MESSAGE ROUTING IN PURE PEER-TO-PEER NETWORKS

Mojca Ciglarič, Matej Trampuš, Matjaž Pančur, Tone Vidmar
University of Ljubljana, Slovenia
Faculty of Computer and Information Science
Tržaška cesta 25, Ljubljana
Slovenia
{mojca.ciglaric, matej.trampus, matjaz.pancur, tone.vidmar}@fri.uni-lj.si

ABSTRACT

Pure peer-to-peer systems are loosely coupled, highly autonomous systems with a simple architecture, typically without hierarchy. Often, flooding-based routing mechanisms are used for sending messages through the overlay (application level) network. Due to the limited scalability and high network load experiences, the paper suggests an improvement to reduce the traffic and overall performance: if a peer remembers recently forwarded answers, it may route the next query messages asking about the same contents only to the relevant neighbour, not to all of them. The routing simulations have shown promising results.

KEY WORDS

File sharing protocol, distributed search, peer-to-peer system, message routing.

1. INTRODUCTION

Peer-to-peer (P2P) systems are relatively new form of distributed systems. Within the same application, peers can act as clients and as servers simultaneously: they can issue requests as well as respond to other peers' requests. A typical request is to access other peer's resources (files, processor cycles etc.) – information or services. Presently, they are being used in distributed search with content sharing, in collaboration tools (so-called groupware) and in real time communication (chat, instant messaging).

Detailed discussion of the most important P2P applications can be found in [1]. Pure P2P system consists only of peers, communicating with each other directly. Other architectures incorporate simple or more advanced servers that mediate peer connections or assist in resource discovery and management. Explanation of various pseudohierarchical system forms can be found in [2] and [3].

In a file-sharing system, users choose the files they want to share and they define the lifetime (TTL – time to live) of their queries, while query routing mechanism is coded into the P2P application. An existing mechanism of Gnutella protocol, implemented by the Gnutella client LimeWire is illustrated in [4].

Due to the protocol properties, network changeability and the differences among various P2P implementations, we cannot obtain an exact snapshot of the application-level network topology, neither we know how to obtain a realistic topology model. The problem is closely related to modeling the Internet topology on the network level. Faloutsos et. al. [5] have identified the existence of mathematical rules in the form of power laws in the Internet topology. Power laws are believed to hold in various natural networks and recent studies ([6], [7]) have shown that also P2P networks exhibit power-law properties. Adamic et al. [8] study exploitation of power law properties in P2P networks. Bu and Towsley [9] discuss weaknesses of several power-law topology generators and suggest a new generator, based on generalized linear preference.

Another interesting feature of the P2P network topology was found to be a small-world property. It was first described by Watts and Strogatz [10] and its relevance to P2P networks is discussed in [6], [7], [11], and [12].

Several authors address message routing improvement since it is obvious that flooding causes high amount of unnecessary message transfers. Their suggestions include partial ([14], [15]) or complete indices [13], but new problems arise due to higher vulnerability and index maintenance. In [8] and [20] authors discuss epidemic and gossip protocols: message spreading is more efficient when message visits more highly connected nodes. Iterative deepening of the flood and selection of promising neighbors is discussed in [15]. Experimental systems [15] – [19] demonstrate high efficiency but it is only achieved by means of strict topology and data placement rules. The issue of highly repetitive queries is not addressed properly.

Our paper addresses two issues. First, how to model a P2P virtual network topology. It is not a problem to generate a random graph. But where it comes to exhibition of the above mentioned properties (power laws, small world), many generators admit defeat. The second issue is an improvement of the default flooding-based message routing in a system where considerable amount of repetitive queries is detected. Our basic idea is that it is not needed to flood a query searching for a piece of data that we already found recently. It is enough to forward it along the way where the answer was found. In case the answer cannot be found again, we can still flood the query.

The extended protocol behaviour was simulated on the generated overlay topology model and it produced promising results.

The Section 2 of the paper describes *power law* and *small world* properties in P2P networks, which constitute a measure of topology evaluation, and presents a few suggestions for P2P topology simulation. Section 3 explains potential improvements of message routing in flooding-based P2P networks, and describes the simulations results. Section 4 suggests directions for further research and concludes the work.

