

Análise comparativa da seleção de pares em uma P2P SON

Guilherme da Costa de Andrade¹, Omir Correa Alves Junior¹, Adriano Fiorese¹

¹Departamento de Ciência da Computação

Universidade do Estado de Santa Catarina (UDESC) – Joinville – SC – Brazil

guilhermel11@hotmail.com, omalves@gmail.com, fiorese@joinville.udesc.br

Abstract. *This paper consists in a comparative of the distribution of selected best and worst peers in a P2P SON. This distribution is the result of the selection of peers in a peer-to-peer service overlay network based in a performance indicator. To validate this distribution the selection mechanisms leading to the aforementioned distribution, several simulations were realized and each one selects the best and the worst peer of the network. With the results it is expected to answer questions related with the peer's distribution such as the location of the peers.*

Resumo. *Este artigo tem como objetivo fazer um comparativo da distribuição dos melhores e dos piores pares em uma P2P SON. Esta distribuição é resultado da seleção dos pares na rede de sobreposição par-a-par de serviços baseada em um indicador de desempenho. Para validação desta distribuição foram realizadas diversas simulações onde cada uma seleciona o melhor e o pior par presente na rede. Com os resultados obtidos deseja-se responder questões como, por exemplo, a localização destes pares.*

1. Introdução

Uma rede de sobreposição (*Overlay Network*) define com a sua própria estrutura uma nova topologia sobre a rede de computadores já existente. Ela tem as suas próprias regras e técnicas de roteamento usando as conexões na rede de computadores já existente [Esen 2008]. Recentemente, o uso destas redes tem crescido. Um dos motivos para o seu crescimento é o uso dessas redes para o fornecimento de serviços como, por exemplo, as aplicações de compartilhamento de arquivos onde os participantes trocam arquivos (dados) entre si. Ao serem utilizadas para o fornecimento de serviços, as redes de sobreposição ganharam uma nova denominação, chamadas de redes de sobreposição de serviços ou *Service Overlay Networks* (SON).

Segundo Lee *et al.* (2011) as redes de sobreposição de serviços (SON) foram criadas com o objetivo de serem empregadas como camadas intermediárias para apoiarem a criação e desenvolvimento de serviços com valores agregados, como por exemplo, transcodificação de mídias e encriptação de dados, da *Internet* sobre redes heterogêneas. Contudo, para fornecer serviços para os usuários finais, é necessário garantir que este seja entregue com qualidade.

O termo qualidade de serviço é amplo, devido às diversas necessidades específicas de cada um, sendo que essas necessidades nem sempre são as mesmas para todos os serviços. Por isso, os participantes dessas redes tendem a utilizar os serviços que melhor atendam as suas necessidades, selecionando o melhor participante da rede para se comunicar e trocar informações.

Uma das formas de construir uma SON é utilizar a tecnologia Par-a-par (P2P). Sistemas *Peer-to-peer* (P2P) ou Par-a-par são sistemas e aplicações distribuídas onde seus participantes (pares) comunicam-se entre si para fornecer certo serviço a um par requisitante.

Este artigo apresenta uma abordagem relativa à seleção dos melhores pares em uma rede P2P SON (*Peer-to-peer Service Overlay Network*). A partir desta abordagem, o objetivo deste artigo é selecionar os melhores pares dentro da P2P SON que contenham o serviço desejado, baseando-se em um indicador de desempenho da rede, o qual é composto pela razão entre duas métricas, largura de banda e distância euclidiana. Esta última métrica é obtida pela modelagem, em um plano cartesiano, do atraso de transmissão de mensagens na Internet com dados reais, de acordo com [Kaune, S. et al. 2009].

Estas duas métricas foram escolhidas para compor este indicador devido ao interesse em comprovar se os pares que possuem a melhor relação entre a largura de banda e a distância euclidiana são aqueles que transmitem os dados em menor tempo e com menor quantidade de retransmissões.

Além disto, com o uso deste indicador, pode-se indicar aos clientes uma forma de selecionar o melhor e o pior fornecedor, verificar e comparar a distribuição dos melhores pares e dos piores pares. Este trabalho também propõe respostas para questões como, por exemplo, se o melhor par está sempre localizado no mesmo domínio geográfico do par requisitante, ou se há variação na distribuição dos melhores e dos piores pares presentes na SON.

