2010 Sixth International Conference on Networking and Services

CARMA based MST Approximation for Multicast Provision in P2P Networks

Gennadiy Poryev
National Technical University of Ukraine
03056 Kiev, Ukraine
poryevg@fh-trier.de

Hermann Schloss
University of Trier
54286 Trier, Germany
schloss@syssoft.uni-trier.de

Rainer Oechsle
Trier University of Applied Sciences
54293 Trier, Germany
oechsle@fh-trier.de

***Abstract*—The Minimum Spanning Tree (MST) problem is one of the most popular and important problems in the research area of distributed computing and networks. Contrary to the theoretical models where we usually have a global knowledge of all nodes and the corresponding distances for MST construction, in a realistic network (e.g., Internet) a node always has to rely on local knowledge only, that is it neither knows all other nodes nor exact distances between these nodes. In this paper we propose an approach for MST approximation based on local knowledge of a small subset of existing nodes by using the CARMA metric as a distance substitute. According to the evaluation results, our approach achieves a good MST approximation with respect to a communication cost and avoids extraneous communication needed for latency measurements.**

***Keywords*-Minimum Spanning Tree Problem; Application Level Multicast; P2P Networking**

I. Introduction

During the last few years the application scope of P2P systems has been notably extended. Being previously considered as a means of file sharing, nowadays P2P networks serve as a basic infrastructure for a wide range of innovative application scenarios such as massive multiuser environments or games.

While the problems of scalable data localization have been exhaustively addressed, the problem of reducing *multicast cost* in very large, global scale environments still remains inadequately considered. As shown by related work (Section II), the *Minimum Spanning Tree* (MST) problem is one of the most popular and important problems in the research area of graph theory, distributed computing and networks. In opposition to the theoretical models where we usually have a global knowledge of all nodes and the corresponding distances for MST construction, in a realistic network (e.g., Internet) a node neither knows all other nodes involved in the same application scenario nor exact distances between these nodes. In this paper we propose an approach for MST approximation (Section IV) based on local knowledge of a small subset of existing nodes and using the *Combined Affinity Reconnaissance Metric Architecture* (CARMA) metric as a distance substitute. We describe the computation of the CARMA metric in Section III. As the evaluation results show (Section V), our approach achieves a good MST approximation with respect to a communication cost metric and avoids extraneous communication needed for latency measurements. In Section VI we provide a brief overview of our contribution and discuss our future plans.

II. Related Work

ALMI (Application Level Multicast) project [1] uses a central instance for MST computation. However, since here all network nodes are considered for MST construction, high management and maintenance cost can be expected. We however avoid the additional communication, by using the CARMA metric and address the question of distributed approximation of a MST, that is, constructing a suboptimal spanning tree, whose communication cost is near-optimal.

Similar to our approach [2] have considered the problem of the construction of suboptimal spanning trees. In [3] the authors propose the construction of a *Nearest Neighbor Tree* (NNT) spanning tree instead of an MST. To ensure both acceptable multicast cost and latency delays JXTA[1] nodes always connect to the nearest (in terms of latency) node by this means achieving a MST approximation too. However, according to the quote in [4] "There is no satisfactory approximation algorithm known for MST problem". This encouraged us to address this problem in our work.

In [5] the authors propose a *binning* scheme by adjusting adjacent nodes to certain bins depending on their *Round Trip Time* (RTT) distance to certain landmark severs. To be more exact, a node measures its round-trip time to each of these landmarks and orders these landmarks in ascending order. Nodes having the same order of landmarks are then closer than nodes having different order. This approach significantly reduces the amount of communication necessary for capturing of node distance. However, at least the communication with landmark servers for RTT measurements is required.

In [6] the authors claim that most P2P systems use application level routing based on the overlay topology and completely neglecting the topology of the underlying physical network. Because of this, P2P systems cause a lot of extraneous traffic. In order to avoid this traffic, the authors propose the *Internet Service Provider* (ISP) aided neighbor selection by considering the node proximity in the underlying network at the application level.

