P2P MANETs – New Research Issues

Mario Gerla, UCLA; Christoph Lindemann, University of Dortmund; Ant Rowstron, Microsoft Research

1. What are P2P MANETs?

A mobile ad hoc network (MANET) is a collection of mobile nodes that dynamically self organize in a wireless network without using any pre-existing infrastructure. In a MANET, the applications are typically “peer-to-peer” rather than “client-server”. Moreover, a MANET is often built to support a specific application, thus the networking is application-driven. For these reasons, we often refer to the ad hoc networking done in a MANET as peer-to-peer (P2P) networking. This however is not quite consistent with the general meaning of P2P in the broader Internet context. In the Internet, a P2P network is basically an overlay network justified by the need for specialized functions that are not possible (or not cost-effective) in the IP layer. These functions must be performed at the middleware or application layer. Classic examples of Internet overlay networks are: real time multicast overlays (which overcome the lack for multicast support in the IP routers), unstructured P2P systems such as Gnutella [Gnu], BitTorrent [PG+05], and KaZaA [Sha], and; structured P2P systems such as Pastry [RD01], Chord [SM+01], and CAN [RF+01]. Structured P2P overlays are typically implemented as content addressable networks and permit efficient “content” based routing, which is otherwise not possible in the IP layer.

MANETs are based on a network architecture that is quite different from the wired Internet architecture. In fact, a typical MANET architecture reminds us of the architectures introduced in the ‘80s to interconnect LANs, with level 2 routing which is based on MAC addresses. The MANET routing functionality is non-hierarchical on the level-2 and, thus, is several layers below the functionality offered by a P2P network. This implies that in a MANET the P2P concept will be implemented on top of level-2 routing. There is a definite need for overlay/P2P networking in a MANET, for the following reasons: (a) the MANET routing layer is often inadequate to provide the services needed by sophisticated mobile applications, and; (b) the unpredictability of the radio channel combined with the mobility of the users can pose major challenges to routing, requiring upper layer intervention. The leading strategy today is to keep MANET routing and transport protocols simple (mainly for ease of standardization), and; to complement them when necessary with upper layer functions via overlays and P2P networking.

Dagstuhl Seminar Proceedings 05152
Perspectives Workshop: Peer-to-Peer Mobile Ad Hoc Networks - New Research Issues
http://drops.dagstuhl.de/opus/volltexte/2005/213

As an example of wireless overlay, consider a large ad hoc network deployed to overcome a major natural disaster. A sad and very timely example is the South Asia Tsunami disaster. The ad hoc network may be a combination of heterogeneous technologies – from satellites to ground ad hoc radios and improvised cellular and mesh network services. Different teams are formed – a few teams will cooperate in a particular mission. For instance, three or four different teams may search for survivors; others will be in charge of distributing food and supplies; there will be medics teams providing first aid and medications; engineering teams for reconstruction; police teams preventing looting, etc). These teams move and operate as groups. They must coordinate their operations. They require multicast and possibly content based routing (where is the proper tool or medication?). It thus makes sense to develop "team based multicast" and "content based routing" schemes at the user level, on top of the very basic routing service provided by the "instant", ad hoc, heterogeneous infrastructure.

As another example, consider a "delay tolerant" file sharing application that includes hosts partly in the Internet and partly on peripheral wireless ad hoc networks. Wireless nomadic users can rapidly change their connectivity to the Internet from Kbps (say GPRS) to Mbps (say, 802.11). Occasionally, the users may become disconnected. The use of the standard network routing protocols may lead to inefficiencies, violation of delay constraints and possibly retransmission of large portions of the file. A P2P overlay network can keep track of connectivity among the various hosts. The overlay network can extend to wired, wireless and ad hoc network segments. It can predict disconnection/reconnection dynamics and can exploit them to deliver files efficiently and within constraints (for example, using intermediate proxy nodes for "bundle" store-and-forwarding).

As a third example, consider an instant messaging (IM) system for a pure MANET as a "delay sensitive" application [LLW05]. "Presence" technology enables users of an IM system to determine if their contacts are online, signed onto the IM application, and ready to communicate. Presence technology is used in a number of commercial applications other than classical IM systems e.g., in numerous computer-supported-cooperative-work applications. In fact, the 2003 release of Microsoft's Office constitutes one of the most prominent examples of a presence-enabled application. The protocol design for disseminating presence information in the Internet has been matured and organizations such as the IETF and the Jabber software foundation have developed protocol standards. However, due to the dynamic network topology and the lack of fixed infrastructure the dissemination of presence information in MANET poses a challenging research problem.

