Internet Research Task Force (IRTF)
Request for Comments: 6029
Category: Informational
ISSN: 2070-1721

I. Rimac
V. Hilt
M. Tomsu
V. Gurbani
Bell Labs, Alcatel-Lucent
E. Marocco
Telecom Italia

October 2010

A Survey on Research on the Application-Layer Traffic Optimization (ALTO) Problem

Abstract

A significant part of the Internet traffic today is generated by peer-to-peer (P2P) applications used originally for file sharing, and more recently for real-time communications and live media streaming. Such applications discover a route to each other through an overlay network with little knowledge of the underlying network topology. As a result, they may choose peers based on information deduced from empirical measurements, which can lead to suboptimal choices. This document, a product of the P2P Research Group, presents a survey of existing literature on discovering and using network topology information for Application-Layer Traffic Optimization.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Research Task Force (IRTF). The IRTF publishes the results of Internet-related research and development activities. These results might not be suitable for deployment. This RFC represents the consensus of the Peer-to-Peer Research Group of the Internet Research Task Force (IRTF). Documents approved for publication by the IRSG are not a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6029.

Table of Contents

1. Introduction

A significant part of today's Internet traffic is generated by peer-to-peer (P2P) applications, used originally for file sharing, and more recently for real-time multimedia communications and live media streaming. P2P applications pose serious challenges to the Internet infrastructure; by some estimates, P2P systems are so popular that they make up anywhere between 40% and 85% of the entire Internet traffic [Karagiannis], [LightReading], [LinuxReviews], [Parker], [Glasner].

P2P systems ensure that popular content is replicated at multiple instances in the overlay. But perhaps ironically, a peer searching for that content may ignore the topology of the latent overlay network and instead select among available instances based on information it deduces from empirical measurements, which in some particular situations may lead to suboptimal choices. For example, a shorter round-trip time estimation is not indicative of the bandwidth and reliability of the underlying links, which have more of an influence than delay for large file transfer P2P applications.

Most Distributed Hash Tables (DHT) -- the data structures that impose a specific ordering for P2P overlays -- use greedy forwarding algorithms to reach their destination, making locally optimal decisions that may not turn out to be globally optimized [Gummadi]. This naturally leads to the Application-Layer Traffic Optimization (ALTO) problem [RFC5693]: how to best provide the topology of the underlying network while at the same time allowing the requesting node to use such information to effectively reach the node on which the content resides. Thus, it would appear that P2P networks with their application-layer routing strategies based on overlay topologies are in direct competition against the Internet routing and topology.

One way to solve the ALTO problem is to build distributed application-level services for location and path selection [Francis], [Ng], [Dabek], [Costa], [Wong], [Madhyastha] in order to enable peers to estimate their position in the network and to efficiently select their neighbors. Similar solutions have been embedded into P2P applications such as Vuze [Vuze]. A slightly different approach is to have the Internet service provider (ISP) take a proactive role in the routing of P2P application traffic; the means by which this can be achieved have been proposed [Aggarwal], [Xie], [Saucez]. There is an intrinsic struggle between the layers -- P2P overlay and network underlay -- when performing the same service (routing); however, there are strategies to mitigate this dichotomy [Seetharaman].

This document, initially intended as a complement to RFC 5693 [RFC5693] and discussed during the creation of the IETF ALTO Working Group, has been completed and refined in the IRTF P2P Research Group. Its goal is to summarize the contemporary research activities on the Application-Layer Traffic Optimization problem as input to the ALTO working group protocol designers.

1.1. Terminology

Terminology adopted in this document includes terms such as "ring geometry", "tree structure", and "butterfly network", borrowed from P2P scientific literature. [RFC4981] provides an exhaustive definition of such terminology.

Certain security-related terms are to be understood in the sense defined in [RFC4949]; such terms include, but are not limited to, "attack", "authentication", "confidentiality", "encryption", "identity", and "integrity". Other security-related terms (for example, "denial of service") are to be understood in the sense defined in the referenced specifications.

2. Survey of Existing Literature

Gummadi et al. [Gummadi] compare popular DHT algorithms, and besides analyzing their resilience, provide an accurate evaluation of how well the logical overlay topology maps on the physical network layer. In their paper, relying only on measurements independently performed by overlay nodes without the support of additional location information provided by external entities, they demonstrate that the most efficient algorithms in terms of resilience and proximity performance are those based on the simplest geometric concept (i.e., the ring geometry, rather than tree structures, butterfly networks, and hybrid geometries).