[1]JXTA ia a hybrid P2P network: https://jxta.dev.java.net/ [10.12.2009]

978-0-7695-3969-0/10 $26.00 © 2010 IEEE
DOI 10.1109/ICNS.2010.25

The authors of [7] have recently described the design, deployment and evaluation of an approach minimizing the expensive cross-ISP traffic. The authors show that the application of their approach significantly reduces the latency delays. The *P4P* architecture [8] is aiming at the minimization of the network traffic, too. In order to achieve their objective, the authors take into account the conditions of the underlying network layer during the overlay construction.

III. CARMA Metric

In our previous research in the area of file sharing networks [9], we figured out a method which could be used to improve the performance of these file sharing networks. This method may also be applied for a wide range of P2P networks and even for networks in general. As we will show in this paper, the method may for example be used for MST construction.

There is wide variety of P2P file sharing networks. To name a few, these are: *eDonkey2000*[2] (ED2K) and *BitTorrent* [10]. Regardless of the differences in their protocols and implementations, there is something obviously in common in all file-sharing networks. That is, after the request for published entity is processed by either the indexing server or other nodes, a response is obtained in the form of a list of peers - network nodes that may serve the requested content.

From this point on, it is completely up to the client software to decide which nodes should be queried and in what order. From our previous experiences of analyzing ED2K and BitTorrent network traffic from a single node, we found out that client software usually performs queries in an unsorted order initially reported by network or index server. In the popular BitTorrent tracking servers, the number of peers for highly-demanded content could easily reach tens of thousands, whereas for most end-user nodes it is quite impractical to initiate more than hundred connections simultaneously even when having high-speed links. We believe that it is essential to not leave the peer selection process to a pure luck. In order to increase overall performance of a network, we introduce a special affinity metric providing the querying order. As mentioned before this metric has a general value and may also be used in MST construction as the distance substitute.

A. Analysis of existing differentiation methods

By their design ED2K clients will query every known source and will attempt to place themselves in the download queue of every source they managed to successfully negotiate with. The other (receiving) side will organize the download queue initially according to *First In, First Out* (FIFO) principle. Modern clients (eMule[3] and its numerous clones) also feature reward system, which advances inbound client in the queue according to the amount of related traffic they had provided to the node. This is supposed to discourage leeching but also have obvious drawback in delaying new nodes that do not have any part of the content yet.

Although eMule provides a few tuning methods such as queue rotation speed and chunk management based on file's popularity, none of it takes into account anything related to connectivity (client bandwidth, network latency etc).

BitTorrent protocol is simpler than ED2K. It does not feature any reward system, and due to per-content swarm isolation BitTorrent is generally faster. Also, tracker may not report all peers to the client initially, however this may be circumvented later by the peer exchange and *Distributed Hash Table* (DHT) mechanisms.

Recently there have been some advances in the locality awareness for BitTorrent networks. Popular nationwide trackers (torrents.ru, for instance) have introduced so-called "retrackers" - dedicated secondary servers. These servers are optionally connected to primary database but mainly supposed to only return peer list local to specific network scope. This scope usually consists of an IP address pool allocated to customers of a particular ISP. This provides for a significant speed burst for affected ISP clients, but it is very simple method that only allows for two-tier locality awareness.

B. Geographic of the Internet

Internet is in no way uniform structure. There are large backbone networks involved in international and intercontinental links, national-tier ISPs, end-user-servicing ISPs, hosting companies and, of course, end-users. Network latency and quality of service are accordingly very different depending on the link speed from tens of Gbps to the speeds of dial-up modems, less than 56 Kbps. On the scale of a country, Internet structure used to be organized rather sporadically - individual ISPs established arbitrary links among themselves and to foreign upstream ISPs. This had lead to peering conflicts and situations in which a message to a neighboring house traveled halfway the continent.

To alleviate this problem *Internet Exchange* points (commonly abbreviated as IX) were introduced. Usually, a number of national telecom operators create the dedicated facility to which all national ISPs then connect. Thus, consumer traffic within the scope of IX does not travel expensive international or satellite links. This helps to balance mutual peering and to ensure lower costs of maintenance per ISP allowing lower costs to end-users. Some developed countries have more than one nationwide IXes. From the customer point of view it is generally assumed that traffic within single IX flows faster and is cheaper than external. The presence of an IX can also provide for a lower hop count in packet path. The Figure 1 depicts an Internet segment covered by an IX.

