Peer-to-Peer Traffic Optimization Techniques: Postprint

Authors: Zhang Guoqiang, Tang Mingdong, Cheng Suqi, 张国清

Date: 2017-03-09T00:00:00+00:00

Abstract

Peer-to-peer (P2P) content distribution systems can substantially reduce distribution costs for content providers and enhance system scalability. However, the mismatch between overlay networks and underlying networks results in wastage of network resources, generates significant redundant traffic, and exacerbates conflicts between content providers and Internet Service Providers (ISPs). Therefore, effectively utilizing bandwidth resources of underlying networks and mitigating traffic pressure imposed by P2P content distribution systems on ISP networks constitutes a current research focus in P2P content distribution systems and is crucial for the sustainable development of P2P systems. This paper elaborates on various P2P traffic optimization techniques from three perspectives: P2P caching, location-aware technology, and data scheduling algorithms, examines applicable environments for each technique, and identifies existing challenges and future research directions.

Full Text

Abstract

Peer-to-peer (P2P) content distribution systems significantly reduce distribution costs for content providers and enhance system scalability. However, the mismatch between overlay networks and underlying physical networks leads to wasteful consumption of network resources, generates substantial redundant traffic, and intensifies conflicts between content providers and Internet Service Providers (ISPs). Therefore, how to effectively utilize underlying network bandwidth resources and reduce the traffic pressure of P2P content distribution systems on ISP networks has become a hot research topic and a key to the sustainable development of P2P systems. This paper elaborates on different P2P traffic optimization techniques from three dimensions: P2P caching, location-aware technology, and data scheduling algorithms. We discuss the applicable

environments for various techniques and identify existing problems and future directions.

Keywords: P2P traffic localization, P2P traffic optimization, P2P caching, location awareness, network coding, network topology

1 Introduction

In recent years, P2P-based content distribution systems have gained widespread popularity among users, including Kazaa, Gnutella, Emule, BitTorrent, PPLive, and PPStream. Compared with traditional client/server or Content Delivery Network (CDN) models, P2P content distribution dramatically reduces distribution costs for content providers and improves system scalability.

However, the enormous traffic generated by P2P applications has become a bottleneck restricting their further development. Measurements indicate that P2P traffic has surpassed Web traffic to become the dominant contributor to network traffic. Currently, in the triangle formed by end users, P2P content providers, and ISPs, the first two parties benefit significantly from the proliferation of P2P applications: end users experience faster download speeds, while content providers enjoy reduced distribution costs. ISPs, however, become the sole victims. ISP investments in bandwidth expansion are relentlessly consumed by P2P applications, preventing them from achieving returns on their investments. Moreover, excessive bandwidth occupation affects the normal operation of other services. Although ISPs could balance their cost-revenue ratio by increasing user fees, they face the risk of losing customers.

A primary reason for the massive P2P traffic is the inefficient utilization of underlying resources, resulting in substantial waste. Therefore, optimizing P2P traffic represents the only path to alleviating and eliminating tensions between content providers and ISPs. P2P traffic comprises control flow and data flow, with actual data transmission dominating network traffic. Consequently, optimizing data transmission is the key.

Efforts by content providers and ISPs to address massive P2P traffic can be divided into three stages: confrontation, independent action, and win-win cooperation. In the confrontation stage, ISPs employed Deep Packet Inspection (DPI) and other technologies to identify P2P traffic for blocking, rate-limiting, or shaping. In response, P2P content providers used dynamic ports, message encryption, and other mechanisms to hide their traffic. This led to a vicious cycle similar to that between viruses and anti-virus software, which industry insiders jokingly called a "war" between the two parties. In the second stage, both sides adopted proactive measures to optimize network resource utilization. Content providers used reverse engineering methods to infer underlying topology and network state information, leveraging this data to assist in building overlay topologies that matched the underlying network, thereby improving bandwidth utilization. ISPs, meanwhile, cached P2P traffic or employed Proxy-Tracker technologies to influence overlay topology construction and optimize network resource usage.