Regardless of the geometrical properties of the distributed data structures involved, interactions between application-layer overlays and the underlying networks are a rich area of investigation. The available literature in this field can be divided into two categories (Figure 1): using application-level techniques to estimate topology, and using some kind of layer cooperation to estimate topology.

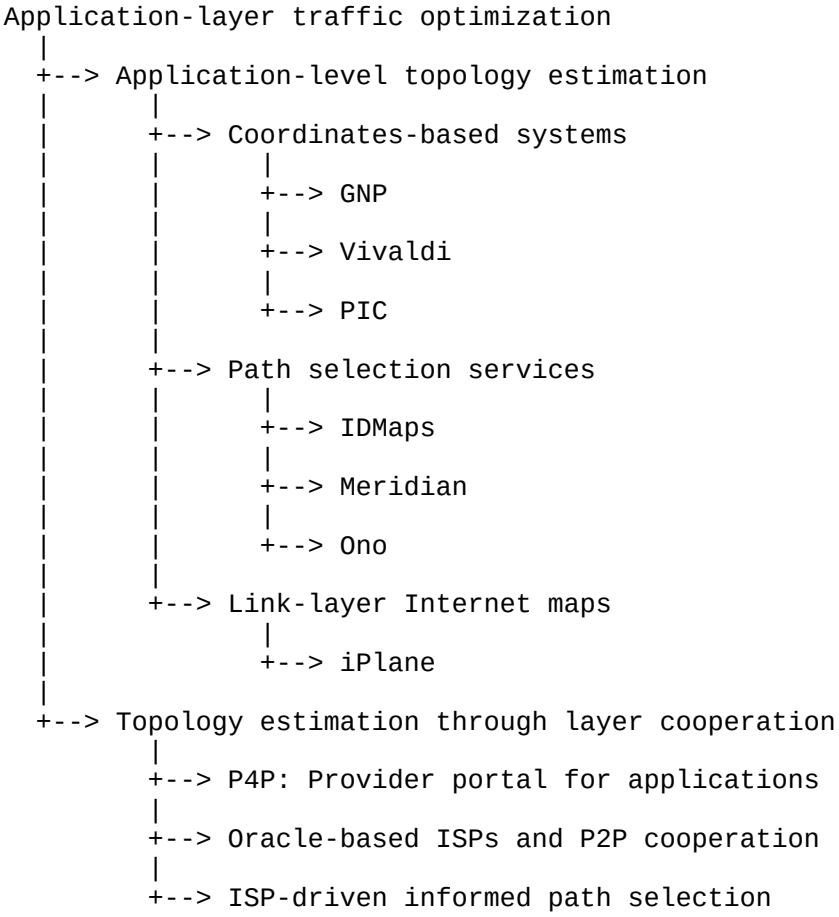

Figure 1: Taxonomy of Solutions for the Application-Layer Traffic Optimization Problem

2.1. Application-Level Topology Estimation

Estimating network topology information on the application layer has been an area of active research. Early systems used triangulation techniques to bound the distance between two hosts using a common landmark host. In such a technique, given a cost function C, a set of vertexes V and their corresponding edges, the triangle inequality holds if for any triple {a, b, c} in V, C(a, c) is always less than or equal to C(a, g) + C(b, c). The cost function C could be expressed in terms of desirable metrics such as bandwidth or latency.

We note that the techniques presented in this section are only representative of the sizable research in this area. Rather than

trying to enumerate an exhaustive list, we have chosen certain techniques because they represent an advance in the area that further led to derivative works.

Francis et al. proposed IDMaps [Francis], a system where one or more special hosts called tracers are deployed near an autonomous system. The distance measured in round-trip time (RTT) between hosts A and B is estimated as the cumulative distance between A and its nearest tracer Ta, plus the distance between B and its nearest tracer Tb, plus the shortest distance from Ta to Tb. To aid in scalability beyond that provided by the client-server design of IDMaps, Ng et al. proposed a P2P-based Global Network Positioning (GNP) architecture [Ng]. GNP was a network coordinate system based on absolute coordinates computed from modeling the Internet as a geometric space. It proposed a two-part architecture: in the first part, a small set of finite distributed hosts called landmarks compute their own coordinates in a fixed geometric space. In the second part, a host wishing to participate computes its own coordinates relative to those of the landmark hosts. Thus, armed with the computed coordinates, hosts can then determine interhost distance as soon as they discover each other.

