitioning. *Proceedings of the VLDB Endowment*, 3(1-2):48–57.
- Facebook (2011). Facebook statistics. <http://www.facebook.com/press/info.php?statistics>.
- Pujol, J., Erramilli, V., Siganos, G., Yang, X., Laoutaris, N., Chhabra, P., and Rodriguez, P. (2010). The little engine (s) that could: Scaling online social networks. In *ACM SIGCOMM Computer Communication Review*, volume 40, pages 375–386. ACM.
- Pujol, J., Siganos, G., Erramilli, V., and Rodriguez, P. (2009). Scaling online social networks without pains. In *Proc of NETDB*.
- Scalability, H. (2011). Friendster lost lead because of a failure to scale. <http://highscalability.com/blog/2007/11/13/friendster-lost-lead-because-of-a-failure-to-scale.html>.
- Viswanath, B., Mislove, A., Cha, M., and Gummadi, K. P. (2009). On the evolution of user interaction in Facebook. *Proceedings of the 2nd ACM workshop on Online social networks - WOSN '09*, page 37.
- Wikipedia (2011). List of social networking websites. http://en.wikipedia.org/wiki/List_of_social_networking_websites.