In the third stage, content providers and ISPs began experimenting with cooperative approaches to optimize network resource utilization. Under this architecture, ISPs provide network topology and state information services, which content providers access to optimize their topology construction and data scheduling. ISPs can control the granularity of provided network topology and state information, hide necessary private information, and generate revenue through these services. Content providers can obtain accurate location information to optimize resource scheduling, reduce topology probing overhead, and avoid the risk of being blocked by ISPs.

P2P traffic optimization mainly includes three categories of technologies: P2P caching technology, location and topology awareness technology, and data scheduling algorithms. This paper uses these as a starting point to deeply analyze and compare these three categories. Before detailing these technologies, we first provide a brief classification of P2P content distribution systems.

2 Classification of P2P Content Distribution Systems

Based on the nature of distributed content, P2P content distribution systems can be divided into non-real-time file distribution systems (e.g., Gnutella, Kazaa, BitTorrent) and real-time streaming media distribution systems (e.g., PPLive, PPStream). Based on P2P network structure, they can be categorized as unstructured or structured P2P. Structured P2P is built upon Distributed Hash Table (DHT) technology, offering low search costs but suffering from high network maintenance overhead and ineffective support for keyword-based fuzzy queries, thus failing to gain widespread adoption in content distribution systems. Currently, the majority of P2P traffic originates from unstructured P2P applications. Within unstructured P2P, we can further identify three generations based on employed technologies and chronological order. Napster represents the first generation of P2P content distribution systems, featuring a centralized index server. The second generation includes fully distributed Gnutella and hybrid systems like Gnutella and Kazaa. The third generation comprises file distribution systems like BitTorrent and real-time streaming systems like PPLive. However, despite all using P2P technology, Gnutella, BitTorrent, and PPLive exhibit significant differences. The key distinction between BT/PPLive and Kazaa is that in BT or PPLive, a node can serve other nodes after downloading only a portion of the entire file. Their overlay network serves as both a carrier for control information and a conduit for actual data transmission. In Kazaa, however, only nodes possessing complete file data can provide download services to others—the overlay network serves solely as a carrier for control information, while actual data downloads do not depend on the overlay network, and the concept of a cooperative network does not exist in the download process. As we will see later, different P2P traffic optimization techniques have varying applicability to these two distinct P2P file distribution models.

3 P2P Caching Technology

Caching technology is a solution used by ISPs to alleviate traffic pressure and reduce user response time, initially widely employed for Web services. With the surge in P2P traffic, caching it has become the preferred choice for ISPs to relieve traffic pressure. However, P2P differs from Web in traffic characteristics, and their caching design objectives also differ. A primary goal of Web caching is to reduce user-perceived response time, whereas the main design objective of P2P caching is to reduce network traffic. Web caching typically uses object hit rate and average user response time as evaluation metrics, while P2P caching generally uses byte hit rate as its evaluation metric. Therefore, efficient caching algorithms should be tailored specifically for P2P applications based on their traffic characteristics and caching design objectives.

3.1 P2P Traffic Characteristics

P2P application traffic characteristics differ significantly from traditional Web traffic, as shown in Table 1 . These differences impact various aspects of caching design. For instance, object size and popularity are major factors affecting caching efficiency, while protocol openness influences caching implementation complexity and scalability.