Both IDMaps and GNP require fixed network infrastructure support in the form of tracers or landmark hosts; this often introduces a single point of failure and inhibits scalability. To combat this, new techniques were developed that embedded the network topology in a low-dimensional coordinate space to enable network distance estimation through vector analysis. Costa et al. introduced Practical Internet Coordinates (PIC) [Costa]. While PIC used the notion of landmark hosts, it did not require explicit network support to designate specific landmark hosts. Any node whose coordinates have been computed could act as a landmark host. When a node joined the system, it probed the network distance to some landmark hosts. Then, it obtained the coordinates of each landmark host and computed its own coordinates relative to each landmark host, subject to the constraint of minimizing the error in the predicted distance and computed distance.

Like PIC, Vivaldi [Dabek] proposed a fully distributed network coordinate system without any distinguished hosts. Whenever a node A communicates with another node B, it measures the RTT to that node and learns that node’s current coordinates. Node A subsequently adjusts its coordinates such that it is closer to, or further from, B by computing new coordinates that minimize the squared error. A Vivaldi node is thus constantly adjusting its position based on a simulation of interconnected mass springs. Vivaldi is now being used in the popular P2P application Vuze, and studies indicate that it scales well to very large networks [Ledlie].

Network coordinate systems require the embedding of the Internet topology into a coordinate system. This is not always possible without errors, which impacts the accuracy of distance estimations. In particular, it has proved to be difficult to embed the triangular inequalities found in Internet path distances [Ledlie]. Thus, Meridian [Wong] abandons the generality of network coordinate systems and provides specific distance evaluation services. In Meridian, each node keeps track of a small fixed number of neighbors and organizes them in concentric rings, ordered by distance from the node. Meridian locates the closest node by performing a multi-hop search where each hop exponentially reduces the distance to the target. Although less general than virtual coordinates, Meridian incurs significantly less error for closest node discovery.

The Ono project [Ono] takes a different approach and uses network measurements from Content Distribution Networks (CDNs) such as Akamai to find nearby peers. Used as a plugin to the Vuze bittorrent client, Ono provides 31% average download rate improvement [Su].

Comparison of application-level topology estimation techniques, as reported in literature. Results in terms of number of (D)imensions and (L)andmarks, 90th percentile relative error.

GNP vs. IDMaps(a) (7D, 15L)	PIC(b) vs. GNP (8D, 16L)	Vivaldi vs. GNP (2D, 32L)	Meridian vs. GNP (8D, 15L)
GNP: 0.50, IDMaps: 0.97	PIC: 0.38, GNP: 0.37	Vivaldi: 0.65, GNP: 0.65	Meridian: 0.78, GNP: 1.18

(a) Does not use dimensions or landmarks.
(b) Uses results from the hybrid strategy for PIC.

Table 1

Table 1 summarizes the application-level topology estimation techniques. The salient performance metric is the relative error. While all approaches define this metric a bit differently, it can be generalized as how close a predicted distance comes to the corresponding measured distance. A value of zero implies perfect prediction, and a value of 1 implies that the predicted distance is in error by a factor of two. PIC, Vivaldi, and Meridian compare their results with that of GNP, while GNP itself compares its results with a precursor technique, IDMaps. Because each of the techniques uses a different Internet topology and a varying number of landmarks and dimensions to interpret the data set, it is impossible to

normalize the relative error across all techniques uniformly. Thus, we present the relative error data in pairs, as reported in the literature describing the specific technique. Readers are urged to compare the relative error performance in each column on its own and not draw any conclusions by comparing the data across columns.

