Approach for Service Search and Peer Selection in P2P Service Overlays

Adriano Fiorese*†
*UDESC, University of the State of Santa Catarina
CCT, Department of Computer Science - DCC
89223-100, Joinville-SC, Brazil
e-mail:fiorese@joinville.udesc.br

Paulo Simões†, and Fernando Boavida†
†CISUC, University of Coimbra
Pólo II, Department of Informatics Engineering - DEI
3030-290 Coimbra, Portugal
e-mail: {fiorese,psimoes,boavida}@dei.uc.pt

***Abstract*—Peer-to-peer service overlay networks (P2P SON) are increasingly being infrastructure providers for networking services, allowing service providers to cooperatively offer and run a flexible set of services. Regarding this condition, the selection of peers is a key issue for improving resource usage; service performance, and ultimately end users Quality of Experience (QoE). This paper presents an approach to best peer selection in a three-tier P2P SON architecture, allowing the splitting of service business functions and peer selection functions. The proposed best peer selection approach is evaluated by simulation, using a literature available geographic positioning metrich that takes into account real delay and jitter made available by the CAIDA project and MaxMind's free database. The simulation results show the consistency and good performance of the proposed peer selection approach.**

***Index Terms*—Services management, P2P, Peer Selection**

I. Introduction

Services are becoming one of the primary sources of revenue on the Internet. Currently, data transport is being sold as a commodity for a broad audience. This business model allows rising a wide service consumer market. On the other hand, it allows the appearing of new service providers at the stage. This also leads to a competition for the service consumers among the whole set of service providers, eventually leading to a global competition depending on the quality and originality of the offered service.

In this context, service providers can enhance their ability to make their service or service components available to a broader set of customers, through the utilization of a Service Overlay Network (SON) [1], [2]. In this case, a SON acts as an infrastructure where services are published/offered and to which the users access in order to select and use these services. Moreover, a SON can be created by a consortium of service providers using the Internet to make their services available to the user community at large. The peer-to-peer (P2P) technology is a well suited way for constructing that kind of SON. This technology leads to a self-organizing overlay and, additionally, to sharing maintenance costs among service providers.

This work was partially funded by FCT under scholarship contract SFRH/BD/45683/2008.

Although P2P eases the construction and maintenance of SONs, it does not guarantee adequate service performance. In order to maximize performance, the best peer must be found in the P2P SON, among all the potential partners that provide the desired service. Naturally, the choice of best peer should take into account one or more of a set of Quality of Service (QoS) parameters, such as delay, jitter, available bandwidth, etc.. Nevertheless, to minimize inter-provider traffic and, thus, reduce costs for the user and for the service provider, the choice of peers belonging to a different, remote domain should be avoided as much as possible. Thus, locality should also be taken into consideration when choosing a SON peer as the best peer to serve a requested service.

Previous work from the same authors [3]–[5] proposed an architecture for services management in P2P Service Overlay Networks (SON). The architecture, named OMAN, takes into account several aspects of services management, particularly the use of a second overlay - called Aggregation Service (AgS) - to provide efficient service search in the context of multi-domain P2P SONs. In its turn, this paper deals with the proposal and assessment of a third-tier component of OMAN, whose purpose includes the searching and selection of the best peer with which a service-requesting peer should interact in the context of the P2P SON.

Having in mind the stated goal and approach, this paper is organized as follows. Section II discusses related work and also provides the context of the current work, by briefly presenting the OMAN architecture, which the BPSS component belongs to. Section III describes the proposed BPSS service. Subsequently, Section IV presents the evaluation and discusses the simulation results, after describing the simulated scenarios. Finally, Section V summarizes the contributions and presents further work.

II. Related Work

A. *Service Overlay Networks*

According to Tran and Dziong [1], a Service Overlay Network (SON) is a network composed of interconnected nodes, whose generic purpose is to provide the required Quality of Service (QoS) to applications that execute on those nodes. The same authors establish a difference between SON and P2P overlay network claiming that the purpose of the latter is related with providing efficient searching and retrieval.