2. P2P NETWORK PROPERTIES

Peers communicate with each other by passing messages over the network. In a pure P2P system, messages are routed the from the source peer to the destination peer at the application level – over (possibly) large number of in-between peers, connected by means of a P2P application. In a P2P system with a central server, messages usually travel through the server, thus taking the first application-level step from the source peer to the server and the second application-level step from the server to the destination peer.

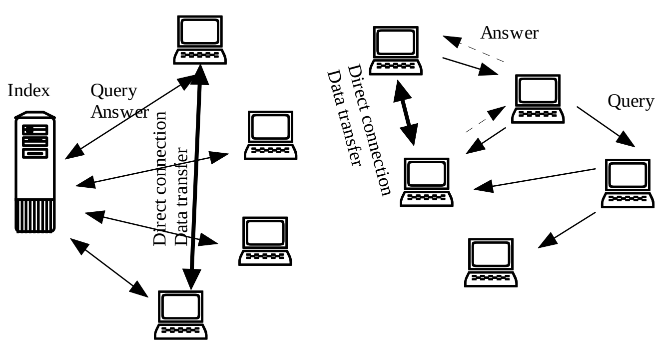

Figure 1: A server-mediated P2P system versus a fully decentralized one (pure P2P).

Fig. 1 shows the differences in the structure of a pure P2P system (on the right) and a server-mediated P2P system. The data transfer usually happens via a direct connection, not over a P2P network. While this is the most effective solution, it reduces the level of anonymity and unobservability since a destination network address has to travel through the P2P network all the way to the source peer.

2.1. Power laws

As the Faloutsos et al. have shown, the power laws are formulated as

$$f(x) \propto x^{\beta}. \tag{1}$$

The exponent is obtained by performing a linear regression on $f(x)$, drawn on the log-log graph. One of the four identified power laws concerns node out degree (the number of connections to other nodes). The fraction of the nodes with a given degree is proportionally related to the degree of the node – this is the so-called **out-degree** exponent.

Other power laws describe: **rank-exponent** (we sort the nodes in decreasing order of their out degree and the out degree is proportional to the rank of the node), **hop-plot exponent** (the total number of pairs of nodes within a given number of hops) and **eigen exponent** (the sorted eigenvalues of a graph are proportional to their order in a sorted sequence).

2.2. Small world

It was shown that many large networks occurring in nature, technology or sociology exhibit small world properties. Everyone of us has some small world experience: talking with a complete stranger at the party, we discover we have friends or relatives that know each other. In mathematical terms, small-world graphs exhibit small **diameter** (the average shortest path among all pairs of nodes) and are highly clustered. Clustering means that many of the neighbours of a certain node also know each other or are connected to each other. The measure for clustering is *C*, the so-called **clustering coefficient**, representing the ratio of number of existing connections among the node's neighbours to all possible connections among them (full graph), averaged over all the nodes.

Intuitively we can say the flooding-based mechanisms are not very well suited for small world networks: since more than a few neighborhood connections exist in the graph, flooding would cause unnecessary high message replication and network could become congested.

2.3. Topology simulation

Peer-to-peer network exhibits a virtual topology as an overlay on the IP network topology. The protocol itself does not define how exactly new nodes connect to the network, neither can we accurately guess which of the