Most of the work on estimating topology information focuses on predicting network distance in terms of latency and does not provide estimates for other metrics such as throughput or packet loss rate. However, for many P2P applications latency is not the most important performance metric, and these applications could benefit from a richer information plane. Sophisticated methods of active network probing and passive traffic monitoring are generally very powerful and can generate network statistics indirectly related to performance measures of interest, such as delay and loss rate on link-level granularity. Extraction of these hidden attributes can be achieved by applying statistical inference techniques developed in the field of inferential network monitoring or network tomography subsequent to sampling of the network state. Thus, network tomography enables the extraction of a richer set of topology information, but at the same time inherently increases complexity of a potential information plane and introduces estimation errors. For both active and passive methods, statistical models for the measurement process need to be developed, and the spatial and temporal dependence of the measurements should be assessed. Moreover, measurement methodology and statistical inference strategy must be considered jointly. For a deeper discussion of network tomography and recent developments in the field, we refer the reader to [Coates].

One system providing such a service is iPlane [Madhyastha], which aims at creating an annotated atlas of the Internet that contains information about latency, bandwidth, capacity, and loss rate. To determine features of the Internet topology, iPlane bridges and builds upon different ideas, such as active probing based on packet dispersion techniques to infer available bandwidth along path segments. These ideas are drawn from different fields, including network measurement as described by Dovrolis et al. in [Dovrolis] and network tomography [Coates].

2.2. Topology Estimation through Layer Cooperation

Instead of estimating topology information on the application level through distributed measurements, this information could be provided by the entities running the physical networks -- usually ISPs or network operators. In fact, they have full knowledge of the topology of the networks they administer and, in order to avoid congestion on critical links, are interested in helping applications to optimize the traffic they generate. The remainder of this section briefly

describes three recently proposed solutions that follow such an approach to address the ALTO problem.

2.2.1. P4P Architecture

The architecture proposed by Xie et al. [Xie] has been adopted by the Distributed Computing Industry Association (DCIA) P4P working group [P4P], an open group established by ISPs, P2P software distributors, and technology researchers, with the dual goal of defining mechanisms to (1) accelerate content distribution and (2) optimize utilization of network resources.

The main role in the P4P architecture is played by servers called "iTrackers", deployed by network providers and accessed by P2P applications (or, in general, by elements of the P2P system) in order to make optimal decisions when selecting a peer to which the element will connect. An iTracker may offer three interfaces:

1. Info: Allows P2P elements (e.g., peers or trackers) to get opaque information associated to an IP address. Such information is kept opaque to hide the actual network topology, but can be used to compute the network distance between IP addresses.

2. Policy: Allows P2P elements to obtain policies and guidelines of the network, which specify how a network provider would like its networks to be utilized at a high level, regardless of P2P applications.

3. Capability: Allows P2P elements to request network providers' capabilities.

The P4P architecture is under evaluation with simulations, experiments on the PlanetLab distributed testbed, and in field tests with real users. Initial simulations and PlanetLab experiment results [P4P] indicate that improvements in BitTorrent download completion time and link utilization in the range of 50-70% are possible. Results observed on Comcast's network during a field test trial conducted with a modified version of the software used by the Pando content delivery network (documented in RFC 5632 [RFC5632]) show average improvements in download rate in different scenarios varying between 57% and 85%, and a 34% to 80% drop in the cross-domain traffic generated by such an application.

2.2.2. Oracle-Based ISP-P2P Collaboration

In the general solution proposed by Aggarwal et al. [Aggarwal], network providers offer host servers, called "oracles", that help P2P users choose optimal neighbors.

The oracle concept uses the following mechanism: a P2P client sends the list of potential peers to the oracle hosted by its ISP and receives a re-arranged peer list, ordered according to the ISP's local routing policies and preferences. For instance, to keep the traffic local, the ISP may prefer peers within its network, or it may pick links with higher bandwidth or peers that are geographically closer to improve application performance. Once the client has obtained this ordered list, it has enough information to perform better-than-random initial peer selection.

Such a solution has been evaluated with simulations and experiments run on the PlanetLab testbed, and the results show both improvements in content download time and a reduction of overall P2P traffic, even when only a subset of the applications actually query the oracle to make their decisions.