A P2P solution may considerably enrich the IM service. Depending on the delay tolerance of the application, the instant message may be forwarded on the P2P overlay to proxies even if the source has no up to date information of destination "presence".

As the wireless, mobile network structures grow large (e.g., battlefield, urban vehicular grid, etc), different applications may emerge with different customized routing requirements, say. To support these diversified requirements, it may be cost-effective to maintain multiple P2P and overlay networks above the same basic routing and transport architecture. Moreover, some MANETs may grow as an "opportunistic" extension of the wired Internet. In this case, some of the opportunistic ad hoc network users will want to participate in Internet applications already supported by a P2P overlay in the Internet (eg, games, file sharing). This will again create the need to extend the P2P concept to wireless.

2. The Dagstuhl Workshop

The past five years have seen researchers in mobile ad hoc networking and peer-to-peer systems for the wired Internet to investigate their fields separately without considering investigating various cross-cutting issues. This workshop addressed the emergence of "P2P architectures" in MANETs, the applicability and transfer of the wired P2P models and techniques to the wireless scenarios and more generally the nature of the research problems emerging in the ad hoc P2P area. As part of this process, in April 2005, 34 researchers met at Schloß Dagstuhl, Germany (http://www.dagstuhl.de) to examine the current state of affairs with respect to P2P systems and MANETs. Throughout the workshop, we focused on several key issues:

- Which methods, tools and results known for wired P2P systems can be adopted for developing and deploying P2P applications for ad hoc networked systems?
- What kinds of new methods, tools and results are needed for developing and deploying P2P applications for ad hoc networked systems?
- What are the current research challenges for P2P MANETs?

The authors were organizers of the Dagstuhl workshop. Workshop participants were: Ian Akyildiz, Victor Bahl, Suman Banerjee, Christian Bettstetter, Ernst Biersack, Miguel Castro, Peter Druschel, Joe Evans, Kevin Fall, Alain Gefflaut, Mario Gerla, Andreas Haeberlen, Wolfgang Kellerer, Anne-Marie Kermarrec, Eng-Keong Lua, Michela Meo, Petri Mähönen, Maria Papadopouli, Giovanni Pau, Michael Parker, Marcello Pias, Ansley Post, Milena Radenkovic, Ant Rowstron, Ralf Steinmetz, Maarten van Steen, Oliver Waldhorst, Ben Zhao, Martina Zitterbart.

The format of the workshop consisted of intensive small group discussions during which the following three paper sections were developed, interspersed with meetings in plenary session to take stock and chart direction. As a result, workshop participants contributed to just one of the following paper sections, and attribution is indicated under individual section headings.

Work in P2P MANET is diverse and touches on various aspects of computer science and communication engineering. The organization of P2P MANET that emerged during the workshop was along six system building blocks and for four application scenarios as illustrated in Figure 1. Throughout the workshop we focused on three system building blocks for person-to-person MANETs: *Modeling and performance evaluation, Lookup services and key-based routing*, and *Reliable data delivery*. The three subsequent sections report the discussion results on these subareas.