Table 1. Main Differences Between P2P Traffic and Web Traffic

Feature	P2P Traffic	Web Traffic
Object Size	Roughly three load levels: small files under 10M, medium files of several hundred MB, and large files over 1G	Generally small
Object Popularity	Popularity does not follow Zipf' s law; the most popular objects are far less popular than Zipf' s law predicts	Popularity follows Zipf' s law
Popularity Dynamics	Popularity can change abruptly. Objects may become popular overnight and see rapid popularity decline in short periods	Object popularity changes relatively smoothly
Object Mutability	Objects are immutable	Increasingly mutable objects
User Access Patterns	Most objects are downloaded only once by users	Users may access the same object multiple times

Feature	P2P Traffic	Web Traffic
Concurrent Sessions	Dozens of sessions can be opened simultaneously for downloading one object	One or a few sessions
Session Duration	May last for hours	Sessions typically complete within seconds
Protocol Openness	Numerous proprietary protocols	Standard HTTP-based protocols
Network Ports	Different networks use different ports	Single port (80)

3.2 P2P Caching Algorithms

Caching algorithms involve two core issues: (1) whether to cache entire objects or partial objects, and (2) cache replacement algorithms–i.e., how to replace cached objects when cache space is insufficient.

In Web caching, since objects are typically small, entire objects are usually cached. In P2P systems, however, objects can be extremely large. Caching entire objects would limit the number of objects that can be stored in a cache, potentially reducing caching effectiveness. Therefore, dividing objects into blocks or segments and having the caching system store only portions of objects becomes an attractive option.

Cache replacement algorithms are another major determinant of caching efficiency. Traditional algorithms include LRU (Least Recently Used), LFU (Least Frequently Used), MINS (MINimum Size), and MAXS (MAXimum Size). However, these algorithms do not work well for P2P systems. LSB (Least-Sent-Byte) [2] is a cache replacement algorithm designed for P2P that always replaces objects that have provided the fewest bytes to users.

P2P object popularity and popularity duration patterns also provide guidance for cache design. Research shows that P2P object popularity peaks quickly and then declines rapidly, dropping to 1/6 of its peak within 5-10 weeks [3]. This means that if frequency-based replacement algorithms like LFU are used, such objects would remain in the cache long after their popularity has waned, leading to inefficient caching. On the other hand, object popularity duration is not extremely short, typically lasting 2-3 months. This suggests that caching partial objects can gradually increase the cached proportion of an object, giving the caching system sufficient time to build popularity profiles. Consequently, the system can identify truly popular objects and progressively increase their cached proportion until the entire object is cached.

3.3 Problems with P2P Caching Technology

Although P2P caching technology is currently the preferred solution for ISPs to optimize P2P traffic, it suffers from several problems:

1. **Effectiveness**: P2P caching technology works well only when P2P nodes have weak location awareness capabilities. As P2P nodes' location awareness improves, the effectiveness of P2P caching will correspondingly decline.

2. **Scalability**: Since each P2P system uses different protocols, some of which are proprietary and non-public, the scalability of P2P caching systems is limited. To cache traffic from a new P2P application, one must first understand the protocol details of that application, increasing implementation complexity and reducing system scalability. In contrast, Web applications are based on open protocols, making Web caching more universal and easier to implement.

3. **Copyright Issues**: While Web caching also faces copyright issues, the problem is more pronounced for P2P content distribution systems due to the higher copyright sensitivity of distributed objects. P2P caching systems must carefully avoid violating copyright protection laws. It is worth noting that China still lacks clear relevant laws and regulations. Globally, only the U.S. Digital Millennium Copyright Act (DMCA) provides clear definitions, leaving a vast gray area in digital copyright protection for caching systems [4].

4. **Business Model Issues**: P2P caching violates the operational model of the entire network. ISP deployment of P2P caching is a reluctant, forced measure that cannot fundamentally resolve the conflict between ISPs and P2P content providers.