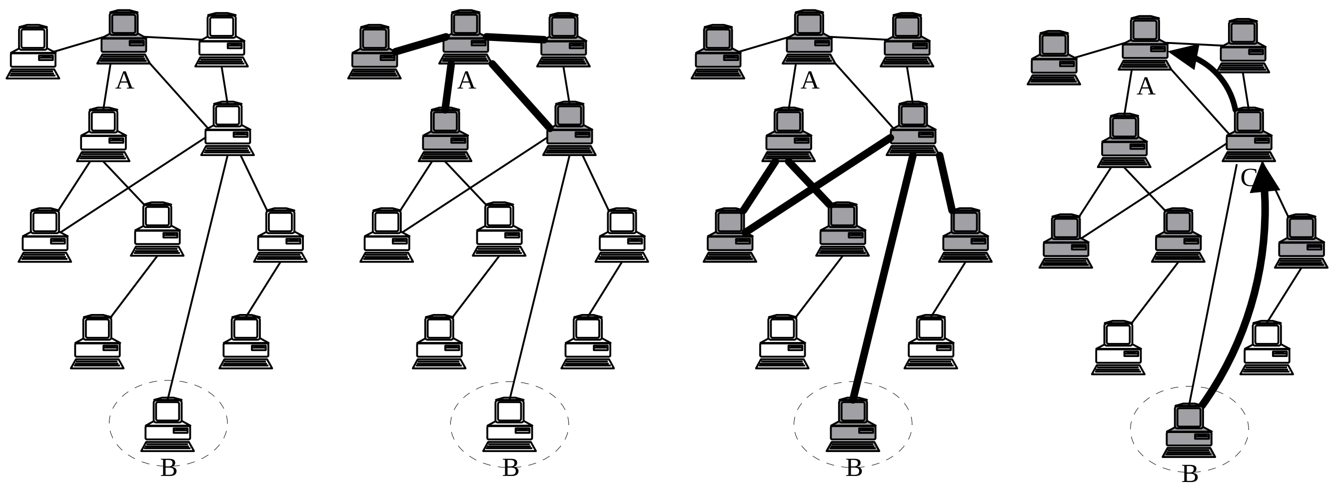

Figure 2: Flooding mechanism. A query is originated by A, an answer is found at B. Peers who have seen the query are shaded, thin lines are P2P connections, thick lines represent query hops, thick arrows represent answer hops.

contacted nodes it will connect to later on. If we want to test new or possibly improved protocols on some topology, we have two options. First, actual network topology can be discovered using a dedicated crawler program (for example *gnutellasim* from LimeWire.org) on an existing network. The accuracy of the result depends on the crawler suitability and on network changeablility – in cases with highly transient resources, the obtained topology might show distortions.

The second option is topology simulation. Several topology generators exist, some developed for Internet topology simulation with emphasis on power laws, and the others centered on small world properties, aimed at P2P overlay networks.

Waxman's generator places random points in the graph and connects them with a probability, given by a sort of power law formula. N-level hierarchical method generates several smaller random graphs and connects them hierarchically. Other options are Poisson random graphs, regular lattices and lattices with random rewiring. The latter model is a typical case of small-world generator.

As noticed in organic networks, also in large self-organizing structures like P2P networks new nodes have a tendency to connect to the nodes with higher degree (number of existing connections). Based on this feature, we decided to use a GLP (Generalized Linear Preference) generator from Bu and Towsley for preliminary simulations.

By now, we have modeled graphs of up to 100 nodes. GLP method reportedly produces graphs that follow all the four above mentioned power laws. We also found that the generated graphs of 100 nodes exhibit a clustering coefficient more than an order of magnitude smaller thab the clustering coefficient in a comparable random network, and short average shortest path lengths of typically less than 3 hops – similar as in random network. The observed properties help us admit that the generated network is appropriate for a message routing simulating and testing.

3. IMPROVED MESSAGE ROUTING

Our research focuses on flooding-based pure P2P systems (i.e. without hierarchy), without knowledge about the system structure and without global indexing scheme. Each peer only knows its neighbours and the messages he passes from one neighbour to another one (or all of them).

The described system structure is particularly well suited for a distributed search implementation. Query messages are flooded from originating peer to all of its neighbor peers, as shown in Fig. 2. Those peers who don't know the answer forward the message to all of their neighbors. In fact this is a breadth-first search. When an answer is found, the destination peer sends an answer message only to the peer from whom the query was forwarded earlier. In this way, answers are not replicated and follow the shortest route to the query originator.

Our research takes end user habits in account. What triggers a certain query? A situation develops, causing the user's urge to know or read or hear or see some information. Often the triggering event activates not one but many users, causing large amount of equal queries in a short time period, originated by distinct users. In addition to the message replication caused by the protocol itself, even more replication occurs and the P2P network reiterates same queries and answers on and on.

Our improved protocol assigns a new task to peer computers. A peer monitors the answers passing through and remembers where they came from and what the query

