Improving User and ISP Experience through ISP-aided P2P Locality

Vinay Aggarwal
Deutsche Telekom Labs/TU Berlin

Obi Akonjang
TU Berlin

Anja Feldmann
Deutsche Telekom Labs/TU Berlin

{vinay,obi,anja}@net.t-labs.tu-berlin.de

***Abstract*—Despite recent improvements, P2P systems are still plagued by fundamental issues such as overlay/underlay topological and routing mismatch, which affects their performance and causes traffic strains on the ISPs. In this work, we aim to improve overall system performance for ISPs as well as P2P systems by means of traffic localization through improved collaboration between ISPs and P2P systems. More specifically, we study the effects of different ISP/P2P topologies as well as a broad range of influential user behavior characteristics, namely content availability, churn, and query patterns, on end-user and ISP experience. We show that ISP-aided P2P locality benefits both P2P users and ISPs, measured in terms of improved content download times, increased network locality of query responses and desired content, and overall reduction in P2P traffic.**

I. Introduction

P2P systems are self-organizing systems of autonomous entities, called peers, that cooperate for common goals. These common goals range from sharing of resources, e.g., music and video files, processing power, or storage space [1] to collaborative routing as in Skype and P2P-TV. A fundamental characteristic of these systems is their distributed topologies and resources, which is in stark contrast to the centralized resources of client-server systems. Measurement studies consistently indicate that $50-70\%$ of Internet traffic is caused by popular P2P applications [1][2].

Despite their many advantages, P2P systems are plagued by some fundamental issues, such as overlay/underlay topological and routing mismatch [3], inefficiencies in locating and retrieving resources, and scalability and performance issues caused by uncontrolled traffic swamps [1].

A. Overlay-Underlay Correlation

Several of these drawbacks can be addressed by collaboration between the P2P overlay and the Internet routing underlay. In previous work [4], we proposed that each ISP offers an "oracle" service to the P2P users which explicitly helps P2P users to choose "good" neighbors. The P2P user can supply its ISP's oracle with a list of possible P2P neighbors, during bootstrapping and/or content exchange. The ISP's oracle then returns a ranked list to the querying user, according to its preference (e.g., nodes within the AS, AS-hop distance) and knowledge of the ISP topology and traffic volume, while at the same time keeping the interest of the P2P user in mind. We showed that in principle, P2P systems as well as the ISPs profit from the use of the oracle even when only considering the AS-distance for ranking nodes [4], because the overlay topology is now localized and respects the underlying Internet topology, and the P2P user profits from the ISP's knowledge. Moreover, using an oracle explicitly avoids caching content, contrary to [5], thus absolving ISPs of potential legal issues. While there are alternate proposals to localize P2P traffic, e.g., [6][7][8], the oracle proposal [4] is simple, scalable, applicable to all overlays, and promotes active collaboration between ISPs and P2P systems.

B. Our Contributions

In this paper, we build on our previous work by

- extending the ISP's oracle to also consider last-hop bandwidth of P2P users while ranking possible neighbors
- studying the impact of different ISP/P2P topologies and user behavior patterns on end-user performance. This task comprises three stages: (i) design of different ISP and P2P topologies, (ii) design of different user behavioral patterns, namely, content availability, churn, and query patterns, (iii) extensive experimental studies to determine the impact of different topologies and behavioral patterns on end-user experience, a task unaddressed as yet to the best of our knowledge.

Our findings show that in contrast to the unmodified P2P system, the ISP-aided localized P2P system shows consistent improvements in the observed end-user experience, measured in terms of content download times, network locality of query responses and desired content, and number of query responses. A significantly large portion of P2P traffic remains local to the ISP network, and ISPs notice a substantial reduction in the overall P2P traffic. This can lead to immense cost savings for the ISPs [2]. The oracle consistently shows performance gains even across different topologies under a broad range of user behavior scenarios.

II. Using Bandwidth to Select Neighbors

We propose to extend the ISP-hosted oracle [4] to go beyond using only network proximity (nodes within the AS, AS-hop distance) to keep traffic within its network. It should also use last-hop bandwidth of P2P users within its network to help querying P2P users select high-performance neighbors. This is possible as the ISP knows its customers' last-hop bandwidth and hence does not have to measure it, yet this metric is difficult and traffic-intensive to reverse engineer accurately [9][10] for the P2P users.