The bandwidth provisioning in a SON composed of nodes that lease links from several link providers is studied in [6].

We advocate a P2P overlay network can also provide QoS services when the participants are in a consortium of service providers that establish well-defined SLAs to regulate the contribution of each participant to the network. In this sense, a platform called ALASA is presented in [2]. It uses a structured P2P overlay network over the Internet to describe, discover, compose, and compute the reputation of services.

P2P is also used in [7] as support for the SON architecture. In that piece of work the authors address discovery of services considering QoS aspects in their approach.

B. Peer Selection

Haase et. al [8] explores neighboring peers relationships and shared peers expertise in order to select peers. The use of artificial intelligence techniques, like machine learning, is another approach to peer selection, which also takes advantage of the peers' expertise [9]. This latter work aims at adapting the selection process to the peers' requirements.

In file sharing, the free-riding problem encourages the adoption of incentive mechanisms as part of the selection scheme. Thus, the fairness between uploads and downloads is used as a metric to the best peer selection. Bittorrent [10] is an emblematic example for this. However, our scheme does not target file(data)-sharing environments. Rather, the problem is to select the best peer that satisfies the requirements of the intended service. Therefore, in our case, performance instead fairness is used. In addition, unlike file sharing applications, our approach considers the best peer selection process for long lasting sessions, as opposed to relatively short burst chunk downloads/uploads.

In the P2P multimedia stream services field, several pieces of work proposing the use of P2P as the delivery mechanism face peer selection issues [11], [12]. Similarly to our approach they aim at optimizing peer choice regarding the performance of the service.

Furthermore, P2P traffic is an issue faced by ISPs and over the Internet in general. In order to assist ISPs in avoiding costs with the choice of best peers out of their own domains, several proposals have been put forward [13]–[15]. These advocate the collaboration between providers and P2P applications. On the other hand, our BPSS scheme selects best peers for service interactions inside the geographical domain of the service requester, without the need for explicit collaboration between the service providers.

C. OMAN

In order to contextualize the Best Peer Selection Service (BPSS), this section briefly presents the underlying OMAN architecture, previously proposed by the authors in [3]–[5].

OMAN is a P2P SON architecture that handles aspects ranging from the composition of the SON until the interaction aspects between the services and the SON, including how to take advantage of the information at the P2P overlay level to leverage the services and applications. Fig. 1 provides an overview of the OMAN architecture.

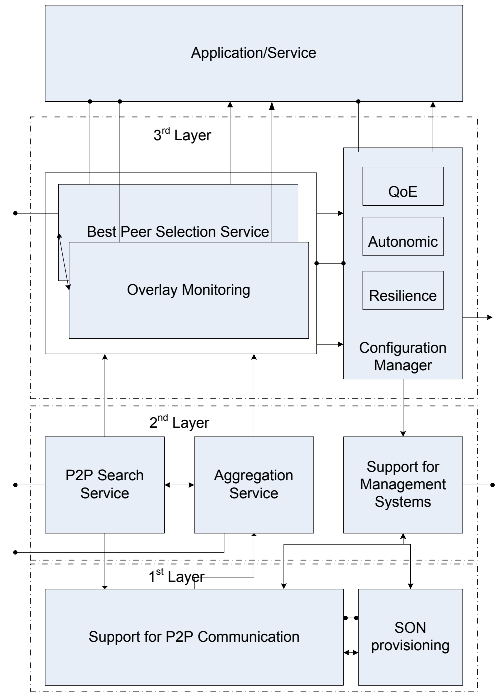

Fig. 1. The OMAN architecture [4]

The lower layer of the architecture is the basic P2P SON layer. A service provider can make available more than one peer to be represented at the P2P SON layer.