Assim, para resolver essas questões, diversas simulações foram realizadas utilizando o indicador proposto, o qual será detalhado durante o artigo. Assim, a partir dessas simulações, uma análise dos dados resultantes é apresentada de forma a inferir respostas sobre as questões apresentadas.

Por fim, este artigo está organizado da seguinte maneira. Seções 2 e 3 discutem trabalho relacionado além de contextualizar este trabalho. A Seção 4 descreve o serviço desenvolvido para selecionar os pares dentro da rede P2P SON. Na sequência, a Seção 5 apresenta a abordagem proposta. Na sequência serão apresentados os resultados na Seção 6 e por fim, a Seção 7 apresenta as conclusões obtidas com este trabalho.

2. Trabalho Relacionado

2.1. OMAN

Descrita por [Fiorese et al. 2010], OMAN é uma arquitetura de gerenciamento de serviços para Redes de Sobreposição Par-a-Par de Serviços que manipula aspectos, desde a composição de uma SON até a interação entre a SON e os serviços. Esta arquitetura está diretamente relacionada a este trabalho, pois entre seus módulos, dois deles, Serviço de Agregação (AgS) e Serviço de Seleção de Melhores Pares/Nós (BPSS), estão presentes em sua implementação, os quais serão discutidos nas próximas seções.

Esta arquitetura é dividida em três níveis, como pode ser visto na Figura 1. O primeiro nível é a base de uma P2P SON. Nele a comunicação P2P irá sustentar toda a arquitetura. Este nível deve ser comum a todos os nós participantes da P2P SON e é

responsável por gerenciar todos os aspectos relacionados ao gerenciamento da rede de sobreposição P2P, como por exemplo, a entrada e a saída de nós. O segundo nível da arquitetura é o Serviço de Agregação (AgS). Ele é uma camada P2P não estruturada, deste modo, não existe forte acoplamento entre a topologia da rede de sobreposição e a localização ou posicionamento físico das informações. No terceiro nível, o *Overlay Monitoring* ou Monitoramento da (rede) de Sobreposição (OM) coleta informações sobre o funcionamento da rede como, por exemplo, os seus recursos e condições de execução. Já o Serviço de Seleção de Melhores Pares (BPSS) é um serviço que informa aos seus solicitantes, um conjunto de melhores pares relacionados a um determinado serviço e de acordo com um indicador particular. O OM e o BPSS auxiliam o Gerenciador de Configuração (CM), o qual possui três sub-módulos chamados: QoE, Autonomia e Resiliência [Fiorese et al. 2010].

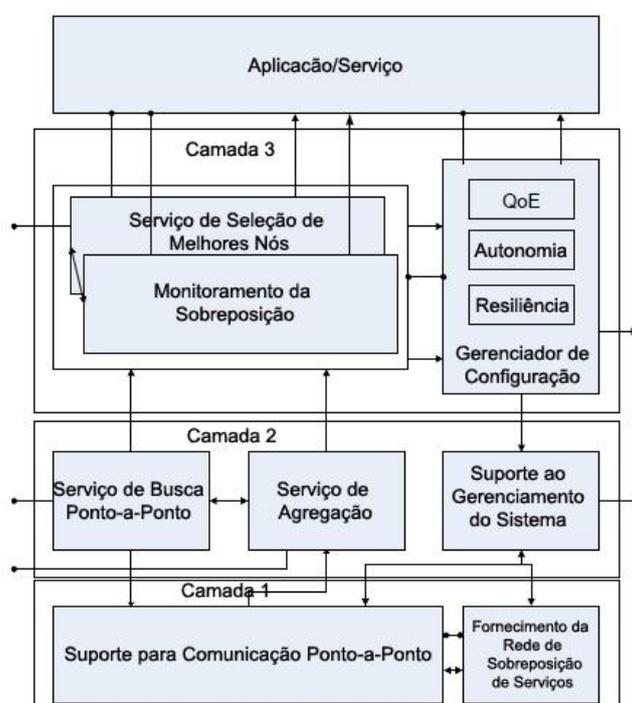


Figura 1. Arquitetura OMAN [Fiorese et al. 2010].