[2]http://tinyurl.com/ed2klink/ [10.12.2009]
[3]http://www.emule-project.net/ [10.12.2009]

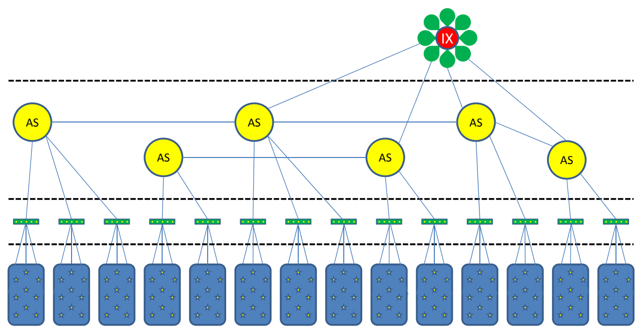


Figure 1. Schematic of an Internet Segment covered by an IX

C. CARMA Metric

In this section we describe the *Combined Affinity Reconnaissance Metric Architecture* (CARMA). Modern network modeling environments that deal with network topology rarely take locality into account. Most of them use either network latency metric measured in time units between request and response (ping) or hop count metric measured as number of nodes between source and destination hosts [11]. We deem ping method as generally unreliable as it heavily depends on link speeds and bandwidth conditions. For example, zero-loaded end-user ADSL link can produce slower pings than almost fully loaded Gbps link. As shown in [12] standard routing trace methods may also be unreliable and affected by bandwidth conditions or indicating non-existent links due to traffic switch-overs.

We propose the combined affinity metric which is calculated locally on each node and is only meaningful within the scope of this node. This metric is calculated given the remote IP address of the peer and all information than can be inferred from it. When dealing with Internet topology, CARMA is using a set of topology entities influenced by the database structure provided by RIPE registry [4] [5] (similar databases are accessible from other four regional Internet registries). The entities are listed below.

IPv4 range - a subset of IPv4 address space defined by origin node and host count. Note that host count is not necessary the power of 2 as implied by the Classless InterDomain Routing (CIDR) rules now commonly used for Internet routing. There are records that specify arbitrary number of nodes, but for practical reasons such definitions are subsequently augmented by CARMA to contain power of 2 number of nodes. The source for the range definitions is *delegated-ripencc-latest* file (see Listing 1).

AS - registered Autonomous System [13]. AS definitions are also listed in *delegated-ripencc-latest* file along with ISO country code and date of allocation. However, this file does not specify relation between ranges and AS, for

[4]ftp://ftp.ripe.net/ripe/dbase/split/ [10.12.2009]
[5]ftp://ftp.ripe.net/pub/stats/ripencc/ [10.12.2009]

```
ripencc|EU|ipv4|143.65.0.0|65536|19900326|assigned
ripencc|EU|ipv4|143.93.0.0|65536|19940413|assigned
ripencc|NO|ipv4|143.97.0.0|65536|20070104|assigned
ripencc|EU|ipv4|143.99.0.0|65536|19900907|assigned
```

Listing 1. Excerpt from database file with IP ranges

```
ripencc|EU|asn|2857|1|19931227|allocated
ripencc|EU|asn|2858|1|19940112|allocated
ripencc|SE|asn|2859|1|19940127|allocated
ripencc|EU|asn|2860|1|19940118|allocated

route:       143.93.192.0/18
descr:       FH-RPL-NET
origin:      AS2857
mnt-by:      AS2857-MNT
changed:     weiss@uni-mainz.de 20001220
source:      RIPE
```

Listing 2. Excerpt from database files with AS definitions and relations

which CARMA uses *ripe.db.route.gz* file (see Listing 2). The latter file contain definition blocks, each block specifies IPv4 range (this time in proper CIDR notation) and related AS. This information is used to establish relations between ranges and ASes listed in *delegated-* files. Note that relation between range and AS is not unambiguous - same range can be announced under different ASes; some ASes or ranges listed in *delegated-* file may not be linked and some relations specified in *ripe.db.route.gz* may contain AS and ranges unspecified in *delegated-* file. Every AS has a numerical identifier.