4 Location Awareness Technology

P2P networks are overlay networks built on top of underlying physical networks. The degree of matching between them largely determines how P2P nodes utilize underlying network resources. Early P2P networks selected neighbors randomly, resulting in severe mismatches between the constructed overlay and underlying network topology. This not only degraded application-layer performance but also generated substantial redundant traffic. Figure 1 [Figure 1: see original paper] illustrates examples of overlay networks constructed using location-aware methods (Figure 1(a)) and random methods (Figure 1(b)). Random overlay construction can cause serious mismatches with the underlying network, producing many long-distance transmissions during application-layer data transfer and wasting precious bandwidth resources. Measurements show [5] that in Gnutella networks, only 2-5% of connections are within the same Autonomous System (AS), yet over 40% of Gnutella nodes reside in the top 10 ASes, indicating that Gnutella networks essentially lack location awareness. Using location awareness technology can achieve routing-layer consistency between the constructed overlay and underlying topology. As shown in Figure 1(a), nodes that are close in the underlying network become more densely connected, thereby optimizing network resource utilization.

Figure 1. Two P2P Topology Construction Approaches
(a) P2P network constructed based on location awareness technology; (b) P2P network constructed based on random neighbor selection

The above example demonstrates that building an overlay network matching the underlying topology is crucial for improving P2P system performance and reducing network traffic pressure. This involves two research aspects: how to acquire and provide location awareness information, and how to use this information to assist topology construction.

4.1 Acquisition of Location Awareness Information

The purpose of location awareness is to predict network distance between nodes or estimate the proximity of candidate nodes. Based on the real-time nature of employed techniques, methods can be categorized as real-time or non-real-time.

4.1.1 Real-Time Measurement-Based Location Awareness Technology

Depending on whether network coordinates are used, real-time measurement-based location awareness can be divided into non-coordinate and network coordinate-based techniques.

Non-coordinate location awareness can rely on newly built dedicated infrastructure (such as IDMaps and Binning) or leverage existing network infrastructure (such as DNS and CDNs) to perceive relative network distance or proximity between nodes.

IDMaps [6] includes two concepts: tracer servers (Tracer) and Address Prefixes (AP). Tracers are a set of dedicated measurement servers distributed across the Internet. Tracers periodically measure latency between nodes and also measure latency to nodes adjacent to their address prefixes. Clients accessing IDMaps services collect all this latency information to build a virtual topology of the Internet composed of Tracers and address prefixes (as shown in Figure 2 [Figure 2: see original paper]), and estimate latency between any two IP addresses based on this virtual topology. Specifically, for two given IP addresses, the estimated latency equals the sum of the latency from each IP address’ s AP to its nearest Tracer plus the latency between these two Tracers.

The binning method proposed in [7] is a landmark-based measurement approach. Let landmark nodes be L1, L2, ···, Lk. When node P joins the system, it first measures its latency to these k landmark nodes, obtaining l1, l2, ···, lk. This yields a permutation L1’, L2’, ···, Lk’of landmark nodes ordered by latency. This permutation itself can serve as a classification basis. If nodes P1 and P2 obtain the same landmark node ordering through the above measurement and sorting, they are classified into the same category. This is because topologically close nodes should have similar latencies to different landmarks and thus produce the same ordering. Based on landmark ordering, absolute latency values can be further classified into levels to obtain finer-grained categorization.

Figure 2. IDMaps Architecture

King [8], developed by the University of Washington, leverages the existing Domain Name System (DNS) architecture for latency prediction, with the advantage of requiring no additional measurement platform deployment or protocol modifications. King works as follows: for any two given nodes A and B, it first finds the authoritative name servers NS(A) and NS(B) close to A and B, then relies on DNS recursive queries to obtain latency between NS(A) and NS(B), using this latency as the estimate for nodes A and B. Figure 3 [Figure 3: see original paper] illustrates King' s workflow. To prevent DNS servers from performing hierarchical queries from root servers, a pre-resolution is performed before latency measurement to allow name server NS(A) to cache NS(B)' s address. However, not all authoritative servers are close to their target hosts, so this method may produce non-negligible errors. Turbo King [9] improves upon this by reducing error margins and overcoming the large-scale DNS cache pollution problem caused by King.