3. Serviço de Agregação

O Serviço de Agregação (AgS) é o módulo responsável por otimizar a busca de serviços. Ele compreende a agregação das informações dos fornecedores dos serviços, e assim pode melhorar a busca [Fiorese et al. 2010]. Os pares dentro do serviço de agregação são chamados de pares AgS e são escolhidos entre os pares pertencentes à rede P2P SON, tornando-os pares especializados. Os pares SON não especializados são aqueles pertencentes aos provedores de serviços que efetivamente executam os serviços e formam a P2P SON. Ao contrário destes, os pares AgS não executam os serviços. Estes apenas agregam as ofertas dos serviços feitas pelos pares SON.

A ideia dos pares de agregação é agregar as ofertas de serviços para facilitar a busca. Vale lembrar que os pares de agregação podem estar localizados no mesmo domínio geográfico dos pares SON, como também podem estar localizados em domínios geográficos diferentes.

4. Serviço de Seleção de Melhores Pares/Nós

Baseando-se na arquitetura OMAN desenvolvida por [Fiorese et al. 2010], este trabalho terá como foco principal o módulo de Seleção de Melhores Pares (BPSS) localizado no terceiro nível da arquitetura. Ele é construído utilizando-se das funcionalidades de busca fornecidas pelo Serviço de Agregação (AgS) para selecionar o melhor par desejado por outro par SON.

O objetivo do BPSS é fornecer pares SON identificados como melhores pares para um determinado serviço em relação a um indicador específico. Desse modo, a sua implementação torna-se importante em uma P2P SON devido à possibilidade dos seus membros ao buscarem os serviços desejados comunicarem-se com aqueles que melhor satisfizerem um indicador específico. Como exemplo, podemos citar um usuário buscando um serviço de envio de vídeos, e este pode ser escolhido entre vários fornecedores.

4.1. Funcionamento

A Figura 2 apresenta a arquitetura do Serviço de Seleção de Melhores Pares (BPSS). Com ela, é possível entender o funcionamento do módulo de BPSS.

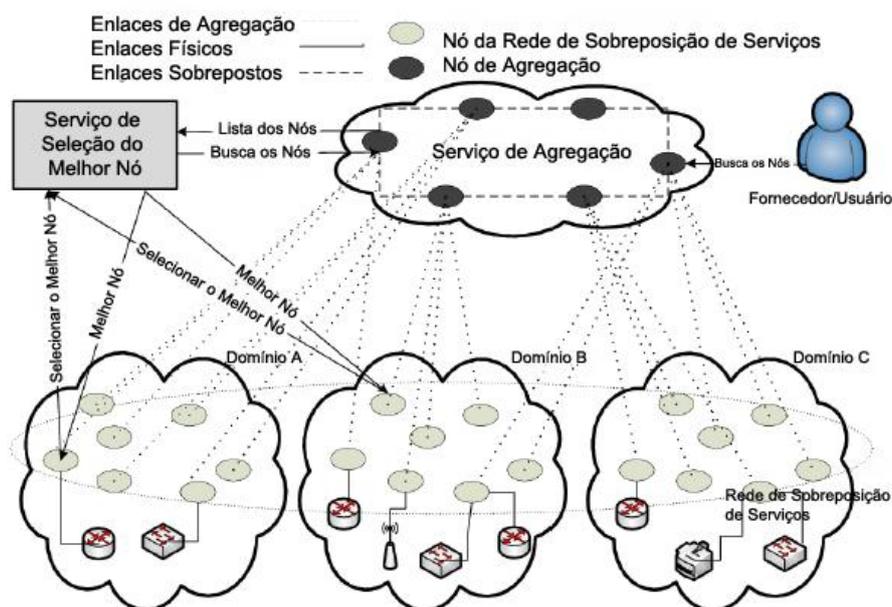


Figura 2. Arquitetura do BPSS [Fiorese et al. 2011].

Partindo-se do pressuposto de que a rede de sobreposição já está criada e o serviço de agregação disponível, inicialmente um par SON solicita ao BPSS informações sobre o melhor par que possui um determinado serviço utilizando a mensagem (Selecionar o Melhor Par/Nó). Assim, quando solicitado, o BPSS pede ao serviço de agregação a lista de todos os pares que possuem esse serviço utilizando a mensagem *Query*. Após receber esta mensagem de busca, o Serviço de Agregação procura os pares SON fornecedores do serviço desejado e retorna esta lista ao BPSS.

Quando recebida esta lista, o BPSS calcula os melhores pares e retorna sua referência para o par SON que a tinha solicitado. O cálculo do melhor par é feito

utilizando um indicador definido, como por exemplo, a razão entre a distância e a largura de banda entre o par requisitante e o par que possui o serviço.