Moreover, this has advantages over neighbor selection using latency measurements [8] as network latency can change quickly [11]. Also, latency is difficult to predict reliably [12][13], especially in the face of newer breed of Internet applications characterized by large data content and high churn. While we agree that similar arguments hold to some degree regarding estimating available last-hop bandwidth [10] as well, we argue that utilizing the ISP knowledge via the oracle helps to (i) improve accuracy (ii) mitigate ISP's concerns about traffic management and respect for routing policies (iii) reduce the excessive traffic swarm [14] that results from frequent pinging of the network to deduce latency and/or available bandwidth. Besides, latency between Internet hosts is dominated by the cable/DSL bandwidths at the last-mile connections [15], thus making neighbor selection based on last-hop bandwidths a good option.

We show that keeping the P2P traffic localized allows users to benefit from the significant geographic and interest-based clustering [16] for audio/video P2P content. One may argue in favour of bypassing the ISP's oracle service to utilize geolocalization techniques [17] to choose neighbors. However, we caution that even the best such techniques can identify a node to within 22 miles of its actual position [13], hence making differentiation of nodes even within the same city difficult. On the other hand, the ISP being aware of the minute details of its PoP-level backbone topology, can easily use this information to better rank the querying node's neighbors, even within the same city. Indeed, oracles from multiple ISPs can collaborate to build a global coordinate system [18]. We thus believe that ISP-aided P2P neighbor selection is a win-win solution for ISPs as well as P2P systems.

III. Methodology

To be able to study user behavior and its impact on P2P system performance, we require a packet level simulator, even though this limits the scale of our experiments to roughly 700 P2P nodes. We use the Java-based, discrete-event network simulator SSFNet [19][20], which provides multi-protocol, multi-domain Internet modeling at and above the IP packet level of detail. Network entities like IP, TCP, UDP, BGP, OSPF, HTTP, hosts, routers, links and LANs are supported within it. At the application layer, we code a P2P protocol [21] similar to Gnutella, an unstructured file-sharing system relying on flooding connectivity pings and search queries to locate content [22]. Each node floods query search messages to all its connected peers, which iteratively forward it to their connected peers, until the query reaches a peer that possesses the searched content. Such a peer then sends a query response to the originating peer, retracing the overlay path traversed by the corresponding query. Each message carries a TTL (time to live) and message ID tag. To improve scalability, nodes are classified in a two-level hierarchy, with high-performance ultrapeer nodes maintaining the overlay structure by connecting with each other and forwarding only the relevant messages to a small number of shielded leaf nodes. File exchange is carried out between the peers directly using HTTP, similar to most other P2P file sharing systems. Recent developments like efficient querying, Gnutella2, and support for mobile users have helped keep the number of Gnutella users around $1-2$ million [23].

A. Topology models

In order to study the effects of ISP topologies as well as the distribution of P2P customers per ISP on P2P locality, we design 5 different AS topologies: Germany, USA, World1, World2 and World3. Each topology consists of 700 P2P nodes distributed within various ASes (recall the memory limitations of packet-level simulators [20][6]). We take a subset of the AS topology of Germany as published in [24], and distribute the P2P nodes in each of the 12 ISPs according to the actual number of DSL customers of these major ISPs [25]. For USA, we model several regional providers at most of the major cities, and connect them with peering links [20][26], distributing the P2P nodes in the 25 ASes according to the ratio of the population of the cities. To model the World topologies, we design inter-AS connections as derived from BGP routing information in [27], and distribute P2P nodes based on results in [27][28]. Each World topology has 1 level-1 AS, 5 level-2 ASes, and 10 level-3 ASes, hence resulting in a 16-AS network. The number of P2P nodes assigned to (level-1,level-2,level-3) ASes are as follows: World1: $10, 46, 46$; World2: $355, 23, 23$; World3: $50, 46, 42$. We thus have 3 different topologies, and for the World topology, 3 different ways of distributing P2P nodes within the ASes.

Bearing in mind the memory limitations and that it is fundamentally difficult to simulate the Internet [29], we model the topologies within SSFNet as follows. Each AS has 2 routers, one for intra-AS node connections, and one for the inter-AS connections between ASes. Within each AS, all the nodes are connected in a star topology to the intra-AS router. The nodes have network interfaces representing typical last-hop DSL and cable modem bandwidths, ranging from 1 Mbps to 16 Mbps, and top-tier ASes have a larger proportion of higher bandwidth customers than the lower-tier ASes [9][20][25][28]. The links between level-1 and level-2 ASes have a delay of $4-6$ ms, while links between level-2 and level-3 ASes have a delay of $14-18$ ms [20][30].