IPv4 subrange - a subset of IPv4 address space defined by origin node and end node. These definitions are listed in *ripe.db.inetnum.gz* file (see Listing 3). The subranges differ from ranges in that they are not explicitly related to AS. Subranges are generally smaller in terms of address space. Vast majority of them are derived from splitting up ranges. It is therefore possible to establish a relation between one or more subrange and single range, although not all ranges are split into subranges. When parsing information from this file, one should take care to check for sanity of the subranges specified. For instance, one of subranges specify an entire IPv4 address space, another subrange have netmask length of 3 bits - such cases are obviously invalid for CARMA.

AS set or *asset* - topological junction point that may contain arbitrary number of ASes and facilitate connectivity among them. It is assumed that information flow between

```
inetnum:     143.93.32.0 - 143.93.63.255
netname:     FH-RPL-NET
descr:       Fachhochschule Trier
descr:       Rechenzentrum
descr:       Schneidershof
descr:       D-54293 Trier
country:     DE
admin-c:     KM624-RIPE
tech-c:      RB373-RIPE
status:      ASSIGNED PI
mnt-by:      TRANSKOM-MNT
changed:     hostmaster@transkom.net 20050207
source:      RIPE
```

Listing 3. Excerpt from database file with IPv4 subranges

```
as-set:        AS-DECIX-CONNECTED
descr:         ASN of DE-CIX members
descr:         DE-CIX, the German Internet Exchange
admin-c:       AN6695-RIPE
tech-c:        WT6695-RIPE
tech-c:        DM6695-RIPE
tech-c:        SJ6695-RIPE
notify:        notify@de-cix.net
mnt-by:        DECIX-MNT
source:        RIPE
changed:       auto-upd@de-cix.net 20091011
members:       AS42
...
members:       AS2828
members:       AS2857
members:       AS2914
...
members:       AS65333
```

Listing 4. Excerpt from database file with asset definitions

two ASes belonging to the same asset takes no intermediate route. Unlike AS, assets have alphanumeric identifiers. In terms of CARMA, Internet Exchange point is an asset with significant number of member ASes (usually hundreds), although, technically, every asset can be considered as a kind of IX. The definitions for asset may be found in *ripe.db.as-set.gz* file (see Listing 4).

When all database files are processed, the resulting incomplete graph reflects Internet topology as close as it could be done without having access to BGP information. It is not necessary to devise any graph-walking algorithm to calculate *flavor* (affinity value) because the purpose of CARMA is to estimate affinity of two given nodes, not calculating an exact hop count. The proposed flavors of the remote node in relation to the originator node are given below in the order of corresponding tests undertaken by CARMA.

1) Flavor 0 "Same subrange" identifies the presence of remote node's IP address in the same IPv4 subrange specified in *ripe.db.inetnum.gz* database file. However, if such presence is found, CARMA does not immediately return this flavor because subranges may overlap with different netmask length, which in turn may happen to be shorter than that of corresponding range. This flavor identifies the presence of remote node most likely within the scope of operation of single router or same network operations center. For example, this could be end-users connected to the same point of presence of telecom operator, or nodes within the university network, which usually have single upstream ISP.
2) Flavor 1 "Same range" identifies the presence of remote node's IP address in the same IPv4 range specified in *delegated-* or *ripe.db.route.gz* files. If subrange lookup had yielded any results, the ranges found are examined and compared in terms of netmask length. In this case, "range" flavor is only returned if shortest range netmask is shorter or equal than that of a subrange, otherwise "subrange" flavor is returned. This ensures that subrange flavor is never returned for allocations larger than corresponding range, even if they overlap. This flavor suggests that the traffic between nodes is unlikely to travel outside of the single business network of their ISP.
3) Flavor 2 "Same AS" identifies the presence of remote node's IP address within the address space allocated to the same AS as defined in *ripe.db.route.gz*. This flavor suggests that the traffic between nodes is handled by the ISP internally and that incoming traffic going from outside part of Internet to the remote node undergo the same routing rules as traffic for the originator node.
4) Flavor 3 "Same asset" states that both originator's and target's nodes belongs to the ASes announced by the same asset (not every asset is IX, all IXes are assets). The immediate advantage of this knowledge is not obvious but in the developing countries the difference in quality of service may largely depend on this flavor to the extent that network speed and latency differ by the *orders of magnitude* for nodes within and outside of IX.
5) Flavor 4 "Distant" identifies that the affinity of originator's and target's nodes can not be reliably estimated and therefore they assumed to be located topologically far away.