The central module of the second tier of OMAN is the Aggregation Service (AgS) [3], [5]. AgS is an unstructured P2P overlay-tier, meaning there is no tight coupling between overlay topology and information location/placement. AgS executes on top of the P2P SON and it consists of peers that belong to potential providers interested in advertising/offering their services. The purpose of AgS is to aggregate the offerings of services and service components. This is accomplished by concentrating the service offerings in its peers (nodes), in order to facilitate and optimize the services searching process.

The peers that form the AgS P2P overlay are called *aggregation peers (AgS peers)*. They are chosen among the peers that form the underlying P2P SON tier, which makes them specialized *SON peers*.

Each SON peer plays a double role: to execute the services (as any other P2P SON peer) and to maintain and publish references to the available services. Note that service references are made available (i.e., published) to AgS peers in order to optimize service searching. Each SON peer involved in the AgS can take care of several service offerings. A single service offering can be spread over multiple AgS peers in order to allow some redundancy and to overcome churn. These aggregation peers may be located at the same domain as the SON peers or, in some cases, at different domains.

Searching for a service, using the AgS framework, therefore results in a set of references to SON peers that offer an interface to services matching the search criteria. This preserves the internal details of the service, since the external entity, i.e. a user, a third party service provider, or other SON peer in the P2P SON, is only granted with a mediated access (by means of the SON peer), which may hide sensitive information and filter undesired operations.

The Best Peer Selection Service (BPSS) is one of the modules of the third tier of the OMAN architecture. The third layer is the most specialized layer on the OMAN architecture and it interfaces the application. Therefore, BPSS is cornerstone for the architecture and for services provided using the envisaged business model made available by P2P SON.

III. Best Peer Selection Service

The objective of the Best Peer Selection Service (BPSS) is to provide SON peers with the identification of the best peer for a particular service, in the context of services offering in a P2P SON. Naturally, the best peer depends on several aspects, including the service objectives and characteristics. Key to the choice of a suitable peer is the intended service performance, which is tied not only to measurable networking characteristics, such as bandwidth, delay, loss, etc., but also to the peers underlay location.

A. BPSS Architecture

Regarding the use on OMAN, service developers can implement an interface with the BPSS module in order to request and receive best peer information, allowing the splitting of service business functions and best peer determination. This decoupling enhances modularity and best peer selection metric independence, thus leading to high flexibility when choosing the particular metric to use for a particular service type. In order to take advantage of this aspect, the BPSS builds on the AgS [5] service toward select the best peer on request of a P2P SON peer.

Fig. 2 illustrates the use of BPSS. SON peers can request best peer information (select_BP), regarding a particular service, from the BPSS module. On the reception of a best peer request, the BPSS module asks the AgS service the list of all SON peers that have published a service profile for the intended service. After receiving the requested list, the BPSS module calculates the best peer and returns its reference to the requesting SON peer. The selection of the best peer is done using one of the supported metrics. In Section III-C one of such metrics is presented.

It is worth mentioning that with this decoupled approach it is also possible for an external entity (e.g. a user or a particular application/service or service component from outside the P2P SON) to request a best peer selection, as long as the request is compliant with the BPSS interface and the requested metric is supported.