B. User Behavior models

While we have implemented a specific protocol in SSFNet, our goal is to perform experiments that represent a large section of P2P systems in use today. Studies [31][32] have shown that user behavior is largely invariant across P2P systems, both structured and unstructured. This means that factors like session lengths, content availability (free-riding), query patterns and search strings are similar across different P2P systems.

We note that user behavioral patterns are in constant transition, although the broad characteristics across different systems are comparable. Hence, we use different distributions to simulate the behavioral patterns, some very close to observed behavior, e.g., Weibull distributions, some that serve as a comparison standard, and some that reflect worst-case or utopian scenarios, e.g., exponential or uniform distributions. We derive the parameters for each P2P user characteristic via careful *sensitivity analysis*, by exploring multiple parameters for each distribution, until we achieve a representation that reflects observed user behavior within the limitations of a simulation environment.

Content availability: The presence of a large number of free-riders has been confirmed by extensive measurement studies [9][16][23][33]. The distribution of the number of files shared by each peer appears to be heavy-tailed, though there is no agreement on the exact parameters. Hence, we take different models to represent file distribution. While Weibull (scale=42, shape=0.5) and Pareto (k=100, alpha=10) cases represent realistic behavior (i.e., large number of free-riders), the Uniform case (min=0, max=100) is used as a comparison base, and the Poisson case (mean=50) represents a scenario where every peer shares a constant number of files.

Session lengths: Churn in P2P systems has attracted much attention from researchers [31][32][34]. Again, while most studies agree that online session length is a heavy-tailed distribution, different P2P systems have been shown to fit different distributions (or different parameters of the same distribution) at different times of measurement [32]. Hence, we represent online session lengths using different distributions where Pareto (k=600, alpha=0.5) and Weibull (scale=600, shape=0.2) cases represent realistic behavior, Uniform case (min=1, max=600) is used as a comparison base, and Poisson case (mean=300) represents the scenario where almost every peer has a constant online duration.

Query strings: Most P2P systems are characterized by query search phrases of two kinds [35]: constant phrases that aim to find content of a particular type, e.g., mp3, rap, dvd; and volatile phrases that search for a specific content, e.g. artist or album name. Query popularity distributions and load across time and region are reported in [33][35]. We reflect this by using 45% constant phrases and 45% volatile phrases for query strings. The rest 10% query strings are chosen such that they do not match any content in the network. Besides, 20% of all queries match only 1 or 2 content files. This enables us to analyze the effect of P2P locality on content search.

IV. Results

In this paper, we go beyond our previous work, by using the following metrics to judge end-user as well as ISP experience: number of responses that each Query generates, the AS distance and overlay hop count of Query-responses, time taken to download a single file, amount of exchanged content that remains within ISP network boundaries, and total reduction in P2P negotiation traffic. We perform two sets of experiments: (i) to study the effects of *various topologies* on the above metrics with realistic user behavior, comparing oracle-aided P2P with unmodified P2P, (ii) to measure the effects of *various user behavior patterns* on the above metrics for oracle-aided P2P. All the results are based on experiments with $10,000$ successful queries that result in $10,000$ file transfers. Each file is of 512KB size (the typical file piece size used in popular P2P systems) and is exchanged directly between the peers using HTTP. The oracle sorts the candidate list of neighbors based on the following algorithm: (i) identify the nodes within its AS, and sort them based on last-hop bandwidth, (ii) for nodes not within its AS, sort them using AS-hop distance.

A. Variation in Topology

For each topology model, we run two experiments, one with unmodified(U) P2P, another with oracle-aided and therefore a biased(B) P2P. In the unmodified case, P2P nodes go online, connect to random neighbors, search for content and exchange files, without consulting the ISP's oracle at any stage. In the biased case, P2P nodes consult the oracle while bootstrapping, as well as when downloading files. The bootstrapping phase is used to connect to proximal neighbors, hence setting up a localized P2P topology that is correlated with the Internet AS topology. Nodes search for a specific content by flooding queries. On finding it at a set of nodes, they again consult the oracle to choose the best node for downloading. We model content availability and online session lengths by Weibull distributions (realistic behavior). The results for all 5 topologies are shown in Figure 1.