2.2.3. ISP-Driven Informed Path Selection (IDIPS) Service

The solution proposed by Saucez et al. [Saucez] is essentially a modified version of the oracle-based approach described in Section 2.2.2, intended to provide a network-layer service for finding the best source and destination addresses when establishing a connection between two endpoints in multi-homed environments (which are common in IPv6 networking). Peer selection optimization in P2P systems -- the ALTO problem in today's Internet -- can be addressed by the IDIPS solution as a specific sub-case where the options for the destination address consist of all the peers sharing a desired resource, while the choice of the source address is fixed. An evaluation performed on IDIPS shows that costs for both providing and accessing the service are negligible.

3. Application-Level Topology Estimation and the ALTO Problem

The application-level techniques described in Section 2.1 provide tools for peer-to-peer applications to estimate parameters of the underlying network topology. Although these techniques can improve application performance, there are limitations of what can be achieved by operating only on the application level.

Topology estimation techniques use abstractions of the network topology, which often hide features that would be of interest to the application. Network coordinate systems, for example, are unable to detect overlay paths shorter than the direct path in the Internet topology. However, these paths frequently exist in the Internet [Wang]. Similarly, application-level techniques may not accurately estimate topologies with multipath routing.

When using network coordinates to estimate topology information, the underlying assumption is that distance in terms of latency determines performance. However, for file sharing and content distribution applications, there is more to performance than just the network latency between nodes. The utility of a long-lived data transfer is determined by the throughput of the underlying TCP protocol, which depends on the round-trip time as well as the loss rate experienced on the corresponding path [Padhye]. Hence, these applications benefit from a richer set of topology information that goes beyond latency, including loss rate, capacity, and available bandwidth.

Some of the topology estimation techniques used by P2P applications need time to converge to a result. For example, current BitTorrent clients implement local, passive traffic measurements and a tit-for-tat bandwidth reciprocity mechanism to optimize peer selection at a local level. Peers eventually settle on a set of neighbors that maximizes their download rate, but because peers cannot reason about the value of neighbors without actively exchanging data with them, and because the number of concurrent data transfers is limited (typically to 5-7), convergence is delayed and easily can be sub-optimal.

Skype’s P2P Voice over IP (VoIP) application chooses a relay node in cases where two peers are behind NATs and cannot connect directly. Measurements taken by Ren et al. [Ren] showed that the relay selection mechanism of Skype (1) is not able to discover the best possible relay nodes in terms of minimum RTT, (2) requires a long setup and stabilization time, which degrades the end user experience, and (3) is creating a non-negligible amount of overhead traffic due to probing a large number of nodes. They further showed that the quality of the relay paths could be improved when the underlying network Autonomous System (AS) topology is considered.

Some features of the network topology are hard to infer through application-level techniques, and it may not be possible to infer them at all, e.g., service-provider policies and preferences such as the state and cost associated with interdomain peering and transit links. Another example is the traffic engineering policy of a service provider, which may counteract the routing objective of the overlay network, leading to a poor overall performance [Seetharaman].

Finally, application-level techniques often require applications to perform measurements on the topology. These measurements create traffic overhead, in particular, if measurements are performed individually by all applications interested in estimating topology.

4. Open Issues

Beyond a significant amount of research work on the topic, we believe that there are sizable open issues to address in an infrastructure-based approach to traffic optimization. The following is not an exhaustive list, but a representative sample of the pertinent issues.

4.1. Coordinate Estimation or Path Latencies?

Despite the many solutions that have been proposed for providing applications with topology information in a fully distributed manner, there is currently an ongoing debate in the research community whether such solutions should focus on estimating nodes' coordinates or path latencies. Such a debate has recently been fed by studies showing that the triangle inequality on which coordinate systems are based is often proved false in the Internet [Ledlie]. Proposed systems following both approaches -- in particular, Vivaldi [Dabek] and PIC [Costa] following the former, and Meridian [Wong] and iPlane [Madhyastha] the latter -- have been simulated, implemented, and studied in real-world trials, each one showing different points of strength and weaknesses. Concentrated work will be needed to determine which of the two solutions will be conducive to the ALTO problem.

4.2. Malicious Nodes

Another open issue common in most distributed environments consisting of a large number of peers is the resistance against malicious nodes. Security mechanisms to identify misbehavior are based on triangle inequality checks [Costa], which, however, tend to fail and thus return false positives in the presence of measurement inaccuracies induced, for example, by traffic fluctuations that occur quite often in large networks [Ledlie]. Beyond the issue of using triangle inequality checks, authoritatively authenticating the identity of an oracle, and preventing an oracle from attacks are also important. Existing techniques -- such as Public Key Infrastructure (PKI) [RFC5280] or identity-based encryption [Boneh] for authenticating the identity and the use of secure multi-party computation techniques to prevent an oracle from collusion attacks -- need to be explored and studied for judicious use in ALTO-type solutions.