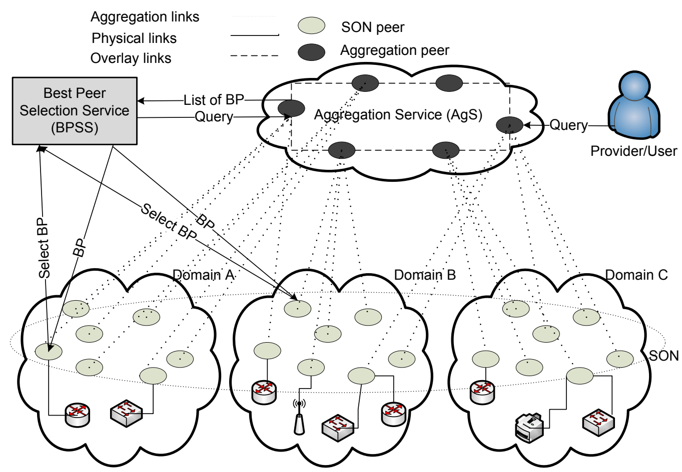

Fig. 2. BPSS Architecture

B. BPSS Model

Based on its architecture, BPSS comprises the model depicted in Fig. 3. The BPSS model is represented as a UML-like diagram. The BPSS model comprises the roles of the involving parts; their interactions; and the information exchanged among them.

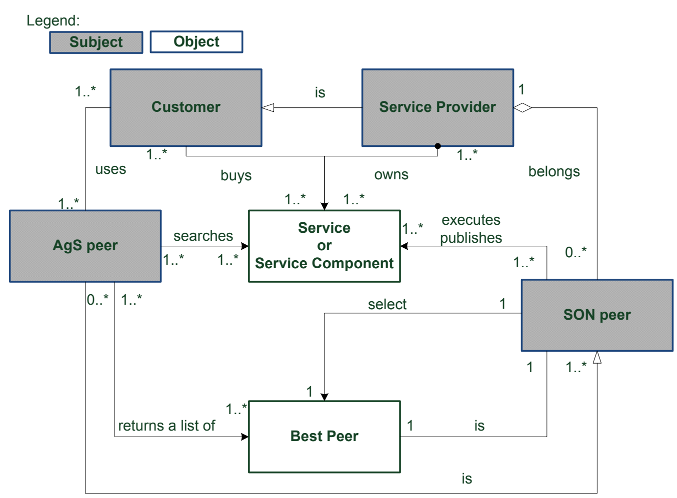

Fig. 3. BPSS Model

The service or service component object and the best peer are the central pieces in the BPSS model. They make part of the interactions with every subject in the model. The service object represents the Service or Service Component that actually executes on the SON peer and that is published in the aggregation peers. Aggregation peers are the peers that form the AgS P2P overlay. The Best Peer object is a SON peer resulting of a best peer selection started by a particular SON peer.

The Service Provider is another cornerstone piece in the BPSS model. It is the owner of the Service or Service Component as well as the SON peers and the AgS peers (since AgS peers are specialized SON peers). Service Providers are

Customers as well, so they can act as third party consumers of service components of other service providers. In this case, this transaction can involve service composition and provisioning, which is out of the scope of this paper. They also can use third party services in a service chain, in order to offer to a home user (customer) a complete set of services. Nonetheless, to accomplish this step, the service provider needs to find the best peer to which its SON peer will interact in order to offer that set of services. Following the BPSS model, the Customer, which can be a third Service Provider, uses the AgS peer to search for services in the AgS service. The AgS peer by its turn returns a list of all SON peers that executes the searched services. Ergo, the SON peer then selects the best one to be used. The interaction among the services involved is not covered by AgS either BPSS.

C. BPSS Metric

The proposed BPSS service is based on one or more service performance metrics to determine best peers. In other words, the peer selected as the best peer should offer optimized interaction with the requesting service/peer, according to some performance criteria. Although BPSS provides independence from the used performance metrics, for the purpose of implementation and evaluation a specific metric was used.

Several performance criteria can be used when determining the best peer. Performance of P2P systems is often very sensitive to the underlying delay characteristics. These are influenced, among other factors, by bandwidth, load and also geographic location. In fact, according to [16], the geographical location of nodes heavily influences jitter and packet loss. This observation points to the need for the node's geographic location when developing a delay prediction model. Having this in mind, the authors of [16] developed a predictive model of the Internet delay space that takes into account the geographical location of the nodes and the delay between them.