Content exchange: The most important metric for the end-user is the time taken to download content. As shown in Figure 1(a), the download time per 512KB file decreases by $1-3$ seconds (a reduction of $16-34\%$) for all 5 topologies, when P2P users consult the ISP-hosted oracle to choose proximal neighbors. We also notice that changing the inter-AS delays does not have a significant effect on file download times. Moreover, additional simulations confirm that exchanging content with a high-bandwidth peer in another AS is consistently faster than a low-bandwidth peer in the same AS. This confirms that file download times are dominated by last-hop bandwidths [15].

From the ISP point of view, the amount of file content that remains within the ISP network boundaries more than doubles for the biased P2P case, see Figure 1(b). This can result in direct cost savings for the ISP, estimated to be in the order of $1 billion world-wide [2]. We note that the improvements for the biased P2P system are more pronounced in World1 and World3, as the peers are more evenly distributed across the ASes in these topologies.

We conclude that while the ISP benefits from AS-distance based neighbor selection, the benefits to P2P users accrue mainly from last-hop bandwidth based selection, thus underscoring the need for both metrics in the oracle.

Query Search: Figure 1(c) shows that there is no adverse effect on the query search phase of P2P systems when nodes actively consult the oracle. We actually notice an increase in the number of query responses per query for the biased P2P case, which is due to a more efficient swarming of the queries (and their responses) within the localized P2P topology. A closer examination reveals that for the same number of unique queries, the negotiation traffic in the overlay, which is emanating from flooding and forwarding of queries and their responses, decreases by $17-40\%$ in the biased P2P topologies. Despite this welcome reduction in P2P traffic, there is no adverse effect, as the number of responses per query actually increases. This implies that a significantly smaller number of duplicate messages is carried in the overlay, thus improving the scalability of P2P systems and reducing the traffic in the ISP network.

The number of queries that fail to find any content remains the same for the unmodified as well as the biased P2P system. This means that even for the case of queries which match only 1 or 2 content files located somewhere in the network, the efficient swarming of queries in the localized topology ensures that queries find such content. Besides, the query responses more often come from peers that are located within the same AS as the originating query, see Figure 1(d). This naturally leads to a decrease in the average AS distance of query responses per query for the biased P2P case.

P2P topology: An investigation of the graph topological properties of biased overlay graphs reveals that localized P2P graphs maintain the nice graph properties which are typical of random overlays, namely, small node degree, small graph diameter, small mean path length and connectedness, even under heavy node churn. The average node degree, shown in Figure 1(e), changes only slightly, from 18 for unmodified P2P to 16 for biased P2P. The graph diameter is found to remain constant at $6-7$ hops, and the mean overlay path length between all pairs of overlay nodes increases only nominally from 2.5 hops for unmodified P2P to 3.3 hops for biased P2P, see Figure 1(f). In other words, the graph structural properties of the overlay are not affected adversely when consulting the oracle even under churn. Importantly, despite heavy node churn, the overlay graph remains connected. Even if a subgraph gets temporarily disconnected, P2P nodes quickly re-establish peerings and form a connected topology.

B. Variation in User Behavior

Now that the benefits of ISP-aided P2P locality have been established across various topology models, we analyse the effects of user behavior on the above metrics. This helps to reveal the effect of aggresive node churn on graph connectivity and query responses. We also study scenarios when a small number of nodes serve most of the files in the P2P network and go offline, to observe their impact on network performance. In other words, we determine if biased P2P maintains its benefits across different scenarios.

As explained in Section III-B, we model content availability as well as session lengths as Uniform, Pareto, Weibull and Poisson distributions, thus giving us 16 possible combinations for the two characteristics. Hence, we run 16 different experiments for the biased P2P case for each topology. In this section, we focus on the World3 topology as the P2P nodes are nearly evenly distributed in each of the 16 ASes, thus minimizing the effect of topology on the metrics.