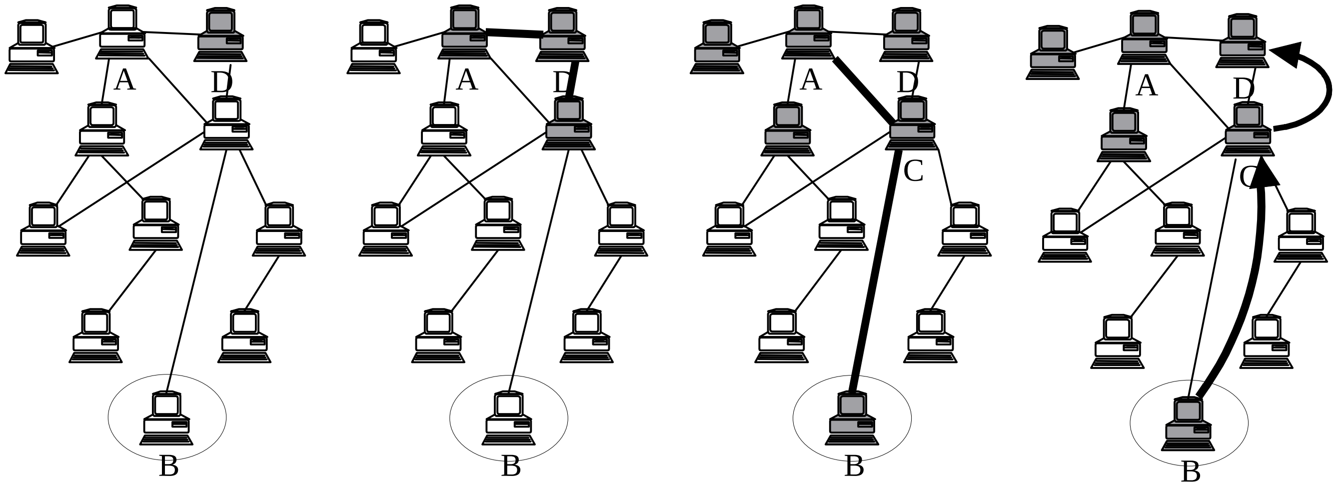

Figure 3: Improved routing mechanism. If a situation from Fig. 2 has happened earlier, C now answer came from B and therefore forwards the query only to B. A knows the answer came from C the query to C. The peers who haven't seen the answer would flood the query

asked for. When a replica of the same query arrives, the peer does not flood it to all of its neighbors. It rather forwards the query only to the peer that produced an answer earlier.

Let's consider the situation shown in Fig. 2. In the basic protocol, if another peer issues the same query, similar situation occurs and the query is forwarded about 10 times (cumulatively). In the improved protocol, the query is forwarded only to the peers from which the earlier answer was passed. Fig. 3 shows the whole procedure. We can see the query was forwarded only 4 times (cumulatively) until the answer was found.

The improvement is quite straightforward and does not demand considerable resources on peer machines. However to retain the hit ratio, the life time of cached information must be set properly to prevent excessive routing to non existing resources (peers which have disconnected from the network since they gave the answer).

The simulations were carried out on GLP-simulated topologies of about 100 nodes. Table 1 shows some typical topology properties.

Topology	Characteristic path length	Clustering coefficient	Avg. node degree
GLP	2,59	0,531	5,91
RANDOM	2,61	0,077	5,88
LATTICE	6,17	0,6	6,00

Table 1: Topology properties

The data is stored in the Oracle 9i database while the simulations and analyses are written partly in PL/SQL and partly in Java. Table 2 shows that using our improved protocol (2) instead of the basic flooding protocol (1) reduces the network query load by more than 20% while keeping the quality of service (the query response times – path the answer has to travel before reaching the query originator) at the same level as before. The result is quite promising and it is quite clear that the protocol exhibits similar efficiency in large networks too.

QUERY ROUTING	ANSWERS FOUND	ALL QUERY HOPS	ALL ANSWER HOPS	AVG. HOPS / QUERY.	AVG. HOPS / ANSWER	TOTAL HOPS
1	1312	5075	3125	38,72	2,38	8200
2	1080	3872	2513	50,75	2,32	6385

Table 2: Simulation results. Avg. hops per query denotes how many times a message was forwarded before it became obsolete, while avg. hops per answer denotes the average number of hops the first answer took from its source to the query originator.

The next step is to experiment with different topology sizes and types (random network, Poisson network, small world network, regular networks like grids or lattices…) and repetition degree. We also have in mind further routing optimizations to reduce repetitive queries with the same contents.

4. CONCLUSIONS

In the paper, basic properties of P2P systems are outlined. Two characteristics of P2P network topology are explained: power laws and small world property. The problem areas of P2P topology simulation and network overload are discussed and flooding-based message routing mechanism is presented. An improvement in message routing protocol is suggested to reduce replicated query traffic. The improved routing simulations on small topologies have shown the query traffic reduction by 20 percent. The suggested protocol is applicable in networks where replicated queries often occur. It is subject to further study to evaluate the minimum amount of replicated queries where the improved protocol still makes noticeable traffic reduction.