Using a rich set of real data, namely measured end-to-end round trip time (rtt) [17] and measured end-to-end link jitter [18], those authors mapped the measured end-to-end nodes into a 5-dimensions Euclidean space model of the Internet, by combining this information with global network positioning information [19]. Using the coordinates of each peer in this 5-dimensions model, it is then possible to calculate the Euclidean distance between peers, which takes into account not only network conditions but also peer location.

For the purpose of the work presented in this paper, BPSS used the mentioned distance metric in order to select the best peer for a particular SON peer.

IV. Evaluation Framework

In order to assess the BPSS behavior, we conducted a simulation study to observe the best peer and the second-best peer selection distributions over well-constrained geographical domains.

A. Used Simulator

The PeerFactSim.KOM [20] discrete event simulator was used in all simulations. Unlike mainstream network simulators, the PeerFactSim.KOM simulator was specifically designed to simulate P2P networks. Among its advantages are the use of modular abstractions for the network and transport layers, and the utilization of well-known and useful software engineering design-patterns. These characteristics ease the experimentation and facilitate code reutilization and maintenance.

B. Simulations setup

The simulation environment was constituted of SON peers whose network identifiers (IP addresses) were taken from the CAIDA project and MaxMind GeoIP database. Therefore, the simulated peers belong to real geographical domains. The Internet delay space scheme presented in Section III-C was used.

The scenarios were modeled based on real nodes available in the compiled data set. To accomplish that, peers belonging to specific geographical country domains were split between SON and aggregation peers. The number of AgS peers was 10% of the total number of SON peers used in each domain. Thus, if a country had 50 SON peers at the P2P SON - which, for instance, can mean that it comprised 50 service providers - then 5 AgS peers belonging to that country would be part of the Aggregation Service. The chosen geographical domains were the following European countries: Portugal, Spain, France, Italy and Germany.

The simulation comprised 11 sets of individual simulations. Each set simulated a particular number of SON and AgS peers. The initial set simulated a scenario with the total of 50 SON peers, corresponding to 10 randomly chosen SON peers from each of the aforementioned countries. The second simulated set comprised the total of 75 SON peers (15 for each country). Thus, with steps of 25 SON peers between each simulated scenario, the last simulated set comprised 300 SON peers.

C. Simulations strategy

For the sake of simplicity, without loss of generality, a particular SON peer could only offer, at most, seven randomly chosen services or service components (using a uniform distribution) from the service set S={S1,S2,S3,S4,S5,S6,S7}. In addition, each SON peer could only publish its service subset on, at most, 10 distinct randomly-chosen aggregation peers (also following a uniform distribution). Nevertheless, it was possible that more than one SON peer could offer the same services and publish them on the same or different aggregation peer.

Each experiment execution had simulated 50 hours of work and it had been repeated 10 times in order to get averaged values. Every operation (e.g. joining, leaving, publishing, searching, and select) was specified in time. Each simulation executed 100 searching operations, i.e., 100 best peer selection requests, since a searching operation is triggered by a best peer selection request.

The SON peers that execute this selection process are chosen randomly using a uniform distribution. This happens when the number of SON peers composing the P2P SON is greater than the 100 best peer select operations, which composes each experiment. Otherwise, each SON peer executes that operation at least once.

The experiments also had chosen second-best peers. They are the peers whose performance regarding the used metric puts them in the second place in an hypothetical ordered list of best peers. Actually, the second-best peer is selected from the same list of SON peers provided by the AgS service, by removing the best peer from that list and repeating the measurement process.

The determination of the second-best peer can aid in the validation of the used metric in two ways: 1) by checking if the service requesting the best peer behaves better in an interaction with the selected best peer, then this provides a measure of the metrics consistency; 2) by measuring the average improvement of the best peer over the second-best peer, an indication of the metrics effectiveness can be obtained.

The wide range of simulated P2P SON sizes intends to cover scenarios with few service providers (e.g. small P2P SON for very specialized services) until scenarios composed of many service providers (e.g. a more competitive scenario).