We see that across all the 16 combinations of content availability and online session lengths, the biased P2P topologies maintain their benefits for the P2P users as well as the ISPs. Consider the median file download time in Figure 2(a). Even though its value varies from $5.5-7$ seconds for biased P2P, it still remains below 7.8 seconds for unmodified P2P. The results for the mean AS distance of query responses are similar. In Figure 2(b), we witness a noticeable reduction in the number of AS hops between peers that send a query and peers that satisfy the query. Also, the mean overlay hop count of query responses in baised P2P cases remains comparable to that of unbiased P2P, as shown in Figure 2(c). This result has positive ramifications for mobile applications, where an increase in the overlay hop count can lead to performance degradations due to processing overhead at each additional node encountered in the path. The success rate of queries remains the same, while the number of responses to queries remains consistenly higher than that with the unmodified P2P system.

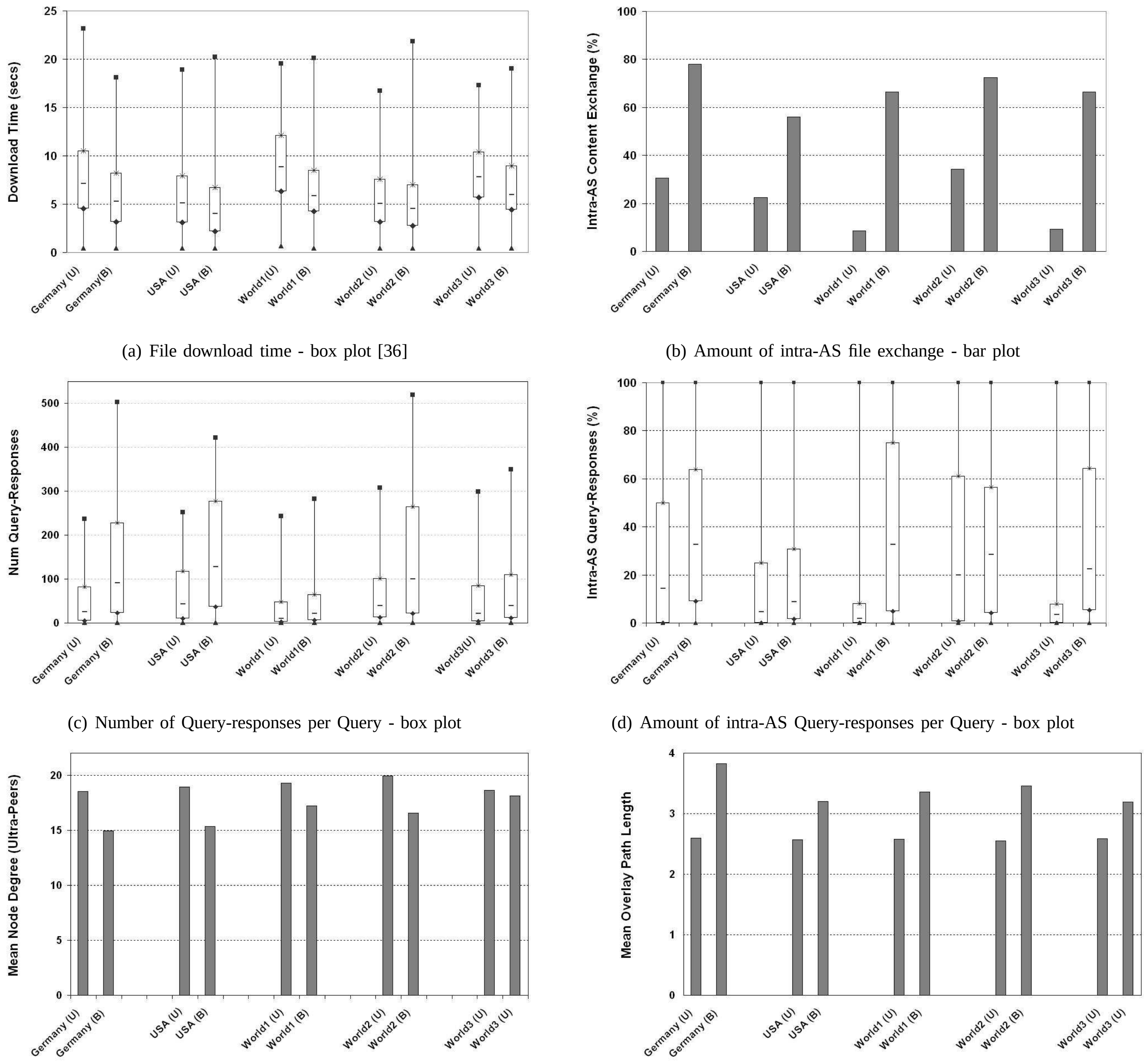
With regards to the graph properties, the node degree in Figure 2(d) remains largely unchanged, except for the case of Poisson session lengths. The results for the mean overlay path length between all pairs of nodes are also similar, see Figure 2(e). Although the graph properties are negatively affected by the Poisson session length distribution, we note that this distribution is not observed in real P2P systems, hence we can ignore this case.