4.3. Information Integrity

Similarly, even in controlled architectures deployed by network operators where system elements may be authenticated [Xie], [Aggarwal],[Saucez], it is still possible that the information returned to applications is deliberately altered, for example, assigning higher priority to financially inexpensive links instead of

neutrally applying proximity criteria. What are the effects of such deliberate alterations if multiple peers collude to determine a different route to the target, one that is not provided by an oracle? Similarly, what are the consequences if an oracle targets a particular node in another AS by redirecting an inordinate number of querying peers to it causing, essentially, a Distributed Denial-of-Service (DDoS) [RFC4732] attack on the node? Furthermore, does an oracle broadcast or multicast a response to a query? If so, techniques to protect the confidentiality of the multicast stream will need to be investigated to thwart "free riding" peers.

4.4. Richness of Topological Information

Many systems already use RTT to account for delay when establishing connections with peers (e.g., Content-Addressable Network (CAN) [Ratnasamy], Bamboo [Rhea]). An operator can provide not only the delay metric but other metrics that the peer cannot figure out on its own. These metrics may include the characteristics of the access links to other peers, bandwidth available to peers (based on operators' engineering of the network), network policies, preferences such as state and cost associated with intradomain peering links, and so on. Exactly what kinds of metrics an operator can provide to stabilize the network throughput will also need to be investigated.

4.5. Hybrid Solutions

It is conceivable that P2P users may not be comfortable with operator intervention to provide topology information. To eliminate this intervention, alternative schemes to estimate topological distance can be used. For instance, Ono uses client redirections generated by Akamai CDN servers as an approximation for estimating distance to peers; Vivaldi, GNP, and PIC use synthetic coordinate systems. A neutral third party can make available a hybrid layer-cooperation service -- without the active participation of the ISP -- that uses alternative techniques discussed in Section 2.1 to create a topological map. This map can be subsequently used by a subset of users who may not trust the ISP.

4.6. Negative Impact of Over-Localization

The literature presented in Section 2 shows that a certain level of locality-awareness in the peer selection process of P2P algorithms is usually beneficial to application performance. However, an excessive localization of the traffic might cause partitioning in the overlay interconnecting these peers, which will negatively affect the performance experienced by the peers themselves.

Finding the right balance between localization and randomness in peer selection is an open issue. At the time of writing, it seems that different applications have different levels of tolerance and should be addressed separately. Le Blond et al. [LeBlond] have studied the specific case of BitTorrent, proposing a simple mechanism to prevent partitioning in the overlay, yet reach a high level of cross-domain traffic reduction without adversely impacting peers.

5. Security Considerations

This document is a survey of existing literature on topology estimation. As such, it does not introduce any new security considerations to be taken into account beyond what is already discussed in each paper surveyed.

Insofar as topology estimation is used to provide a solution to the ALTO problem, the issues in Sections 4.2 and 4.3 deserve special attention. There are efforts underway in the IETF ALTO working group to design a protocol that protects the privacy of the peer-to-peer users as well as the service providers. [Chen] provides an overview of ALTO security issues, Section 11 of [Alimi] is an exhaustive overview of ALTO security, and Section 6 of RFC 5693 [RFC5693] also lists the privacy and confidentiality aspects of an ALTO solution.

The following references provide a starting point for general peer-to-peer security issues: [Wallach], [Sit], [Douceur], [Castro], and [Friedman].

6. Acknowledgments

This document is a derivative work of a position paper submitted at the IETF RAI area/MIT workshop held on May 28th, 2008 on the topic of Peer-to-Peer Infrastructure (P2Pi) [RFC5594]. The article on a similar topic, also written by the authors of this document and published in IEEE Communications [Gurbani], was also partially derived from the same position paper. The authors thank profusely Arnaud Legout, Richard Yang, Richard Woundy, Stefano Previdi, and the many people that have participated in discussions and provided insightful feedback at any stage of this work.