D. Results

The results presented in this section show the distribution of best and second-best peers regarding requests made by SON peers belonging to the Portuguese geographical domain only, though it is obvious that the used methodology can be applied to any other geographical domain.

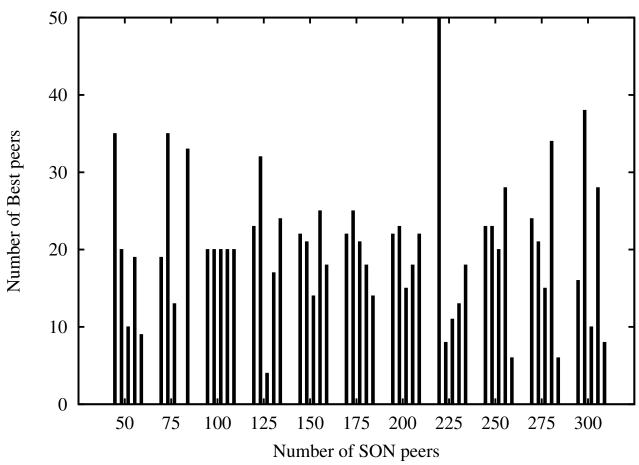

Fig. 4. Clustering Best Peer Ocurrences by Domain by Number of SON peers

Fig. 4 depicts the geographical location of the SON peers selected as best peers. There are eleven 5-bar clusters, each one corresponding to one of the eleven simulated scenarios. In each cluster, each of the five bars represents the number of best peer occurrences in each of the five geographical domains, namely Portugal, Spain, France, Italy, and Germany, respectively.

One can expect the highest number of selected best peers is in the domain of the requester (Portugal, in the case of these simulations), due to geographical distance considerations. Nevertheless, the obtained results clearly show the effect of two key aspects of the OMAN architecture and of its AgS and BPSS services: on one side, in some cases, service searching performed by the AgS service determined that the desired services were not available at any of the SON peers of the requester's domain; on the other hand, the metric used by the BPSS service - based on the Internet model proposed in [16], which takes into account not only the geographical position but also delay and jitter - led to the fact that the closest peer, in terms of the 5-dimension Euclidean distance, resided in a different geographical domain.

Similar results building on the same explanations are obtained regarding the experiments with the second-best peers. Fig. 5 depicts them.