Figure 3. Schematic of King Estimating Node Latency via DNS

A content delivery network information reuse method proposed in [10] predicts node proximity. CDNs use dynamic DNS redirection to provide customers with low-latency mirror servers, enabling optimized node selection by leveraging existing CDN infrastructure redirection behavior. This scheme assumes that any two nodes with similar redirection behaviors have a high probability of being close to each other—for example, they likely belong to the same ISP.

Network coordinate-based location awareness technology fundamentally embeds all network nodes, based on network distances (latency, bandwidth, packet loss rate, etc.), into a virtual space, assigning each node coordinates in this virtual space. Subsequently, network distance between any two nodes (e.g., Round Trip Time) can be calculated based on their virtual coordinates (e.g., Euclidean distance between coordinates). A major advantage of network coordinate systems is eliminating the additional probing overhead nodes need to determine latency, enhancing scalability of relative location awareness systems. For example, GNP [12] developed by Carnegie Mellon University is a static landmark-based network coordinate system, while Vivaldi [11] developed by MIT is a dynamic landmark-based system.

4.1.2 Non-Real-Time Location Awareness Technology Non-real-time location awareness primarily estimates proximity based on IP address structure. The Internet routes based on IP address prefixes. Each IP address corresponds to one or more routable IP address prefixes, expressed as a.b.c.d/l where l denotes prefix length. However, without routing system information support, nodes cannot know the exact prefix length of an IP address. Traditional globally routable unicast IP addresses include Classes A, B, and C, allowing nodes to simply determine their category and thus prefix length (8, 16, or 24). However, widespread use of subnetting and Classless Inter-Domain Routing (CIDR) creates accuracy and granularity issues for this method.

A more appropriate solution leverages routing system information. Routing tables store all IP address prefix information visible from that router. Two types exist: Border Gateway Protocol (BGP) routing tables for inter-AS routing and intra-AS routing tables. Typically, intra-AS routing tables have finer granularity for routes within their domain. However, ISPs generally 不愿 (are reluctant to) disclose their intra-AS routing tables as they are considered private information. Inter-AS routing tables are relatively easier to obtain.

A drawback of these methods is their reliance on pre-collected routing table prefix information, making them unable to adjust promptly to changes in network address allocation. A self-organizing distributed prefix matching method proposed in [13] has each node hash its IP address mask or random k bits into a key, then store its IP address on the node corresponding to that key in a DHT system. This allows newly joined nodes to easily find nodes with the same IP prefix. This method requires no pre-processed IP address prefix set and is fully self-organizing and adaptive but requires maintaining a DHT.

A coarser-grained proximity estimation method maps IP addresses to Autonomous System Numbers to determine whether different nodes reside in the same AS. IP-to-AS mapping can be achieved through BGP routing tables, which store paths (AS paths) to IP address prefixes. The first AS in an AS path is called the origin AS of the IP address prefix. Combining multiple BGP routing tables can build a mapping table between IP address prefixes and origin ASes. Given an IP address, the corresponding AS can be found through longest prefix matching.

4.1.3 ISP-Provided Information The above schemes rely on reverse engineering by content providers or third parties to perceive relative network distances. In reality, ISPs are the best providers of network topology and state information. Recently, technical proposals have emerged for ISPs to provide location information services. However, as this involves cooperation between ISPs and content providers, providing loosely coupled, scalable, secure, generic, and relatively flexible standardized service interfaces is fundamental to such methods. P4P [1] is a typical representative of this technology, enabling effective cooperation between applications and network providers to utilize network resources more efficiently and fairly while improving or maintaining P2P application-layer performance.

The P4P architecture proposes deploying an ingress tracking server (iTracker) within ISP networks. This device collects ISP network information and provides interfaces for interaction with P2P applications. These interfaces include static network policies, P4P distances reflecting network policies and states, and network capabilities (such as caching). These interfaces enable external applications to obtain relevant ISP network information while protecting ISP privacy, allowing joint optimization of network performance by ISPs and content providers. Figure 4 [Figure 4: see original paper] illustrates the P4P control layer entities and information flow. The appTracker is the P2P application' s tracker (which records download peer lists), interacting with iTracker and forwarding P4P control layer information to P2P application nodes.