5. Abordagem proposta

Este trabalho busca estender o trabalho de [Fiorese et al. 2011], apresentando uma abordagem para a escolha do melhor par, também baseada em performance, porém utilizando outro indicador. Além da escolha do melhor par este trabalho busca também a escolha do pior par presente na rede, baseando-se no mesmo indicador.

5.1. Indicador Utilizado

Um dos pontos chave na implementação do Serviço de Seleção do Melhor Par é definir qual indicador será utilizado para a seleção dos melhores pares. Em uma P2P SON com inúmeros participantes fornecendo o mesmo serviço, escolher um par aleatoriamente pode não ser a melhor escolha, por isso, é necessário que a escolha baseie-se em alguns indicadores para selecionar e receber o serviço com qualidade e menor custo. Como exemplo, o preço cobrado pode ser utilizado como indicador para escolha do melhor par, como apresentado em [Balestrin Jr. and Fiorese 2012].

Neste trabalho, a métrica distância euclidiana utilizada por [Fiorese et al. 2011] será levada em conta. Esta métrica baseia-se no modelo desenvolvido por [Kaune, S. et al. 2009] que leva em conta a localização geográfica e o atraso entre os pares. Assim, o cálculo da distância euclidiana não leva em conta somente a distância física entre os pares, mas também as condições da rede. Além disso, outra métrica também será utilizada em conjunto com a distância euclidiana. Tal métrica será a largura de banda entre os pares. Como cada par possui dois tipos de largura de banda, largura de banda de *upload* e *download*, a escolha da largura de banda a ser utilizada será feita da seguinte maneira.

Se a largura de banda de *download* do par requisitante for superior à largura de banda de *upload* do par que possui o serviço desejado, a largura de banda a ser utilizada como parte do indicador será a de *upload* do par possuidor do serviço. Ou seja, mesmo que par requisitante possa receber o serviço com largura de banda elevada, ele somente o receberá no limite máximo de largura de banda do fornecedor.

Caso a largura de banda de *download* do par requisitante for inferior à largura de banda de *upload* do par que possui o serviço, a largura de banda a ser utilizada será a de *download* do par requisitante. Ou seja, mesmo que o par fornecedor possa fornecer o serviço com largura de banda elevada, ele somente fornecerá com o limite máximo de largura de banda que o par requisitante poderá receber.

O indicador desenvolvido é apresentado na Eq. (1).

$$\text{Indicador} = \text{Largura de Banda} / \text{Distância Euclidiana} \quad (1)$$

Deste modo, a seleção dos pares utilizando este indicador ocorre da seguinte forma. Os melhores pares serão aqueles que possuírem o maior valor para este indicador. Ao contrário destes, os piores pares serão aqueles que possuírem o valor mais baixo para este indicador.

6. Avaliação

Para avaliar o indicador utilizado no Serviço de Seleção de Melhores Pares (BPSS) e verificar a distribuição dos melhores e piores pares foram realizadas diversas simulações e os resultados estão dispostos a seguir.

6.1. Simulador e configurações utilizadas

Para realizar as simulações foi utilizado o simulador PeerfactSim.KOM [Kovačević et al. 2007]. Ele é uma plataforma de avaliação de redes Par-a-par escrito em Java que proporciona a criação de redes de sobreposição e simulação de redes P2P de larga escala [Stingl et al. 2011]. Dentre as vantagens de utilizar este simulador podem ser citadas o uso de abstrações modulares para as camadas de rede e de transporte, as quais facilitam o uso do simulador bem como a manutenção de código.

Com este simulador, foi implementado o BPSS assim como o Serviço de Agregação (AgS). Do mesmo modo que [Fiorese et al. 2011], as simulações foram constituídas de pares SON onde seus identificadores de rede fazem parte do projeto CAIDA [CAIDA, 2012] e da base de dados MaxMind GeoIP [MaxMind GeoIP, 2012]. Deste modo, os pares utilizados em cada simulação pertencem a domínios geográficos reais.