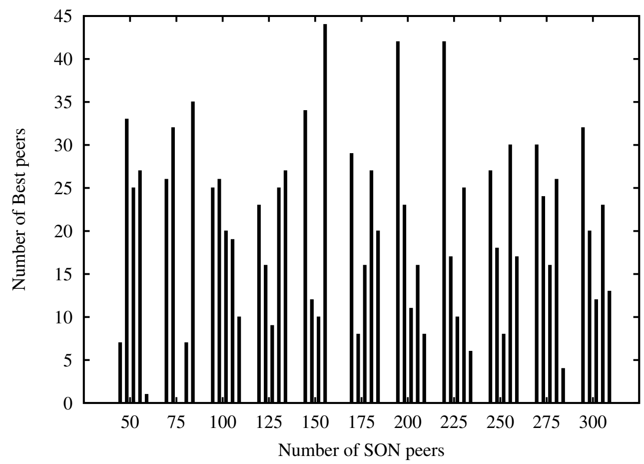

Fig. 5. Clustering Second-Best Peer Ocurrences by Domain by Number of SON peers

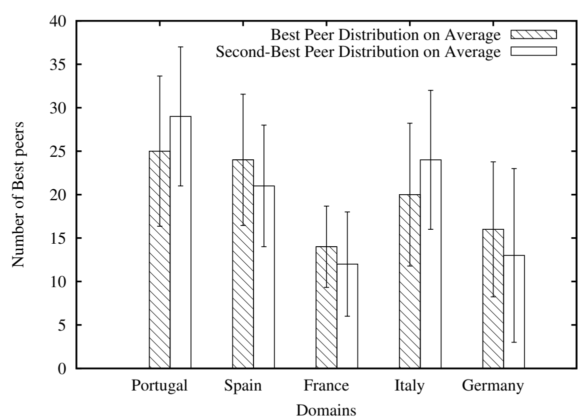

Fig. 6. Average Peer Distribution by Geographical Domain

Nevertheless, even with the mentioned constraints regarding the statistical availability (or unavailability) of the desired services in the requester's domain, averaging the results of all simulations shows that the highest number of best peers was selected in the same domain of the requester. This can be seen in Fig. 6. It is worth mentioning the results rely on and are presented based on a confidence interval of 95% for the mean number of best peer selections on each geographical domain.

Therefore, taking the sum of best peers and second-best peers by domain, SON peers in the requester's domain (Portugal) were selected as best peers in 27% of the time, followed by Spain (22.5%), Italy (22%), Germany (14.5%) and France (13%). This means that almost half the best peer selections resulted in peers belonging to the same geographical domain or to the neighboring geographical domain. This suggests the consistency of the used metric and the good operation of OMAN's AgS and BPSS services.

It is also worth mentioning that the BPSS overhead is negligible. The time taken by the selection process for both the best and second best peer is around 2ms. On the other hand, the overhead associated with the maintenance of the P2P SON and with the AgS service is around 164ms. These values strongly suggest the efficiency, feasibility, and practicality of the proposed BPSS service, in conjunction with AgS and the other modules of the OMAN architecture.

V. Conclusion

In this paper, an approach to the problem of best peer selection in peer-to-peer service overlay networks has been proposed and studied. The proposed Best Peer Selection Service (BPSS) is one of the components of a three-tier service management overlay network architecture (OMAN).

After identifying key issues in related work and providing an overview of the underlying OMAN architecture for the sake of contextualization, the BPSS service was presented. The main features of the presented BPSS are the use of a very efficient aggregation service (AgS) for service searching, and its independence from the best peer selection metric. The latter feature provides flexibility, adaptability and modularity to the overall process of best peer selection.

In order to test and assess the BPSS behavior, a specific distance metric was used. The used metric combines measured delay and jitter real data with geographical location data. The performed simulations involved the selection of best and second-best peers in a universe of five distinct geographical domains.

Obtained results have shown that BPSS performs well and that the overall OMAN architecture - of which the AgS service is a key component - is very effective. Simulations have also shown that the BPSS overhead is negligible and the AgS overhead, which derives from its operations, is very low.

Further work can compare different best peer selection metrics, based not only on different performance parameters, but also on factors such as inter-provider link cost. Still another line of research can be the use of the OMAN approach by service providers in order to identify ways of maximizing user quality of experience or ways of reducing inter-provider traffic, e.g., by deploying specific peers inside their own domain.

References

[1] C. Tran and Z. Dziong, "Service overlay network capacity adaptation for profit maximization," *IEEE Transactions on Network and Service Management*, vol. 7, no. 2, pp. 72–82, Jun. 2010.

[2] S. Zhou, M. Hogan, S. Ardon, M. Portman, T. Hu, K. Wongrujira, and A. Seneviratne, "ALASA: when service overlay networks meet peer-to-peer networks," in *Communications, 2005 Asia-Pacific Conference on*, Perth, WA, USA, 2005, pp. 1053–1057.

[3] A. Fiorese, P. Simões, and F. Boavida, "Service searching based on P2P aggregation," in *Proceedings of International Conference on Information Networking 2010 (ICOIN 2010)*, Busan - South Korea, 2010.

[4] ——, "OMAN – a management architecture for P2P service overlay networks," in *Mechanisms for Autonomous Management of Networks and Services*, ser. Lecture Notes in Computer Science, vol. 6155. Zurich, Switzerland: Stiller, Burkhard; De Turck, Filip, Jun. 2010, pp. 14–25.