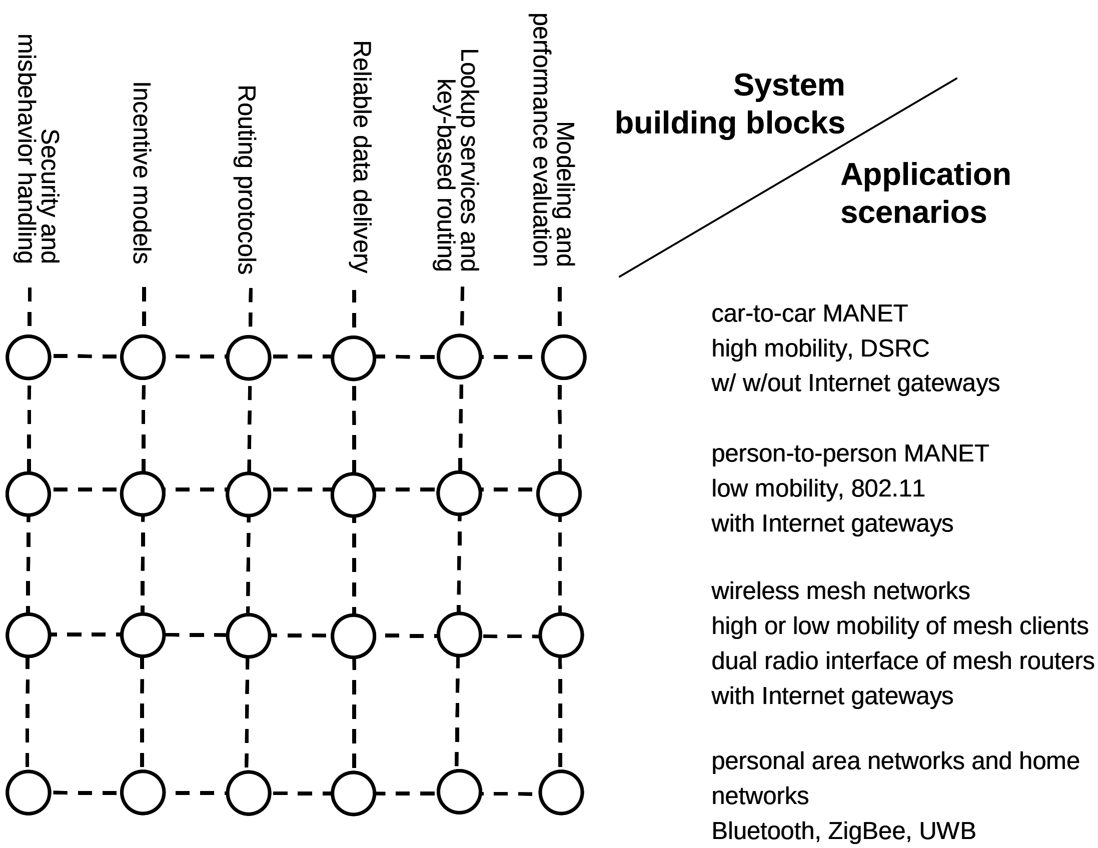

Figure 1: System building blocks and application scenarios of P2P MANET

3. Lookup Services and Key-based Routing in P2P MANETs

Peter Druschel, Miguel Castro, Joe Evans, Alain Gefflaut, Mario Gerla, Andreas Haeberlen, Wolfgang Kellerer, Anne-Marie Kermarrec, Giovanni Pau, Ansley Post, Milena Radenkovic, Ant Rowstron, Ralf Steinmetz, Oliver Waldhorst, Ben Zhao, Martina Zitterbart

In MANET, lookup services can be implemented using either unstructured or structured overlay networks. However, such approaches put some requirements on the MANET environment: (1) The MANET must provide a high degree of connectivity such that a given node can contact each other node at any time with high probability. (2) The nodes in the MANET must exhibit low mobility in order to minimize the required number of updates of routing tables and other structures. In this section, we discuss environmental parameters of lookup services, key-based search, epidemic data dissemination, and complex queries. Finally, research challenges are outlined.

Environmental Parameters

There are several environmental parameters to consider when designing lookup protocols for MANET. Some of this parameters address the application requirements and therefore are similar to those of wired scenarios, some other are peculiar of wireless environments. In particular, we identified the following as major players for a successful design.

Application Related Parameters:

1. Query rate, indicates the amount and distribution of queries in the network; key factors are the statistical characteristics, such as average, standard deviation, median, etc., defining query frequency with a great detail. This parameter influences the workload of the network and impacts the data traffic to be serviced [AC+96].
2. Replication Rate, defines the probability to find a given object in a network [KRR+01].
3. Popularity, defines the query statistics for a particular object. It is worth noting that this is different from the general query rate defined above [KRR+01].
4. Scale of objects, defines the size and the statistics of the object population. In particular it defines the average number of objects per node and its variance. It is worth noting that this parameter may justify a “DHT” based approach even for small networks.

5. Quality of service requirements: delivery ratio and latency, routing consistency for key-based and complex lookups.

Network Related Parameters:

1. Mobility scenario, defines the speed, obstacles, and propagation models. In this field a great contribution would be inferring a realistic scenario from real applications. The mobility is key to determine the success of any MANET application [CB+02] [HG+99].
2. Scale of nodes, defines the size of the join/leave/failure statistics of the participating nodes.
3. Network partitioning and merging is related to the likelihood and extent of network partitions.
4. Network density, defines the average number of nodes per square meter in an Ad hoc network [JL+99] [Bet02].
5. Pure wireless vs. heterogeneous wireless/wired. Any solutions should be suitable for both hybrid networks with ad hoc and infrastructure nodes (i.e., the urban grid) or fully MANET. The latter may introduce additional challenges.
6. Security: Node anonymity and privacy may be required as well as traffic privacy. In particular the network may be required to be resilient at destroying contents or shutdown the network itself. Is worth noting that the security requirements greatly impact the design choice; moreover the network needs to be protected from attacks directed to the overlay structure as well as attacks directed to the underlying technology [ZH01][SM+02][CD+02].