Figure 4. P4P Control Layer Entities and Information Flow

Currently, a P4P working group has been established to promote P4P standardization, which has been welcomed by some U.S. ISPs and content providers. However, P4P still faces challenges, primarily whether ISPs and content providers have sufficient incentive to cooperate and adopt the P4P architecture. The P4P architecture also has not yet solved how to implement itself in a scalable and efficient manner.

Network Matching Service (NMS) is the main technology in the "P2P Traffic Optimization Technology Framework Based on Bearer Network Awareness" [14] proposed by the Institute of Computing Technology of Chinese Academy of Sciences, China Academy of Telecommunication Research, and major ISPs to the China Communications Standards Association (CCSA). It has become a national communications industry draft standard. This framework defines a service that provides information to help P2P applications make decisions, called Network Matching Service. Figure 5 [Figure 5: see original paper] shows the basic architecture of NMS, including functional entities such as bearer network information providers, network matching servers, network matching service discovery servers, and network matching service clients.

Network matching service clients can be ordinary P2P hosts or P2P index servers authorized to access certain network matching servers. They can utilize NMS to optimize provider node selection, P2P network topology, and relay or cache node selection. Network matching servers are entities implemented and deployed by ISPs to provide NMS. An ISP can independently deploy one or more network matching servers within its network, with each server establishing policies specifying which P2P applications and hosts are granted access. The "network matching service discovery server"functions to discover network matching servers and provide registration services. ISPs register their network matching server addresses and policies with the discovery server, and P2P hosts find authorized network matching servers through it. Network topology, policies, capabilities, and dynamic information required for NMS are provided by the bearer network, which specifically refers to ISPs or their proxy programs.

Figure 5. Basic Architecture of Network Matching Service

4.2 Location-Aware Topology Construction

The primary purpose of obtaining location information is to build location-aware overlay topologies. In systems with Trackers (such as BT and PPLive), the Tracker maintains nodes participating in current sessions. When a node attempts to join or replace neighbors, it requests a candidate neighbor set from the Tracker. The Tracker can allocate a group of geographically close candidate nodes based on location information. P2P application nodes can also autonomously optimize their topology during operation based on probed information. ISPs can assist application nodes in building location-aware topologies by deploying Proxy-Trackers.

Proxy-Tracker is deployed at a network' s boundary, maintaining information about nodes in that network participating in a P2P session. It intercepts candidate neighbor response messages sent by the P2P application' s Tracker to nodes, modifying the candidate neighbor set by replacing some distant nodes with nodes located in the same network as the requesting node, thereby assisting in building location-aware topologies. ISPs can also implement location-aware topology construction by providing redirection servers that maintain information about P2P application nodes and objects within their domain, intercepting user connection or data requests and redirecting them.

5 P2P Data Scheduling Algorithms

Data scheduling algorithms represent another option for optimizing P2P traffic. However, research on data scheduling algorithms typically focuses on application-layer performance improvement without considering impacts on underlying ISPs. Data scheduling algorithms are effective for third-generation P2P content distribution technology (where original files are divided into multiple data blocks and nodes form cooperative networks to distribute the original file). In such systems, nodes typically make decisions based only on local neighbor information (i.e., their own and neighbors' data block information) to determine which data block to download. Typical data scheduling algorithms include random scheduling, local rarest first scheduling, and more recent network coding-based scheduling [16][17].