7. Informative References

[Aggarwal] Aggarwal, V., Feldmann, A., and C. Scheideler, "Can ISPs and P2P users cooperate for improved performance?", in ACM SIGCOMM Computer Communications Review, vol. 37, no. 3.

[Alimi] Alimi, R., Ed., Penno, R., Ed., and Y. Yang, Ed., "ALTO Protocol", Work in Progress, July 2010.

[Boneh] Boneh, D. and M. Franklin, "Identity-Based Encryption from the Weil Pairing", in Proceedings of the 21st Annual International Cryptology Conference on Advances in Cryptology, August 2001.

[Castro] Castro, M., Druschelw, P., Ganesh, A., Rowstron, A., and D. Wallach, "Security for Structured Peer-to-peer Overlay Networks", in Proceedings of Symposium on Operating Systems Design and Implementation (OSDI'02), December 2002.

[Chen] Chen, S., Gao, F., Beijing, X., and M. Xiong, "Overview for ALTO Security Issues", Work in Progress, February 2010.

[Coates] Coates, M., Hero, A., Nowak, R., and B. Yu, "Internet Tomography", in IEEE Signal Processing Magazine, vol. 19, no. 3.

[Costa] Costa, M., Castro, M., Rowstron, A., and P. Key, "PIC: Practical Internet coordinates for distance estimation", in Proceedings of International Conference on Distributed Systems 2003.

[Dabek] Dabek, F., Cox, R., Kaashoek, F., and R. Morris, "Vivaldi: A Decentralized Network Coordinate System", in ACM SIGCOMM: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications, vol. 34, no. 4.

[Douceur] Douceur, J., "The Sybil Attack", in Proceedings of the First International Workshop on Peer-to-Peer Systems, March 2002.

[Dovrolis] Dovrolis, C., Ramanathan, P., and D. Moore, "What do packet dispersion techniques measure?", in Proceedings of IEEE INFOCOM 2001.

[Francis] Francis, P., Jamin, S., Jin, C., Jin, Y., Raz, D., Shavitt, Y., and L. Zhang, "IDMaps: A global Internet host distance estimation service", in Proceedings of IEEE INFOCOM 2001.

[Friedman] Friedman, A. and A. Camp, "Peer-to-Peer Security", in The Handbook of Information Security, J. Wiley & Sons, 2005.

[Glasner] Glasner, J., "P2P fuels global bandwidth binge", available from http://www.wired.com/.

[Gummadi] Gummadi, K., Gummadi, R., Gribble, S., Ratnasamy, S., Shenker, S., and I. Stoica, "The impact of DHT routing geometry on resilience and proximity", in ACM SIGCOMM: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications.

[Gurbani] Gurbani, V., Hilt, V., Rimac, I., Tomsu, M., and E. Marocco, "A Survey of Research on the Application-Layer Traffic Optimization Problem and the Need for Layer Cooperation", in IEEE Communications, vol. 47, no. 8.

[Karagiannis] Karagiannis, T., Broido, A., Brownlee, N., Claffy, K., and M. Faloutsos, "Is P2P dying or just hiding?", in Proceedings of IEEE GLOBECOM 2004 Conference.

[LeBlond] Le Blond, S., Legout, A., and W. Dabbous, "Pushing BitTorrent Locality to the Limit", available at http://hal.inria.fr/.

[Ledlie] Ledlie, J., Gardner, P., and M. Seltzer, "Network Coordinates in the Wild", in USENIX: Proceedings of NSDI 2007.

[LightReading] LightReading, "Controlling P2P traffic", available from http://www.lightreading.com/.

[LinuxReviews] linuxReviews.org, "Peer to peer network traffic may account for up to 85% of Internet’s bandwidth usage", available from http://linuxreviews.org/.

[Madhyastha] Madhyastha, H., Isdal, T., Piatek, M., Dixon, C., Anderson, T., Krishnamurthy, A., and A. Venkataramani, "iPlane: an information plane for distributed services", in USENIX: Proceedings of the 7th symposium on Operating systems design and implementation.

[Ng] Ng, T. and H. Zhang, "Predicting internet network distance with coordinates-based approaches", in Proceedings of INFOCOM 2002.

[Ono] "Northwestern University Ono Project", <http://www.aqualab.cs.northwestern.edu/projects/Ono.html>.