IV. CARMA BASED MST APPROXIMATION

According to [14] the consideration of node locality is the key to efficient communication in P2P systems. It improves performance and increases availability, since the probability of transmission failure increases with distance and depends on bandwidth conditions.

One problem of constructing a MST in real networks is that we don't have exact distances between nodes (latency delays) as we do in the graph theoretical research. Measurement of the round trip latency between nodes for the purpose of distance acquisition by sending extraneous ping messages induces an inacceptable high communication overhead in large scale networks and hence has to be avoided. As mentioned before, CARMA flavors indicate the node locality by telling whether a remote peer belongs to the same subnet, same AS or same IX. Therefore, in our approach we utilize the CARMA flavors as a distance substitute for spanning tree approximation.

Another problem we have to address is the approximation of global knowledge needed for spanning tree construction. Most of existing P2P networks designed for provision of application level multicast use the bootstrapping process which involves a list of nodes identified by their IP addresses presumed to be online. That is, the initial knowledge of a node is limited to these nodes from the bootstrapping subset. In our work we assumed this list to contain between $log(N)$ and $\sqrt{N}$ entries, where N is the number of network nodes. In our approach for approximation of a simple MST we choose a node depending on its CARMA flavor. For

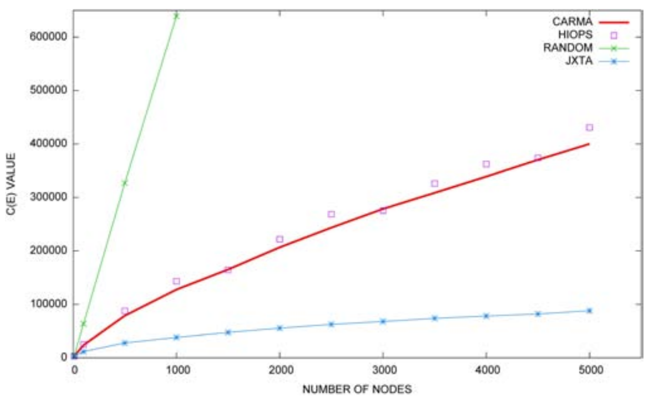


Figure 2. Communication Cost

this purpose we first sort nodes from the bootstrapping set depending on their CARMA flavor and then choose an arbitrary node within the flavor identifying best network conditions. As mentioned before no additional communication and computation is required for CARMA based MST construction.

V. Evaluation Results

For our computation we model the Internet as an undirected and connected graph $G = (V, E, w)$. Hereby V stands for the set of vertices $v_i,\ 1 \leq i \leq |V|$ representing network nodes, E is the set of edges $e_{i,j} = \{v_i, v_j\}$ representing the logical connections between nodes and $w : E \rightarrow \mathbb{N}$ is a weight function assigning a edge its weight. [15] proposes the reduction of multicast cost by using a MST. Hereby the multicast cost $C(E)$ is denoted as the cost for propagating a message to all recipients in the group which is the sum of all edge weights in the tree representing latency delays along any path taken by the message:

$$C(E) = \sum_{e_{i,j} \in E} w(e_{i,j}) \tag{1}$$

We use the $C(E)$ metric as the quality function for comparison of different approaches.