Data scheduling algorithms alone cannot optimize P2P traffic; they must be combined with location-aware download strategies to achieve traffic optimization goals. Figure 6 [Figure 6: see original paper] illustrates the impact of three different data scheduling algorithms on P2P traffic. Assuming initially only node A has data blocks a, b, c, d: Figure 6(a) uses random download decisions, where nodes B, C, D, E might simultaneously request the same data block (e.g., block a) from node A, leaving connections among B, C, D, E underutilized in subsequent data transmission cycles. Figure 6(b) uses local rarest first strategy, but even this leads to uneven data distribution. For example, node B first requests block a from A, then node C requests from A and, following local rarest first, might request block b, then node D requests from A and might request block a again, and node E might request block b. Thus, B, C, D, E only have blocks a and b, making intra-domain links unusable after another transmission cycle, with new data only obtainable from node A. Figure 6(c) shows the case with network coding, where node A sends randomly linearly encoded blocks to nodes B, C, D, E. With a sufficiently large finite field such as F(2^8) or F(2^16), the four blocks transmitted to B, C, D, E are linearly independent with high probability. Consequently, nodes B, C, D, E can complete data transmission via intra-domain links to decode original blocks, achieving optimal resource utilization.

Figure 6. Impact of Data Scheduling Algorithms on P2P Traffic Optimization
(a) Random scheduling strategy; (b) Local rarest first strategy; (c) Network coding

Research [15] shows that with the same location awareness capability, local rarest first scheduling can reduce inter-domain traffic redundancy by 80%-90% compared to random scheduling. However, even with the strongest location awareness settings, local rarest first scheduling still maintains inter-domain traffic redundancy above 3.0, meaning each data block enters a domain more than three times on average. We compared local rarest first scheduling and network coding-based scheduling under the same location awareness capability. Our experiments used China' s AS topology with 4-core as the underlying AS topology, with 1000 nodes and average node degree of 10. Results shown in Figure 7 [Figure 7: see original paper] demonstrate that as location awareness capability strengthens, inter-domain traffic from both algorithms decreases significantly. However, when location awareness is not too weak, network coding produces less than half the inter-domain traffic redundancy of local rarest first scheduling.

Figure 7. Inter-Domain Data Redundancy Produced by Network Coding and Local Rarest First Scheduling Strategies

6.1 Comparison

Although P2P caching, location awareness technology, and data scheduling algorithms can all optimize P2P traffic, they differ in applicability, independence, and implementation approach. Table 2 summarizes these differences.

Table 2. Differences Among P2P Traffic Optimization Techniques in Applicability, Independent Usability, and Implementing Party

Technique	Applicability	Independent Usability	Implementing Party
P2P Caching	Applicable to all P2P applications; effectiveness decreases as P2P network location awareness improves	Can be used independently	ISPs

Technique	Applicability	Independent Usability	Implementing Party
Location Awareness	Applicable to all P2P applications	Can be used independently	ISPs, content providers, or third parties
Data Scheduling	Only applicable to applications with cooperative download networks like BT and PPLive	Must be combined with location awareness	Content providers

6.2 Summary and Outlook

P2P traffic optimization is a prerequisite for the healthy, benign, and sustainable development of P2P applications. This paper surveys different technologies for optimizing P2P traffic from three perspectives: P2P caching technology, location awareness technology, and data scheduling algorithms. While theoretical research on P2P traffic optimization has made considerable progress, widespread adoption by ISPs and content providers still requires effort. On one hand, the effectiveness of many theoretical techniques has only been verified through simulation experiments, lacking large-scale real-world Internet measurement results. In real Internet environments, the presence of Network Address Translation (NAT) and firewalls may significantly reduce the effectiveness of location awareness. Therefore, large-scale testing and evaluation in real Internet environments has become an urgent task to verify the effectiveness of location awareness technology for P2P traffic optimization. On the other hand, beyond technical feasibility, establishing appropriate business models among ISPs, content providers, and end users is key to promoting these technologies.

References

[1] H. Y. Xie, Y. R. Yang, A. Krishnamurthy, Y. B. Liu, and A. Silberschatz, "P4P: Provider portal for applications" , ACM SIGCOMM 2008.