[P4P] "DCIA P4P Working group", <http://www.dcia.info/activities/#P4P>.

[Padhye] Padhye, J., Firoiu, V., Towsley, D., and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation", in Technical Report UM-CS-1998-008, University of Massachusetts 1998.

[Parker] Parker, A., "The true picture of peer-to-peer filesharing", available from http://www.cachelogic.com/.

[RFC4732] Handley, M., Ed., Rescorla, E., Ed., and IAB, "Internet Denial-of-Service Considerations", RFC 4732, December 2006.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", FYI 36, RFC 4949, August 2007.

[RFC4981] Risson, J. and T. Moors, "Survey of Research towards Robust Peer-to-Peer Networks: Search Methods", RFC 4981, September 2007.

[RFC5280] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., and W. Polk, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, May 2008.

[RFC5594] Peterson, J. and A. Cooper, "Report from the IETF Workshop on Peer-to-Peer (P2P) Infrastructure, May 28, 2008", RFC 5594, July 2009.

[RFC5632] Griffiths, C., Livingood, J., Popkin, L., Woundy, R., and Y. Yang, "Comcast’s ISP Experiences in a Proactive Network Provider Participation for P2P (P4P) Technical Trial", RFC 5632, September 2009.

[RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, October 2009.

[Ratnasamy] Ratnasamy, S., Francis, P., Handley, M., Karp, R., and S. Shenker, "A Scalable Content-Addressable Network", in ACM SIGCOMM: Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications, January 2001.

[Ren] Ren, S., Guo, L., and X. Zhang, "ASAP: An AS-aware peer-relay protocol for high quality VoIP", in Proceedings of IEEE ICDCS 2006.

[Rhea] Rhea, S., Godfrey, B., Karp, B., Kubiatowicz, J., Ratnasamy, S., Shenker, S., Stoica, I., and H. Yu, "OpenDHT: a public DHT service and its uses", in ACM SIGCOMM: Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications, August 2005.

[Saucez] Saucez, D., Donnet, B., and O. Bonaventure, "Implementation and Preliminary Evaluation of an ISP-Driven Informed Path Selection", in Proceedings of ACM CoNEXT 2007.

[Seetharaman] Seetharaman, S., Hilt, V., Hofmann, M., and M. Ammar, "Preemptive Strategies to Improve Routing Performance of Native and Overlay Layers", in Proceedings of IEEE INFOCOM 2007.

[Sit] Sit, E. and R. Morris, "Security Considerations for Peer-to-Peer Distributed Hash Tables, Revised Papers from the First", in Proceedings of the First International Workshop on Peer-to-Peer Systems, March 2002.

[Su] Su, A., Choffnes, D., Kuzmanovic, A., and F. Bustamante, "Drafting behind Akamai (travelocity-based detouring)", in ACM SIGCOMM: Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications.

[Vuze] "Vuze bittorrent client", <http://www.vuze.com/>.

[Wallach] Wallach, D., "A survey of peer-to-peer security issues", in Proceedings of International Symposium on Software Security, 2002.

[Wang] Wang, G., Zhang, B., and T. Ng, "Towards Network Triangle Inequality Violation Aware Distributed Systems", in ACM SIGCOMM: Proceedings of the 7th conference on Internet measurement.

[Wong] Wong, B., Slivkins, A., and E. Sirer, "Meridian: A lightweight network location service without virtual coordinates", in ACM SIGCOMM: Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications.

[Xie] Xie, H., Krishnamurthy, A., Silberschatz, A., and Y. Yang, "P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers", in ACM SIGCOMM Computer Communication Review, vol. 38, no. 4.

Authors' Addresses

Ivica Rimac
Bell Labs, Alcatel-Lucent
EMail: rimac@bell-labs.com

Volker Hilt
Bell Labs, Alcatel-Lucent
EMail: volkerh@bell-labs.com

Marco Tomsu
Bell Labs, Alcatel-Lucent
EMail: marco.tomsu@alcatel-lucent.com

Vijay K. Gurbani
Bell Labs, Alcatel-Lucent
EMail: vkg@bell-labs.com

Enrico Marocco
Telecom Italia
EMail: enrico.marocco@telecomitalia.it