REFERENCES

[1] A. Oram (ed.), *Peer-to-Peer: Harnessing the Power of Disruptive Technologies* (Sebastopol, CA: O'Reilly & Associates, 2001).

[2] N. Minar, Distributed Systems Topologies, *O'Reilly Peer-to-Peer and Web Services Conference*, November 2001, http://www.openp2p.com/pub/a/p2p/2001/12/14/topologies_one.html

[3] M. Ciglarič, T. Vidmar, Position of modern peer-to-peer systems in the distributed systems architecture. In: *Communications and information technology for development: proceedings*, 11th IEEE Mediterranean Electrotechnical Conference MELECON 2002, Cairo, Egypt. Piscataway: The Institute of Electrical and Electronics Engineers, pp. 341-345, 2002.

[4] C. Rohrs, Query Routing for the Gnutella Network, *LimeWire LLC*, May 16, 2002, http://www.limewire.com/developer/query_routing/keyword%20routing.htm.

[5] M. Faloutsos, P. Faloutsos, C. Faloutsos, On Power-Law Relationships of the Internet Topology, in L. Chapin (ed.), Proceedings of the Conference on Applications, Technologies, Architecture and Protocols for Computer Communications (New York: ACM, 1999), pp. 251 – 262.

[6] N. Shiode, M. Batty, Power Law Distributions in Real and Virtual Worlds, ISOC2002, http://www.isoc.org/inet2000/cdproceedings/2a/2a_2.htm.

[7] T. Hong, Performance, in A. Oram (ed.), *Peer-to-Peer: Harnessing the Power of Disruptive Technologies* (Sebastopol, CA: O'Reilly & Associates, 2001).

[8] L. Adamic, R. Lukose, A. Puniyani, B. Huberman, Search in Power-Law Networks, *Physical Review* E, Vol.64, 46135 (2001).

[9] T. Bu, D. Towsley, On Distinguishing between Internet Power Law Topology Generators, in *Proceedings of IEEE Infocom*, 2002.

[10] D. J. Watts, S. H. Strogatz: Collective dynamics of Small-World Networks, *Nature*, 393 (June), 1998, 440-442.

[11] M. A. Jovanovic, F. S. Anexstein, K. A. Berman, Modeling Peer-to-Peer Network Topologies through "Small-World" Models and Power Laws, in *Proc. of IX. Telecommunications Forum TELFOR 2001*, Belgrade, 2001.

[12] M. Marchiori, V. Latora, Harmony in the Small World, *Physica* A 285, 2000, 539-546.

[13] M. T. Prinkley: An Efficient Scheme for Query Processing on Peer-to-Peer Networks, *Technical Report*, Aeolus Research Corporation, http://aeolusres.homestead.com, 2002.

[14] A. Crespo, H. Garcia-Molina: Routing Indices For Peer-to-Peer Systems, *Proc. Intl. Conf. on Distributed Computing Systems*, 2002.

[15] S. Ratnasamy, P. Francis, M. Handley, R. Karp, S. Shenker: A Scalable Content-Addressable Network. *Proc. ACM SIGCOMM 2001*.

[16] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, H. Balakrishnan: Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications. In *Proc. ACM SIGCOMM 01*, August 2001.

[17] A. Rowstron, P. Druschel: Pastry: Scalable, distributed Object Location and Routing for Large-Scale Peer-to-Peer Systems. In *Proc. 18th IFIP/ACM Intl. Conf. on Distributed Systems Platforms (Middleware 2001)*, November 2001.

[18] B. Zhao, J. Kubiatowicz, A. Joseph: Tapestry: An infrastructure for Fault-tolerant Wide-area Location and Routing. *Technical Report UCB/CSD-01-1141*, U.C. Berkeley, 2001. http://www.cs.berkeley.edu.

[19] C. Tang, Z. Xu, M. Mahalingam: PeerSearch: Efficient Information Retrieval in Peer-to-Peer Networks, *Technical Report HPL-2002-198*, HP Laboratories Palo Alto, July 2002.

[20] M. Portmann, A. Seneviratne: Cost-Effective Broadcast for Fully Decentralized Peer-to-Peer Networks. *Elsevier Computer Communications*, uncorrected proof, November 2002.