Key based Search

The ability to find an object given a key is useful in several applications. For example, cooperative caching [IRD02], file sharing [Gnu], serverless instant messaging and email [MP+03], voice and video communication, and others. We discuss two basic approaches to implement key lookups in wireless systems.

Unstructured approach

Unstructured overlays organize nodes into a topology that is independent of the keys of the objects that they store. They support queries by performing floods or random walks on the overlay.

Strict layering of unstructured overlay approaches on top of wireless routing protocols is unlikely to work in MANETs. This would require proactively maintaining overlay links that would be multi-hop in the physical network. Keeping the network connected would require a significant amount of traffic to detect and repair failed overlay links. Sharing information across layers may enable layering. For example, when using link-state routing (e.g., OLSR) information about link failures can be used to detect virtual neighbor failures with the need for explicit probe traffic [CJ+01], [KLW03],.

A better approach is to **explore the graph defined by the physical topology and apply the flooding and random walk techniques on this graph**. For example, one could use a DSR- or AODV-type approach that broadcasts to find routes to a particular key identifying an object. Another approach constitutes to replicate the keys of objects stored by a node in its k-hop neighbors and use a DSR- or AODV-type approach to search outside this range. This is similar to ZRP [HP+01] and one hop replication in GIA [CR+03]. We do not know of any previous work that explored random walks to lookup objects in a wireless environment, but many approaches have explored random walks in the wired setting, see for example, GIA [CR+03].

To work effectively, unstructured approaches put certain constrains on the application scenarios. For example, a flooding operation requires media access at each node. Thus, flooding is most effective if both the number of nodes or the number of queries is moderate. Furthermore, a routing-based approach requires that an object is in the same network partition as the node that sends a query for the object in order to be successfully discovered. In scenarios where node density is low and network partitions frequently, replication-based approaches are preferable to routing-based approaches. The number of objects in the system is crucial for both routing-based and replication-based approaches. Storing a network route to each object requires large routing tables. Replicating objects requires a huge amount of storage at the nodes. However, both types of approaches benefit from locality in the query popularity distribution, since it reduces the number of active routes and the size of the "hot set" of objects that must be replicated.

Structured approach

Structured overlays assign identifiers to nodes and constrain the identifiers of overlay neighbors and the keys of the objects that they store. They can locate object efficiently without flooding, for example, lookups can cost O(logN) with O(logN) state per node.

Strict layering of structured overlay approaches on top of wireless routing protocols is unlikely to work for the same reasons mentioned for unstructured overlays. Sharing information across layers may enable layering as discussed above.

Another approach is to use a different routing protocol that supports key lookups directly. There are to possible approaches to do this: topology-dependent structuring and topology-independent structuring. Both can be extended to hybrid wired/wireless environments.

Topology-dependent structuring assigns topology dependent identifiers to nodes and stores objects at the node whose identifier is **closest** to the object’s key under some definition of distance.

A possible approach is to use geographical routing techniques, for example, GPSR [KK00], BVR [FR+05], or GEM [NS03]. This approach was first proposed in the geographical hash table (GHT). GHT is built on top of GPSR. Its keys are coordinates in the geographical space and information about the object is replicated in the nodes whose coordinates are closest to the key.

A second approach is to use landmark routing, for example, LANMAR [PGH00], L+[CM02].

With **Topology-independent structuring**, nodes are assigned identifiers that are independent of the topology and nodes must maintain paths between overlay neighbors. Objects are stored at the nodes whose identifiers are closest to their keys. Virtual Ring Routing (VRR) is an example of this approach.

The structured overlay based approaches (other than strict layering) scale well with the number of nodes and the query rate. For example, the approaches based on geographical routing require routing state proportional to the number of physical neighbors of a node, Strata uses routing state proportional to the number of levels in the landmark hierarchy (O(logN)), and VRR maintains O(sqrt(n)) routing state per node. All the approaches can complete lookups in O(sqrt(n)) steps but paths in VRR will be slightly longer than paths in the topology-dependent structuring approaches.

A large number of objects and or large objects may be an issue for topology-dependent structuring approaches with mobility. The issue is that when nodes move their coordinates change and they may no longer be the nodes with coordinates closest to the keys of some of the objects they store. This requires objects to be transferred to other nodes. With topology-independent structuring, objects only need to be transferred between nodes when nodes fail or new nodes arrive.

Structured overlay approaches will outperform unstructured approaches when the number of nodes, the number of objects, or the query rate increases. Caching information about objects can improve the scalability of unstructured approaches when the number of objects is small or the object popularity distribution is very skewed, for example, zipf or zipf like.