In order to provide meaningful results, we compare our CARMA based approximation with some of the existing P2P approaches supporting application level multicast such as ALMI and JXTA and HiOPS [15]. To extend the range of our evaluation we have also considered a *RANDOM* infrastructure, where a new node build up connections to a randomly selected node from the bootstrapping set. ALMI, JXTA and HiOPS (Figures 3(a), 3(b) and 3(c)) approaches require a certain amount of global knowledge. Moreover, ALMI and HiOPS do adjust their infrastructures on joining or leaving of nodes. This global knowledge and readjustment, makes them useless for large scale application scenarios. On the contrary, CARMA and RANDOM (Figures 3(d) and 3(e)) infrastructures rely only on local bootstrapping sets and insert new nodes into the network graph without any adjustment, therefore providing a scalable solution for large-scale applications.

To compare these networks with respect to communication cost $C(E)$, we have set up a simple simulation environment. Using this environment we can create an arbitrary number of network nodes, interconnect them according to a given algorithm and then compute the communication cost metric $C(E)$. We have performed several evaluation runs where we randomly created 10, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500 and 5000 nodes interconnecting them as ALMI, JXTA, HiOPS, CARMA and RANDOM infrastructures. After each run we computed the communication cost $C(E)$ in milliseconds needed for propagation of a multicast message to all existing nodes.

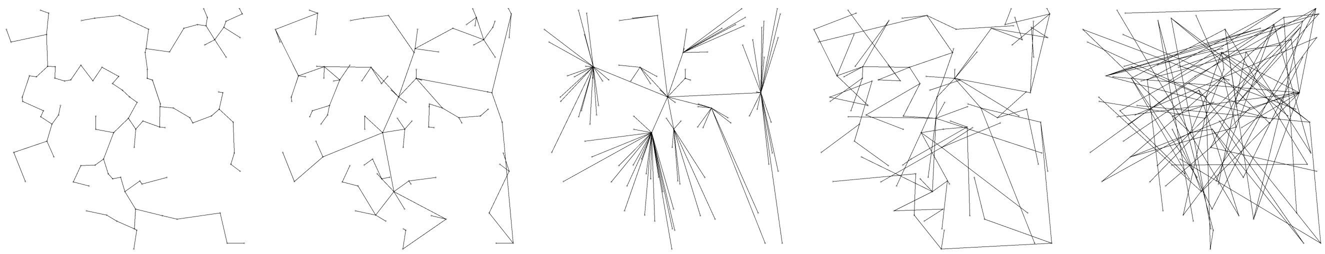
As the Figure 2 show, the JXTA and HiOPS approaches relying on global knowledge do provide low $C(E)$ values. But as mentioned before, these do not scale in terms of large number of users. As expected, the RANDOM infrastructure incurs the highest communication cost. The scalable CARMA approach provides nearly the same $C(E)$ values as the HiOPS overlay and relies only on the local knowledge as the RANDOM infrastructure, by this means providing a good trade-off between construction and communication costs. The binning approach [5] would show almost the same behavior in terms of the communications cost as the CARMA approach. However, CARMA does not require any additional communication for ordering the nodes in the bootstrapping set, whereas nodes following the binning approach have to contact the landmark servers.

VI. Conclusion and Future Work

In this paper we presented a MST approximation approach based on the local knowledge and on the CARMA metric as a distance substitute. As the evaluation results show, even the simple algorithm of connecting to an arbitrary node with a flavor identifying best network conditions provide better communication cost $C(E)$ values than the random selection approach. Contrary to the approaches computing an optimal MST or relying on the global knowledge, our approach does not require any additional communication and computation for spanning tree construction, thus providing a good trade-off between construction and communication costs.

In this paper we largely focused on the aspect of CARMA that provide benefit from preliminary topological knowledge and do not involve any measurement-related traffic. However, this is only first of three CARMA layers we will propose in our future work. In addition to the flavors of affinity, there are several additional components that can be included in metric calculation, as follows:

1) average response time to keep-alive requests;
2) average hop count to the destination, including the possibility of its change during communication [12];
3) bandwidth and average consumption at the moment of decision, including preset constraints;



(a) ALMI (b) JXTA (c) HiOPS (d) CARMA (e) RANDOM

Figure 3. MST Approximations in a Network with 100 Nodes

4) "gratitude" and "greed" values calculated as the amount of traffic the remote party had provided and consumed respectively;

Generally speaking, we consider CARMA three-layered, with first layer being the aforementioned locality awareness expressed in flavors, the second layer that utilizes additional traffic but does not involve actual P2P communication and third layer that require active communication to the remote party over compatible protocol. By its design, CARMA is meant to be dynamically changing as the communication goes, reflecting and adapting to the changes in bandwidth conditions.