Os domínios geográficos escolhidos para as simulações foram os seguintes: Portugal, Espanha, França, Itália e Alemanha. Além disso, foram realizados 11 conjuntos de simulações. Cada conjunto de simulações contém a mesma quantidade de pares SON, divididos em cada domínio geográfico, os quais foram escolhidos aleatoriamente. Iniciando com um conjunto de 50 pares SON e o conjunto de pares AgS sendo 10% do total de pares SON, foram acrescidos sempre 25 pares SON para cada conjunto de simulações. Ou seja, iniciou com 50 pares SON e 05 pares AgS, sendo 10 pares SON e 01 par AgS pertencente a cada domínio geográfico. O próximo conjunto de simulações com 75 pares SON e 05 pares AgS, depois 100 pares SON e 10 pares AgS e assim sucessivamente até 300 pares SON e 30 pares AgS.

Além disso, a origem de cada par requisitante, o qual deseja procurar o melhor par para se comunicar, sempre será localizado no domínio geográfico de Portugal. Outro dado importante é a largura de banda de cada par, o qual possuirá entre 50 e 100Mbps de largura de banda de *download* e *upload*.

6.2. Resultados

Ao utilizar as configurações descritas anteriormente, foram desenvolvidos dois gráficos representando a distribuição dos melhores e dos piores pares fornecedores dos serviços na P2P SON, os quais podem ser vistos a partir da Figura 3, baseada na Tabela 1, e na Figura 4, baseada na Tabela 2. Cada gráfico contém os 11 grupos de simulações e, cada grupo possui cinco barras verticais, as quais representam a quantidade, de melhores pares na Figura 3 e dos piores pares na Figura 4, relativos aos domínios de Portugal, Espanha, França, Itália e Alemanha, respectivamente. Um pensamento inicial seria que o melhor par para o requisitante, localizado em Portugal, se comunicar seria algum par também localizado em Portugal, isto porque a distância euclidiana teoricamente deveria ser a menor possível, deste modo, influenciando no valor calculado pelo indicador.

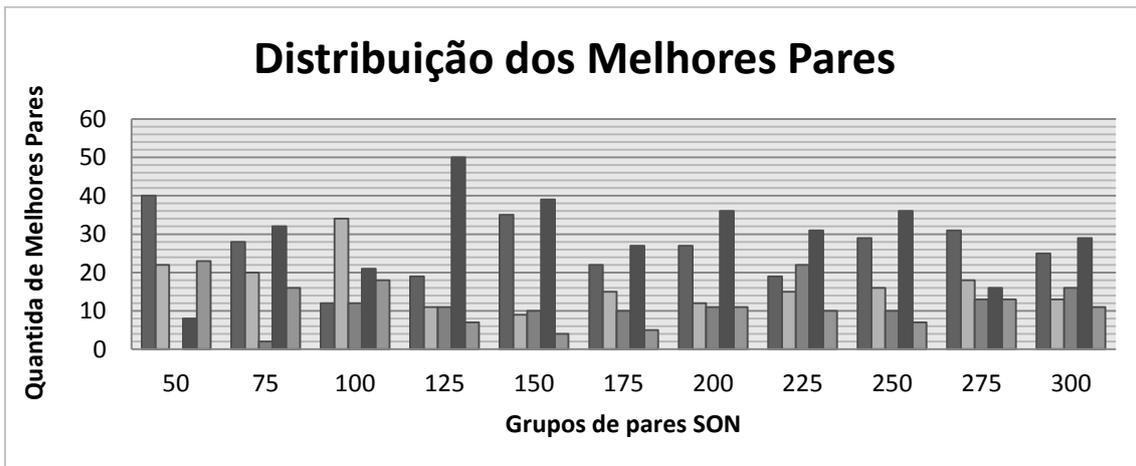


Figura 3. Distribuição dos Melhores Pares.

	Grupos de SON Peers											Média
	50	75	100	125	150	175	200	225	250	275	300	
Portugal	40	28	12	19	35	22	27	19	29	31	25	26,09091
Espanha	22	20	34	11	9	15	12	15	16	18	13	16,81818
França	0	2	12	11	10	10	11	22	10	13	16	10,63636
Itália	8	32	21	50	39	27	36	31	36	16	29	29,54545
Alemanha	23	16	18	7	4	5	11	10	7	13	11	11,36364

Tabela 1. Distribuição dos Melhores Pares.