Epidemic Data Dissemination

Typically, both structured and unstructured approaches will perform poorly in scenarios with low connectivity and high mobility. However, the mobility intrinsic in MANET can also be exploited to efficiently distribute and locate resources across the network by “epidemic dissemination” of information. Grossglauser and Tse provided the theoretical foundation for such approaches. They presented a formal proof and simulation results showing that mobility increases the per-session throughput in a MANET [GT02]. Papadopouli and Schulzrinne introduced seven degrees of separation (7DS), a system for P2P Web document sharing between mobile users based on epidemic dissemination [PS01]. To locate a Web document, a 7DS node broadcasts a query message to all mobile nodes currently located inside its radio coverage. Recipients of the query send response messages containing file descriptors of matching Web documents stored in their local file caches. Subsequently, such documents can be downloaded with HTTP by the inquiring mobile node. Downloaded Web documents may be distributed to other nodes that move into radio coverage, implementing an epidemic dissemination of information.

Lindemann and Waldhorst showed how to apply epidemic dissemination to implement a distributed directory and introduced an analytical performance modeling approach [LW04], [LW05]. They presented a general-purpose distributed lookup service for mobile applications, denoted Passive Distributed Indexing (PDI). PDI stores index entries in form of (key, value) pairs in index caches maintained by each mobile device. Index caches are filled by epidemic dissemination of popular index entries. By exploiting node mobility, PDI can resolve most queries locally without sending messages outside the radio coverage of the inquiring node. To foster data dissemination, PDI introduces a bandwidth-efficient mechanism for message relaying denoted selective forwarding. For keeping index caches coherent, configurable value timeouts implement implicit invalidation of index entries and lazy invalidation caches implement explicit invalidation of index entries by epidemic distributions of invalidation messages.

Further applications of epidemic dissemination include content delivery and the dissemination of presence information in mobile instant messaging systems. The epidemic dissemination mechanism is a novel and powerful form of P2P cooperation which is obviously not available in the wired Internet.

Complex Queries

In this section, we consider the issue of supporting complex queries in wired networks and MANETs. Complex queries include any combination of queries on multiple attributes and range queries with optional preference-based rankings. Specific examples of applications include searching for services or resources in the network, e.g. find the closest uncongested gateway to the Internet, locate a color-capable printer within my security domain, etc.

The service discovery problem on wired networks is well studied. Numerous existing approaches (SLP, SSDS [HC+02], Condor ClassAds [RLS98]) have addressed various combinations of issues including: expressive queries, authentication and access control, network locality, and scalability in number of queries and clients. In the context of structured peer-to-peer networks, resolving fully generalized queries required additional constructs on top of the single key-based routing functionality. Existing efforts include pSearch [TXD03] based on information retrieval techniques, Mercury [BAS04] based on data partitioning on primary fields, and range queries on CONE [SG+04].

A number of approaches address the service discovery problem in wireless networks. The most simplistic approach is to flood a query to all nodes in the wireless network. This approach can be effective and responsive under highly mobile environments, but incurs a high bandwidth/energy cost per query. Other approaches rely on probabilistic approaches to routing queries towards service descriptions. This can be the form of a random walk inside the network, or a directional flood of both queries and descriptions in approximately N, S, E, W directions [TV04]. Finally, nodes can embed encodings of service descriptions alongside their routing information at the routing layer. Nodes discover service descriptions as part of the route discovery process, and can cache them for future lookups.

General rich query is a challenge for both the wired peer-to-peer and wireless network contexts. If P2P overlays can be maintained on top of highly mobile wireless networks, then layering additional rich query functionality on top can be a promising solution to solving service discovery on wireless networks. For highly mobile environments, unstructured approaches might be preferred over approaches that rely more on cached state. Under heavy query load, however, stateful approaches can significantly reduce the amount of traffic and stress on the underlying network.

Research Challenges

Most of the research in area of lookup and key-based routing is still in its early stages, so there are numerous challenges:

- Is structure required, and if yes, how much of it? Do general structured approaches (like DHTs) provide the right abstraction for a large class of MANET applications under different environmental parameters? Do we need specialized solutions?
- Can applications that are currently implemented using structured and unstructured overlays be directly used on MANETs? Is the KBR interface for structured overlays appropriate of the MANET based structured overlays?
- If routing state is maintained in the network how does this scale under more extreme environmental parameters, especially high mobility?
- In scenarios where wired and wireless infrastructure is integrated, is this exposed to the application, and if so to what degree?
- How do we evaluate the different approaches under realistic environmental parameters and workloads?