Analyzing the benefits to the ISPs, we notice in Figure 2(f) that the amount of exchanged content that remains within the ISP network boundaries across all the tested scenarios ranges from $60-80\%$, significantly more than the 10% value observed in the case of unmodified P2P. This convincingly shows that ISP-aided P2P neighbor selection maintains its benefits across different user behavior patterns. Even the presence of a large number of free-riders, or a large number of peers who have very short online durations does not adversely affect localized P2P topologies. The inherent dynamic of P2P systems ensures that the overlay graph remains connected and maintains its nice graph structural properties, while ISPs as well as P2P systems benefit from co-operation.

V. Summary and Future Work

In this paper, we design representative ISP/P2P topology models and user behavior characteristics in a simulation framework, and study their impact on ISP-aided bandwidth-based localized neighbor selection for P2P users. Through extensive experiments, we show that both P2P users and ISPs benefit from collaboration, measured in terms of improved content download times, increased network locality of query responses and desired content, and overall reduction in P2P traffic. While ISPs benefit from a simple AS distance-based neighbor selection, P2P users benefit mainly by peering with nodes possessing higher last-hop bandwidth links. The advantages of ISP-P2P collaboration hold across different ISP/P2P topologies under a broad range of user behavior scenarios.

We now seek to further extend the oracle to consider ISP's router-level topology and its routing policies, and react to dynamic network conditions by considering "available" bandwidth at bottleneck links. We are implementing the oracle



(a) File download time - box plot [36]

(b) Amount of intra-AS file exchange - bar plot

(c) Number of Query-responses per Query - box plot

(d) Amount of intra-AS Query-responses per Query - box plot

(e) Mean node degree of overlay nodes - bar plot

(f) Mean overlay path length - bar plot

Fig. 1. Plots comparing unmodified(U) and biased(B) P2P neighbor selection across 10K queries and 10K file transfers, for 5 topologies

server in a scalable and efficient manner, and will run experiments in the Internet with other P2P systems like BitTorrent, KaZaa and P2P-TV system.

REFERENCES

[1] R. Steinmetz and K. Wehrle, *P2P Systems and Applications*. Springer Lecture Notes in CS, 2005.

[2] "CacheLogic: The True Picture of P2P Filesharing," http://www.cachelogic.com/. [Online]. Available: http://www.cachelogic.com/

[3] S. Seetharaman and M. Ammar, "On the Interaction between Dynamic Routing in the Native and Overlay Layers," in *INFOCOM*, 2006.

[4] V. Aggarwal, A. Feldmann, and C. Scheideler, "Can ISPs and P2P Systems Cooperate for Improved Performance?" in *ACM SIGCOMM Computer Communication Review, 37(3)*, 2007.

[5] G. Shen, Y. Wang, Y. Xiong, B. Zhao, and Z. Zhang, "HPTP: Relieving the Tension between ISPs and P2P," in *IPTPS*, 2007.

[6] R. Bindal, P. Cao, W. Chan, J. Medved, G. Suwala, T. Bates, and A. Zhang, "Improving Traffic Locality in BitTorrent via Biased Neighbor Selection," in *IEEE ICDCS*, 2006.

[7] T. Karagiannis, P. Rodriguez, and K. Papagiannaki, "Should ISPs fear Peer-Assisted Content Distribution?" in *IMC*, 2005.

[8] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically aware overlay construction and server selection," in *INFOCOM*, 2002.