De acordo com a Figura 3, os pares localizados na Itália foram os mais escolhidos como melhores pares. Isto pode contrariar a suposição inicial que, devido ao uso da distância euclidiana como parâmetro para a escolha do melhor par, os pares localizados em Portugal não foram escolhidos com maior frequência. Este fato também contraria outra suposição inicial que, devido a Itália estar geograficamente mais distante do par requisitante, este localizado em Portugal, os seus pares poderiam ser os menos escolhidos para comunicarem-se. Este fato ocorre devido a distância euclidiana não utilizar a distância geográfica como métrica principal, e sim, as condições da rede. Deste modo, ao mapear em um plano cartesiano as diversas métricas referente as condições da rede, um par localizado em Portugal pode estar mais próximo, por exemplo, a um par localizado na Itália do que em Portugal.

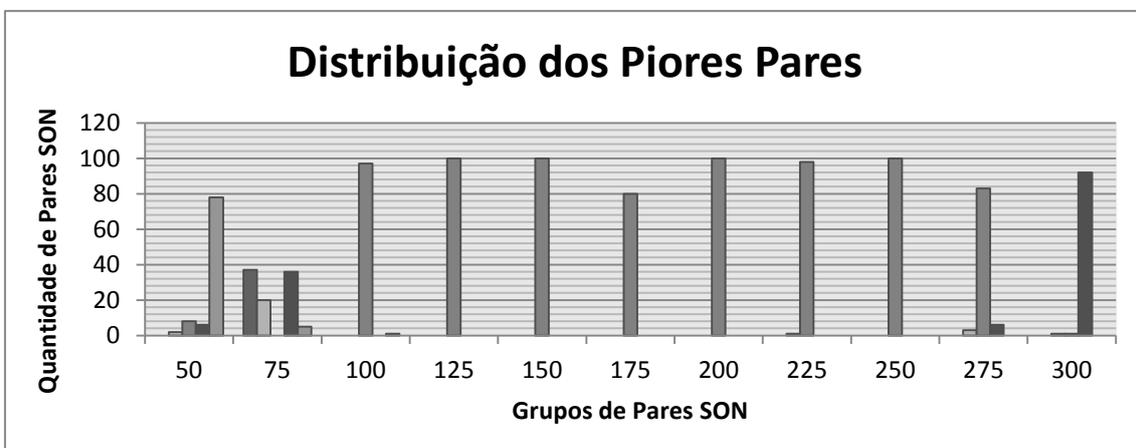


Figura 4. Distribuição dos Piores Pares.

Além disso, ao visualizarmos a Figura 4, derivada da Tabela 2, a suposição de que os piores pares estivessem mais frequentemente localizados na Itália também é contrariada. Isto porque os piores pares estiveram mais frequentemente localizados na França.

	Grupos de SON Peers											Média
	50	75	100	125	150	175	200	225	250	275	300	
Portugal	0	37	0	0	0	0	0	0	0	0	0	3,363636
Espanha	2	20	0	0	0	0	0	1	0	3	1	2,454545
França	8	0	97	100	100	80	100	98	100	83	1	69,72727
Itália	6	36	0	0	0	0	0	0	0	6	92	12,72727
Alemanha	78	5	1	0	0	0	0	0	0	0	0	7,636364

Tabela 2. Distribuição dos Piores Pares.

A Figura 5 apresenta esses dados em termos do número médio de melhores e piores pares selecionados por domínio geográfico.