4. Reliable Data Delivery in P2P MANETs

Kevin Fall, Ian Akyildiz, Victor Bahl, Ernst Biersack, Michael Parker, Marcelo Pias, Milena Radenkovic, Ben Zhao

We consider the application of network overlay technologies in MANETs. The compelling MANET applications appear to be communications among and between nodes where essentially no pre-planned infrastructure is present. This includes military (battlefield) units, emergency response scenarios, and exploration in dangerous or unknown areas. Other applications may exist, although their need for MANET support is less compelling, and many other applications that have been proposed appear to be supportable at reasonable cost with various types of pre-planned infrastructure.

Each of the compelling MANET scenarios mentioned above will have particular requirements as to its need for reliability and timeliness of data delivery. We do not attempt an encyclopedic accounting of each of these, but rather explore the common design issues that arise.

Defining Reliable Data Delivery

An application's requirement for reliable data transport can be defined as a constraint set or tolerance threshold upon the granularity, scale, and timeliness of data delivery. The scale here refers to the number of destinations intended to receive a particular data item. The threshold is the application's relative tolerance to errors made visible to it. A key aspect of this is the requirement (or lack thereof) on the timeliness of delivery. In particular, if timely (low-latency) delivery data is required, the application may be forced to be more tolerant of errors that it must handle itself. For those applications tolerant to delay, the network has more flexibility (and time) to improve its data delivery reliability for the application.

We note that an application's requirements on reliability (and delay) may vary over time. For example, consider an application supporting voice conversations. In some circumstances (e.g. heated debate), interactivity and timeliness of delivery may be of paramount importance. Conversely, when the speaker changes relatively infrequently, enhanced reliability may be desired to avoid drop-outs or other problems. As a conversation progresses and enters and exits various phases, it may be important to vary the application's relative performance preferences.

Mechanisms to Enhance and Enforce Reliable Data Delivery

There are many mechanisms available to network protocol designers to enhance an application's perceived reliability beyond the raw performance of the underlying network links. Examples of these include both pro- and reactive repair mechanisms. Among the proactive mechanisms, forward error correction [BM+01] and intentional replication (possibly in combination) may be used [JD+05]. These procedures may be executed anywhere along the delivery path in the network and potentially by any layer, from physical to overlay. Reactive mechanisms generally operate based upon either an explicit or implicit signal from the network. For mobile networks, signaling may be somewhat problematic. If signaling round-trip-time exceeds node handoff time, control signaling may be sufficiently delayed as to become useless.

The role of overlay nodes in enhancing reliability can be important [AB+02]. While traditional network layer (IP) routers are equipped with packet storage, this storage is transient and generally not accessible to applications (i.e. it is not named). End systems experience the effect of this storage only in statistical multiplexing along a network path. Overlay nodes could be used to enhance reliability by providing storage that may be both accessible to applications and persistent. In cases where disconnection or disruption is common, the ability to buffer data within the network for significant periods of time becomes one of the few critical tools that can be used to provide high reliability [Fal03].

Reliability is 'implemented' by some component of the overall network and its end systems. Today, end-to-end reliability is implemented by TCP with the traditional *fate sharing* argument [Car96]. In effect, if either of the endpoints of a connection are destroyed, it is natural to allow the connection to be cleared. For high reliability in networks where overlay nodes may be present, conjoining operation of the network connection with operation of the end nodes may not be the most effective approach. Overlay nodes may provide the ability to 'capture' the responsibility for ensuring reliable delivery. Such an idea amounts to transferring *custody* of a message or packet to another entity which could be in the overlay.

Trade-Offs

Improving reliability for network communication generally comes at the price of degraded performance in some other performance aspect. Performance measures include delay, throughput, fairness, power and storage utilization. These tradeoffs can be broken down into two major categories: those that affect the pairwise communication of two entities, and those that affect the overall operation of a group of users sharing a communication infrastructure. Groups are affected generally when shared resources are consumed in enhancing reliability; This includes partitioning of throughput, fairness and storage utilization. When additional power is used (e.g. MANET case), this can also affect the noise floor and fairness with respect to other participants.

Impact of Overlays

An overlay can be defined as a subset of nodes in a network that form another network. The overlay nodes may perform different processing functions from the underlying network. There are a rich set of examples of overlay networks, including lookup services (DNS, DHTs such as Pastry [RD01], Chord [SM+01], and CAN [RF+01]), experimental routing systems such as RON [AB+02], store-and-forward capabilities (e-mail, network news, DTN), group communication (CDNs, network news), protocol translators (transcoders), and distributed storage/file sharing (Gnutella [Gnu], Oceanstore, CFS). While some of these services provide services not available with the traditional infrastructure, others effectively circumvent the normal operation of the infrastructure to provide optimization based upon some other metric than the infrastructure ordinarily optimizes. This is often undertaken because the infrastructure itself is not easily modified and evolved.