[5] ——, "An aggregation scheme for the optimisation of service search in Peer-to-Peer overlays," in *Network and Service Management (CNSM), 2010 International Conference on*, Niagara Falls, Canada, Oct. 2010, pp. 481–486.

[6] Z. Duan, Z. Zhang, and Y. T. Hou, "Service overlay networks: SLAs, QoS, and bandwidth provisioning," *IEEE/ACM Transactions on Networking*, vol. 11, no. 6, pp. 870–883, Dec. 2003.

[7] E. Lavinal, N. Simoni, M. Song, and B. Mathieu, "A next-generation service overlay architecture," *Annals of Telecommunications*, vol. 64, no. 3, pp. 175–185, Apr. 2009, publisher: Springer Paris. [Online]. Available: http://dx.doi.org/10.1007/s12243-008-0082-x

[8] P. Haase, R. Siebes, and F. Harmelen, "Expertise-based peer selection in Peer-to-Peer networks," *Knowledge and Information Systems*, vol. 15, no. 1, pp. 75–107, Jan. 2007.

[9] D. S. Bernstein, Z. Feng, B. N. Levine, and S. Zilberstein, "Adaptive peer selection," in *In Proceedings of the 2nd International workshop on Peer-toPeer Sustems (IPTPS03)*, 2003.

[10] K. Huang, L. Wang, D. Zhang, and Y. Liu, "Optimizing the BitTorrent performance using an adaptive peer selection strategy," *Future Generation Computer Systems*, vol. 24, pp. 621–630, jul 2008.

[11] A. Habib and J. Chuang, "Service differentiated peer selection: an incentive mechanism for peer-to-peer media streaming," *Multimedia, IEEE Transactions on*, vol. 8, no. 3, pp. 610–621, 2006.

[12] C. Wu and B. Li, "Optimal peer selection for minimum-delay peer-to-peer streaming with rateless codes," in *Proceedings of the ACM workshop on Advances in peer-to-peer multimedia streaming*. New York, NY, USA: ACM, 2005, pp. 69–78.

[13] V. Aggarwal, O. Akonjang, and A. Feldmann, "Improving user and ISP experience through ISP-aided P2P locality," in *INFOCOM Workshops 2008, IEEE*, 2008, pp. 1–6.

[14] D. R. Choffnes and F. E. Bustamante, "Taming the torrent: a practical approach to reducing cross-isp traffic in peer-to-peer systems," in *ACM SIGCOMM Computer Communication Review*, vol. 38. New York, NY, USA: ACM, Aug. 2008, pp. 363–374.

[15] H. Zhuang, Y. Xu, Y. Hu, and X. Lin, "A peer selection mechanism in P2P networks based on the collaboration of ISPs and P2P systems," in *Information Science and Engineering, International Conference on*, vol. 0. Los Alamitos, CA, USA: IEEE Computer Society, 2009, pp. 1573–1576.

[16] S. Kaune, K. Pussep, C. Leng, A. Kovacevic, G. Tyson, and R. Steinmetz, "Modelling the internet delay space based on geographical locations," in *Parallel, Distributed and Network-based Processing, 2009 17th Euromicro International Conference on*, 2009, pp. 301–310.

[17] "Macroscopic topology measurements," http://www.caida.org/projects/macroscopic/.

[18] "The PingER project," http://www-iepm.slac.stanford.edu/pinger/.

[19] S. Lee, Z. Zhang, S. Sahu, and D. Saha, "On suitability of euclidean embedding of internet hosts," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 34. New York, NY, USA: ACM, Jun. 2006, pp. 157–168.

[20] A. Kovacevic, S. Kaune, N. Liebau, R. Steinmetz, and P. Mukherjee, "Benchmarking platform for peer-to-peer systems," *it - Information Technology*, vol. 49, no. 5, pp. 312–319, Sep. 2007.