[2] A. Wierzbicki, N. Leibowitz, M. Ripeanu, and R. Wozniak, "Cache replacement policies revisited: the case of P2P traffic" , in Proc. 4th Int. Workshop on Global and Peer-to-Peer Computing (GP2P' 04), Chicago, IL, Apr. 2004, pp.182-189.

[3] M. Hefeeda and O. Saleh, "Traffic modeling and proportional partial caching for peer-to-peer systems", IEEE/ACM Trans. Networking, vol.16, no.6, pp.1447-1460, Dec, 2008.

[4] 邹嵘，"基于 P2P Cache 的 P2P 流量优化技术"，电信网技术，2009 年 1 月

[5] M. Ripeanu, L. Foster and A. Iamnitchi, “Mapping the Gnutella Network: Properties of Large-scale Peer-to-Peer Systems and Implications for System Design” , IEEE International Computing Journal Special Issue on peer-to-peer networking, 2002.

[6] P. Francis, S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, and L. Zhang, “IDMaps: A Global Internet Host Distance Estimation Service” , IEEE/ACM Transactions on Networking, 9(5): 525-540, 2001.

[7] Sylvia Ratnasamy, Mark Handley, Richard Karp and Scott Shenker, “Topologically-aware Overlay construction and server selection” , In Proc. INFOCOM 2002, 2002

[8] K. P. Gummadi, S. Saroiu, and S. D. Gribble, “King: Estimating Latency between Arbitrary Internet End Hosts” , IMC’ 02, 2002.

[9] D. Leonard and D. Loguinov, “Turbo King: Framework for Large-Scale Internet Delay Measurements” , INFOCOM’ 08, 2008.

[10] D. R. Choffnes, and F. E. Bustamante, “Taming the Torrent: A Practical Approach to Reducing Cross-ISP Traffic in Peer-to-Peer Systems” , SIGCOMM’ 08, 2008.

[11] F. Dabek, R. Cox, F. Kaashoek, and R. Morris, “Vivaldi: A decentralized network coordinate system” , in Proc. ACM SIGCOMM, Portland, OR, Aug, 2004.

[12] E.Ng and H. Zhang, “Predicting Internet network distance with coordinates-based approaches” , in Proc IEEE INFOCOM, 2002

[13] C. Cramer, K. Kutzner and T. Fuhrmann, “Bootstrapping locality-aware P2P networks” , in.Proc. 12th Int. Conf. Networks(ICON’ 04), vol. 1, pp. 357-361, 2004.

[14] 中国科学院计算所等，“基于承载网感知的 P2P 流量优化技术框架”，通信行业标准，报批稿。

[15] Bindal et al, “Improving traffic locality in BitTorrent via Biased neighbor selection” , in Proc. IEEE ICDCS, 2006

[16] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, “Network Information flow” , IEEE Trans. Information Theory, 2000

[17] S. R. Li, R. W. Yeung, and N. Cai, “Linear network coding” , IEEE Trans. Information Theory, vol. 49, no. 2, 2003

[18] 唐明董，张国清，杨景，傅川，廖祝华，“P2P 流量优化技术综述”，电信网技术，2009 年 1 月

Author Biographies

Zhang Guoqiang is an assistant researcher and Ph.D. at the Institute of Computing Technology, Chinese Academy of Sciences. Email: guoqiang@ict.ac.cn

Tang Mingdong is a Ph.D. candidate at the Institute of Computing Technology, Chinese Academy of Sciences, and a lecturer at the School of Computer Science, Hunan University of Science and Technology.

Cheng Suqi is at the Institute of Computing Technology, Chinese Academy of Sciences.

Zhang Guoqing is a master' s supervisor and Ph.D. at the Institute of Computing Technology, Chinese Academy of Sciences. Email: gqzhang@ict.ac.cn

Note: Figure translations are in progress. See original paper for figures.