When considering the potential uses of an overlay for enhancing the reliability of MANETs, we are faced with the question of what responsibilities we might assign to overlay nodes. For purposes of discussion, we consider the case of multiple MANET not in direct communication with each other that may be aided by the presence of a comparatively reliable packet switched infrastructure (e.g. wired Internet). We now explore the types of questions that must be answered.

Which nodes are members of the overlay?

Given that not all nodes in a network may be members of the overlay, where should those nodes participating in the overlay be placed? Today, nodes in an overlay are typically configured, at least initially, by hand. It is easy to imagine nodes that dynamically and automatically decide whether to participate in an overlay, and this issue may be especially important for MANETs because overlay participation may be dictated by topological location, and with mobile nodes topological position is always in flux. Note that other (e.g. physical) constraints may drive the decision to participate in the overlay. For example, nodes with limited power may not wish to act as overlay routers for other nodes.

A related question is the initialization and bootstrap problem, which is to gather initial information about the operating environment. This includes handling leaves and joins of other nodes, neighbor discovery, topology discovery, and possibly other issues such as address assignment, etc.

What services do overlay nodes provide?

Taking account of the popular types of overlays mentioned above, we observe that the primary services are lookups, dynamic routing, and storage. In some cases, more than one of these capabilities are combined (e.g. DHTs). For enhancing the reliability of MANETs, we first observe that disconnection needs to be addressed. It is unrealistic to assume that a MANET node will *always* be in communication with one or more other nodes. More precisely, however, is that the duration of disconnection may be (qualitatively) short or long. Short-term disruptions can typically be addressed by modifications of the transport or lower layers (e.g. Freeze-TCP) [GM+00], while longer term disruptions (say, more than 30 sec) often induce effects at the application layer.

For short-term interruptions in MANETs, the role of overlay nodes appears to mimic that of nodes in conventional non-MANET networks. That is, they may provide similar types of services to those they are currently providing on conventional Internet-type networks. For long-term interruptions, the role of overlay nodes appears to be more significant. In particular, overlay nodes can offer significant storage capabilities to enhance the reliability of end-to-end delivery. When a node quickly enters a MANET, offloads its data, and exits ('dump and jump'), an overlay node persistently located in the MANET could cache the data, providing a form of reliable delivery to other members of the MANET. In addition, when communication between a MANET and an infrastructure node is necessary, the same situation applies: a particular MANET node with sufficient storage can be tasked to communicate with the infrastructure and 'ferry' the data [ZM+04] originating with the MANET (and vice versa).

Burning Questions

This section considers more directly the role of overlays as a mechanism for enhancing the reliability of MANETs and/or combination MANET/infrastructure systems. To conclude, we would like to pose the two 'burning' questions:

- **Will the ubiquitous deployment of either individual or large-scale MANETs happen?** In cases where MANETs make sense, we believe the most powerful capability overlays offer is their storage, which can be used to help in combating periods of disconnection. In particular, if data can be cached or deposited at *specific* nodes in the network (i.e. in infrastructure nodes), overall reliability may be significantly enhanced. However, this scenario implies an ability to place responsibility (e.g. for buffering) at specific locations in the network, and is inherently asymmetrical. It relates closely to cluster-head and landmark types of routing schemes (e.g. Brocade [ZD+02]).
- **Will overlay systems such as DHTs be employed in small-scale environments such individual MANETs?** We believe that other mechanisms (e.g. lower-layer broadcast) might be a more effective and/or less costly alternative to the use of realiability-enhancing overlays for small-scale MANETs. Note that in discussions of this paper we also considered the arguments supporting visions of large-scale sensor networks, and in most application scenarios for these we find sensor nodes are primarily non-mobile and those that are mobile are frequently proximal to deployable infrastructure and do not appear to require most of the mechanisms forming the basis of MANET designs.

5. Modeling and Performance Evaluation of P2P MANET

Anthony Joseph, Suman Banerjee, Christian Bettstetter, Eng-Keong Lua, Michela Meo, Petri Mähönen, Maria Papadopouli, Maarten van Steen

It is important to recognize that P2P MANET research is not simply about Distributed Hash Tables, rather, it is about decentralized resource management, routing, and control in an overlay network with potentially no single service provider. Also, while the issues we discuss are present in wired networks, we consider them in the specific domain of P2P MANET, which has several unique characteristics relative to wired networks. In particular, power and network resources are constrained in a P2P MANET, while quality of service may vary dramatically depending on the underlying wireless network links and user/node mobility. Also, network proximity plays an important role, as arbitrary routing between two nodes (i.e., direct network connectivity) may not be present.