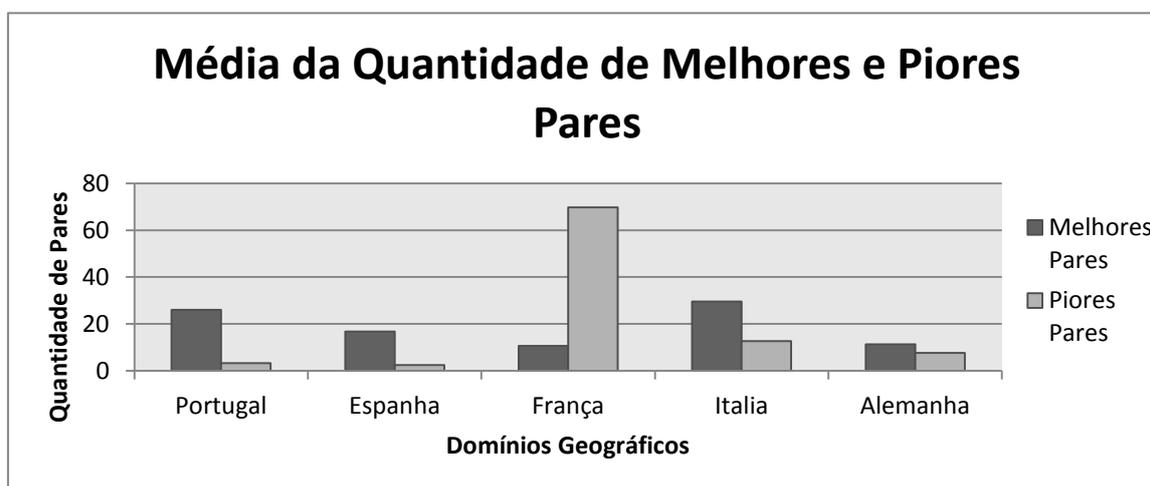


Figura 5. Média da Quantidade de Melhores e Piores Pares

Assim, analisando a Figura 5, pode-se perceber que os melhores pares estão localizados com mais frequência na Itália e com menos frequência na França. Ainda analisando a Figura 5, pode-se perceber que no domínio geográfico França, além de apresentar com menos frequência os melhores pares, apresenta na média, os piores pares. Isto pode acontecer caso estes provedores disponibilizem baixa largura de banda e tenham distância euclidiana elevada, podendo acarretar prejuízos relacionados ao desempenho do serviço fornecido. Exemplo disto pode ser maior tempo para transferência dos dados ou então alta taxa de retransmissões dos pacotes devido a perdas de pacotes.

Como pode-se observar na Figura 4 e 5, o domínio geográfico da França apresentou a maior quantidade de piores de pares de acordo com o indicador utilizado. Essa observação analisada com profundidade revela a grandeza das métricas individuais envolvidas no processo de produção do indicador. A demonstração desse fato, apresentada na Figura 6 pela plotagem dos dados do domínio geográfico França onde a largura de banda é plotada em Mbps e a distância euclidiana em euclidianos, indica que a partir dos experimentos envolvendo 125 pares SON tanto a métrica distância

euclidiana quanto largura de banda apresentaram os maiores valores dentre todos os demais domínios.

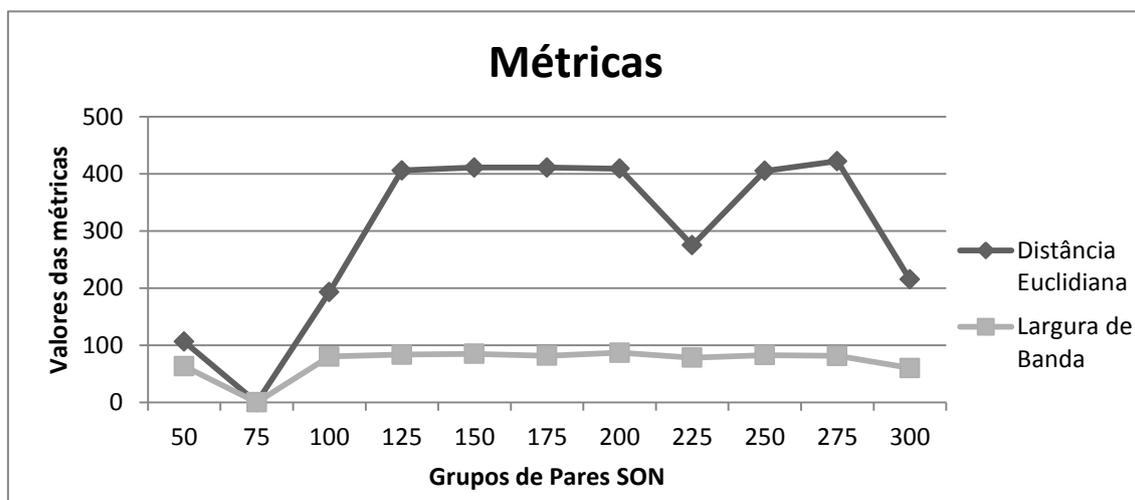


Figura 6. Valores para as Métricas dos piores pares localizados na França

Além disto, entre 125 e 200 pares SON, nenhum outro domínio geográfico teve um par escolhido como sendo o pior, devido a estes altos valores e a possibilidade de que a partir de 125 pares SON ter sido incluso um par com estes altos valores e mesmo aumentando o grupo de pares SON, este continuar sendo escolhido como pior par. Entretanto, no grupo de 225 pares SON, tal par pode ter recebido alto valor para a largura de banda, mesmo mantendo seu alto valor para a distância euclidiana, possibilitando que outros pares possam ter sido escolhidos por mais vezes mesmo no domínio geográfico França.