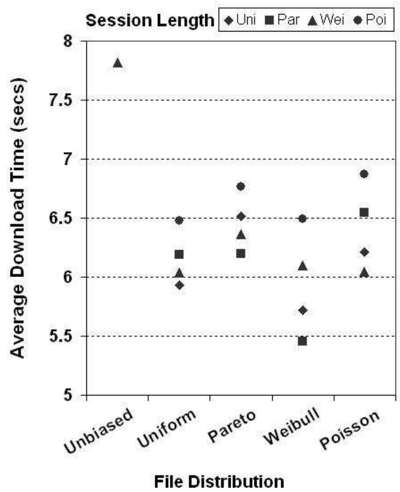
[9] S. Saroiu, P. Gummadi, and S. Gribble, "A Measurement Study of P2P File Sharing Systems," in *Technical Report UW-CSE-01-06-02, Dept of CS, Washington University*, 2002.

[10] R. Prasad, M. Murray, C. Dovrolis, and K. Claffy, "Bandwidth estimation: metrics, measurement techniques and tools," in *IEEE Network*, 2003.

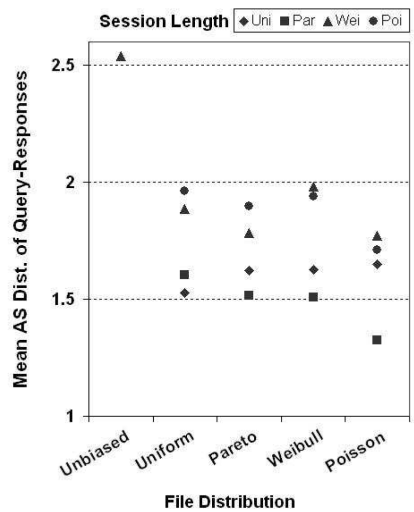
[11] Y. Zhang, N. Duffied, V. Paxson, and S. Shenker, "On the Constancy of Internet Path Properties," in *IMW*, 2001.

[12] J. Ledlie, P. Gardner, and M. Seltzer, "Network Coordinates in the Wild," in *NSDI*, 2007.

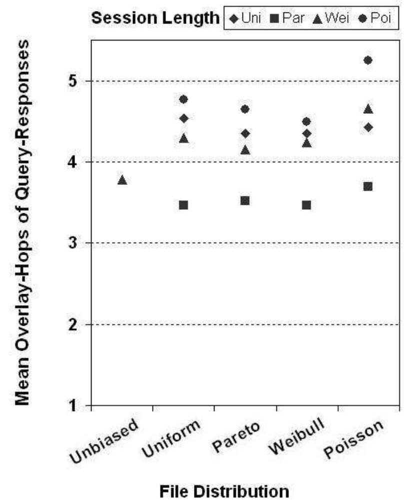
[13] Wong, Stoyanov, and Sirer, "Octant: A Comprehensive Framework for Geolocalization of Internet Hosts," in *NSDI*, 2007.