In our future work we would like to address the second and third layers of CARMA calculated by direct measurements involving additional traffic. We plan to demonstrate the effectiveness of CARMA approach by implementing it in real-life P2P networks and extensive experiments, for which we are developing the software library implementing CARMA method under LGPL license to assist software engineers wishing to optimize performance of their P2P applications.

REFERENCES

[1] D. Pendarakis, S. Shi, D. Verma, and M. Waldvogel, "Almi: an application level multicast infrastructure," in *USITS'01: Proceedings of the 3rd conference on USENIX Symposium on Internet Technologies and Systems*. Berkeley, CA, USA: USENIX Association, 2001, pp. 5–5.

[2] D. Peleg and V. Rubinovich, "A near-tight lower bound on the time complexity of distributed minimum-weight spanning tree construction," *SIAM J. Comput.*, vol. 30, no. 5, pp. 1427–1442, 2000.

[3] M. Khan and G. Pandurangan, "A fast distributed approximation algorithm for minimum spanning trees," *Distributed Computing*, vol. 20, no. 6, pp. 391–402, April 2008. [Online]. Available: http://dx.doi.org/10.1007/s00446-007-0047-8

[4] E. Michael, "Distributed approximation: a survey," *SIGACT News*, vol. 35, no. 4, pp. 40–57, 2004.

[5] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, 2002, pp. 1190– 1199 vol.

[6] A. Feldmann and V. Aggarwal, "Isp-aided neighbor selection in p2p systems," in *IETF P2P Infrastructure Workshop (P2Pi)*, Berlin, Germany, May 2008.

[7] D. R. Choffnes and F. E. Bustamante, "Taming the torrent: a practical approach to reducing cross-isp traffic in peer-to-peer systems," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, pp. 363–374, 2008.

[8] H. Xie, A. Krishnamurthy, A. Silberschatz, and R. Y. Yang, "P4P: Explicit Communications for Cooperative Control Between P2P and Network Providers," DCIA P2P MARKET CONFERENCE, 2008.

[9] G. Poryev, T. Rudyk, and O. Sulima, "Traffic Regulation and Reputation Handling in the BitTorrent Peer-to-Peer Networks," National Technical University of Ukraine "KPI", Tech. Rep., 2008.

[10] B. Cohen, "Incentives build robustness in BitTorrent," in *the First Workshop on the Economics of Peer-to-Peer Systems*, 2003.

[11] G. Lucas, A. Ghose, and J. Chuang, "On characterizing affinity and its impact on network performance," in *MoMeTools '03: Proceedings of the ACM SIGCOMM workshop on Models, methods and tools for reproducible network research*. New York, NY, USA: ACM, 2003, pp. 65–75.

[12] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira, "Avoiding traceroute anomalies with Paris traceroute," in *IMC '06: Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. New York, NY, USA: ACM, 2006, pp. 153–158.

[13] J. Hawkinson and T. Bates, "Guidelines for creation, selection, and registration of an Autonomous System (AS)," RFC 1930 (Best Current Practice), Internet Engineering Task Force, Mar. 1996. [Online]. Available: http://www.ietf.org/rfc/rfc1930.txt

[14] J. Kubiatowicz, "Extracting guarantees from chaos," *Commun. ACM*, vol. 46, no. 2, pp. 33–38, 2003.

[15] H. Schloss, R. Oechsle, J. Botev, M. Esch, A. Höhfeld, and I. Scholtes, "HiOPS overlay - efficient provision of multicast in peer-to-peer systems ," in *16th IEEE International Conference on Networks (ICON 2008)*, New Delhi, India, 2008, pp. 1–6.