Por fim, o que acontece com o grupo de 225 pares SON acontece novamente com o grupo de 300 pares SON. Porém, nesse caso, os pares de outros domínios geográficos são escolhidos mais vezes que os localizados na França. Isto porque, com o acréscimo de mais pares na SON e a alteração da largura de banda dos pares que já existiam na rede faz com que exista maior concorrência entre os pares e que nem sempre o mesmo par seja escolhido para todos os grupos de pares SON.

7. Conclusão

A partir deste artigo foi possível realizar um comparativo da seleção de pares presentes em uma P2P SON através do uso de um indicador desenvolvido para a escolha dos mesmos. Com este indicador foi possível verificar a disposição dos melhores e piores pares que estavam na P2P SON, localizados em cinco diferentes domínios geográficos.

Para isto foram realizadas diversas simulações e como resultado, foram apresentados diversos gráficos contendo informações que auxiliaram na formulação de possíveis respostas para o problema proposto. Dentre as respostas podemos visualizar que os serviços fornecidos no mesmo domínio geográfico em que o par consumidor está localizado é, em segundo lugar, escolhido como melhor par para a comunicação, o que pode ser resultado da influencia de menor largura de banda em relação aos outros domínios geográficos.

Outro dado apresentado foi uma comparação entre a distribuição dos melhores e dos piores pares. Baseando-se nos resultados desta comparação, pode-se visualizar que

o domínio geográfico que obteve menor quantidade de melhores pares, também obteve maior quantidade de piores pares.

A partir desta abordagem, outras podem ser aplicadas à solução da seleção adequada de um provedor de serviço. Por exemplo, a utilização de um indicador referente a um conjunto de métricas relativas ao serviço combinadas com a capacidade de fornecimento do serviço pelos provedores; e, até mesmo unir a abordagem deste artigo com outras, como a citada anteriormente, de forma a construir um indicador abrangente e escalável.

8. Referências

Balestrin Jr., R. and Fiorese, A. (2012). Abordagem de Seleção dos Melhores Pares em Rede de Sobreposição de Serviços Par-a-Par Baseada em Preço do Serviço.

CAIDA (2012). Macroscopic topology measurements. .
<http://www.caida.org/projects/macrosopic/>.

Esen, E. (feb 2008). How to create an Overlay in PeerfactSim.KOM.

Fiorese, A., Simões, P. and Boavida, F. (2010). OMAN: a management architecture for P2P service overlay networks. In *Proceedings of the Mechanisms for autonomous management of networks and services, and 4th international conference on Autonomous infrastructure, management and security*, AIMS'10. Springer-Verlag. <http://dl.acm.org/citation.cfm?id=1875873.1875878>, [accessed on May 7].

Fiorese, A., Simões, P. and Boavida, F. (oct 2011). An approach to peer selection in service overlays. In *2011 7th International Conference on Network and Service Management (CNSM)*.

Kaune, S., Pussep, K., Leng, C., et al. (feb 2009). Modelling the Internet Delay Space Based on Geographical Locations. In *Parallel, Distributed and Network-based Processing, 2009 17th Euromicro International Conference on*.

Kovačević, A., Kaune, S., Liebau, N., Steinmetz, R. and Mukherjee, P. (sep 2007). Benchmarking Platform for Peer-to-Peer Systems (Benchmarking Plattform für Peer-to-Peer Systeme). *it - Information Technology*, v. 49, n. 5, p. 312–319.

Lee, S.-I., Yi, J.-H. and Kang, S.-G. (dec 2011). Context-Aware Service Overlay Network: Concept and Case Study. In *Services Computing Conference (APSCC), 2011 IEEE Asia-Pacific*.

MaxMind GeoIP (2012). MaxMind GeoIP: MaxMind Geolocation Technology. .
<http://www.maxmind.com/en/home>.

Stingl, D., Gross, C., Ruckert, J., et al. (jul 2011). PeerfactSim.KOM: A simulation framework for Peer-to-Peer systems. In *High Performance Computing and Simulation (HPCS), 2011 International Conference on*.