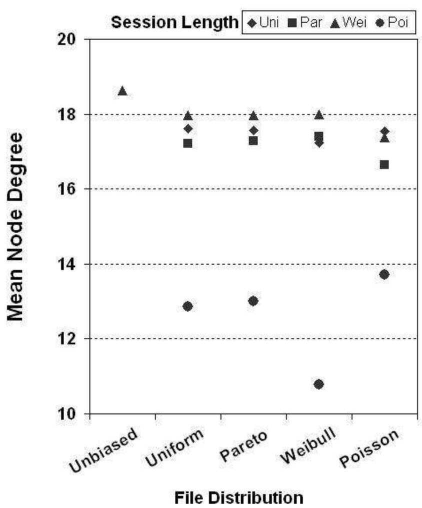
(a) Median file download time



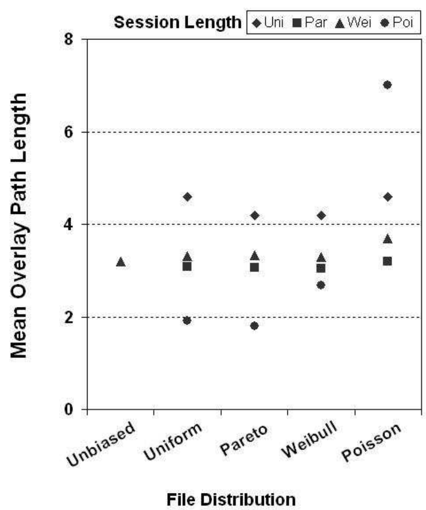
(b) Mean AS distance of Query-responses



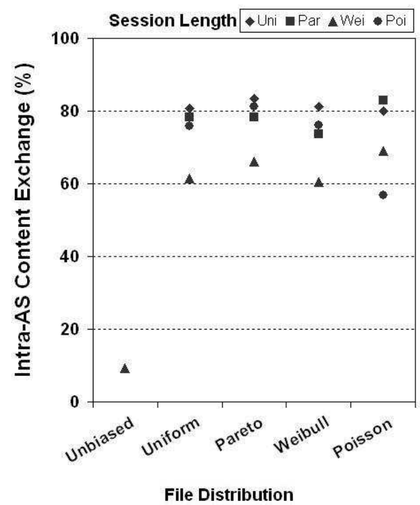
(c) Overlay hop count of Query-responses



(d) Mean node degree of overlay nodes



(e) Mean Overlay path length



(f) Amount of intra-AS file exchange

Fig. 2. Plots showing effect of user behavior (content availability and session length) patterns for World3 topology. X-axis denotes file distribution models, and symbols denote online session length models: Uniform, Pareto, Weibull and Poisson.

[14] S. Rewaskar and J. Kaur, "Testing the Scalability of Overlay Routing Infrastructures," in *PAM*, 2004.

[15] F. Dabek, R. Cox, F. Kaashoek, and R. Morris, "Vivaldi: A Decentralized Network Coordinate System," in *SIGCOMM*, 2004.

[16] F. Fessant, S. Handurukande, A. Kermarrec, and L. Massoulie, "Clustering in P2P File Sharing Workloads," in *IPTPS*, 2004.

[17] V. Padmanabhan and L. Subramanian, "An Investigation of Geographic Mapping Techniques for Internet Hosts," in *SIGCOMM*, 2001.

[18] V. Aggarwal, A. Feldmann, and R. Karrer, "An Internet Coordinate System to Enable Collaboration between ISPs and P2P Systems," in *ICIN*, 2007.

[19] "SSFNet," http://www.ssfnet.org.

[20] M. Liljenstam, J. Liu, and D. Nicol, "Development of an Internet Backbone Topology for Large-Scale Network Simulations," in *Winter Simulation Conference*, 2003.

[21] V. Aggarwal, A. Feldmann, and S. Mohrs, "Implementation of a P2P system within a network simulation framework," in *ECCS P2P-Complex Workshop*, 2005.

[22] "Gnutella wikipedia," http://en.wikipedia.org/wiki/Gnutella.

[23] S. Zhao, D. Stutzbach, and R. Rejaie, "Characterizing Files in the Modern Gnutella Network," in *MMCN*, 2006.

[24] "Germany Internet," http://www.internet-sicherheit.de/internet-deutschland.html.

[25] "TK Fachbegriffe," http://www.tk-fachbegriffe.de/index.php?id2=2400&a=320.

[26] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP Topologies with Rocketfuel," in *SIGCOMM*, 2002.

[27] W. Muehlbauer, O. Maennel, S. Uhlig, A. Feldmann, and M. Roughan, "Building an AS-Topology Model that Captures Route Diversity," in *SIGCOMM*, 2006.

[28] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A First-Principles Approach to Understanding the Internet's Router-level Topology," in *SIGCOMM*, 2004.

[29] S. Floyd and V. Paxson, "Difficulties in Simulating the Internet," in *IEEE/ACM Transactions on Networking, 9(4)*, 2001.

[30] H. Zheng, E. Lua, M. Pias, and T. Griffin, "Internet Routing Policies and Round-Trip-Times," in *PAM*, 2005.

[31] D. Stutzbach and R. Rejaie, "Understanding Churn in P2P Networks," in *IMC*, 2006.

[32] K. Ho, J. Wu, and J. Sum, "On the Session Lifetime Distribution of Gnutella," in *International Journal of Parallel, Emergent and Distributed Systems*, 2007.

[33] A. Klemm, C. Lindemann, M. Vernon, and O. Waldhorst, "Characterizing the Query Behavior in P2P File Sharing Systems," in *IMC*, 2004.

[34] J. Tian and Y. Dai, "Understanding the Dynamic of P2P Systems," in *IPTPS*, 2007.

[35] A. Gish, Y. Shavitt, and T. Tankel, "Geographical Statistics and Characteristics of P2P Query Strings," in *IPTPS*, 2007.

[36] "Box Plot - Wikipedia," http://en.wikipedia.org/wiki/Box_plot.