In this section, we examine the challenges of collecting real-world datasets of network topology and user/node mobility and activity, and using the datasets to create larger scale accurate synthetic ones. We also consider the challenges for performing realistic publication-quality research simulations.

Benchmark Datasets for P2P MANET

Currently, there is a dearth of publicly available benchmark datasets of user/node mobility application usage, and network topology [ABQ02] [BC+03], [BV+02], [HK+04]. Such datasets are critical to enabling repeatable and comparable research in P2P MANETs. However, given the sensitive user information or proprietary network information that may be present in these datasets, there are many challenges: academics are reluctant to expend the time and energy to sanitize the datasets (*e.g.* US academics must comply with FERPA regulations); and companies are often reluctant to disclose information they consider proprietary.

As an example, we consider the characterization the query behavior of peers in a peer-to-peer (P2P) file sharing system [KL+04]. The workload measures include the fraction of connected sessions that are passive (i.e., issue no queries), the duration of such sessions, and for each active session, the number of queries issued, time until first query, query interarrival time, time after last query, and distribution of query popularity. Moreover, the key correlations in these workload measures are captured in the form of conditional distributions, such that the correlations can be accurately reproduced in a synthetic workload. The characterization is based on trace data gathered in the Gnutella P2P system [Gnu] over a period of 40 days.

To enable better research validation and comparison through repeatability, we propose that the scientific community agrees on a set of benchmarks that should be used whenever evaluating new protocols, solutions, algorithms, etc. Such benchmark datasets should include: simple synthetic models, models derived from analyzing real-world data, and combinations of both real data and synthetic models. These models must be defined in terms of different time, spatial, and number of user scales. Also, to enable the development and usage of common tools for experiments and data analysis, each model should use a common standard data representation.

Depending upon a model's intended usage, each model might include some combination of the following non-exhaustive list of elements:

- User/node mobility [BC+03], [CB+02]
- User, node, device, and application types []ABQ02], [AEF+03]
- Network topology [BV+02]
- Traffic characteristics [MW+04], [HK+04]
- Device availability
- Link quality
- Distribution of devices (mobile and infrastructure devices)
- Device capabilities (*e.g.*, energy, memory, transmission, interfaces, etc.)
- Measured application goodput (instead of or in addition to network throughput) [HK+04], [MW+04]
- Possible inputs from the environment
- Wired network characteristics (*e.g.*, access point/node locations, backbone topology, bandwidth and latency distributions, etc.)

The first step in benchmark model creation is collecting the necessary trace data. The current state of the art in tools includes running SNMP gathers, tcpdump processes, and syslog collectors. However, all have problems. For example, there are substantial complexities associated with using SNMP: gathering data from hundreds of wireless access points, dealing with the fact that different vendor's access points use different, and often, inconsistent MIBs, working around buggy SNMP implementations, and accounting for the lack of fine-grain data sampling. Also, time must be synchronized across the data collection tools. The state of the art would be advanced, if there were a common, systematic way of using these tools.

Another problem with the data collection process is that the underlying devices often provide inconsistent or incomplete information across devices types and manufacturers (e.g., many devices provide pseudo SNR values or inconsistent transmitter power). The values reported by devices often require normalization to account for differences between types and manufacturers (e.g., transmit power).

We also need good Quality of Service measurement tools. An open research challenge is how researchers can identify potential systematic errors in their data collection methodologies.

To properly use these benchmarks, researchers should perform sensitivity analysis of the model input parameters and evaluate the statistical confidence of outputs. The research publications that they produce should include these analyses and evaluations to convince readers that the researchers have adequately validated the likely performance of their applications under real-world conditions and scale.

To create models derived from real data, researchers should appropriately sanitize the dataset by performing anonymization, and if necessary, remove outliers and side effects. They should make clear what assumptions they have made and perform appropriate cross-validation against other experimental results. We believe that developing benchmark models is a serious scientific effort that deserves researchers' attention.

Once a researcher creates a benchmark model, they should make the benchmark model available to the community, preferably in a common dataset data base (*e.g.*, Dagstuhl Benchmarks). Ideally, researchers should gather feedback from the community and release updated benchmark models. This process could be encouraged by a workshop dedicated to presenting and evaluating benchmark proposals.

In addition to benchmark models, the benchmark database could include well-documented real and synthetic data traces, along with monitoring and analysis tools.

Challenges for the Simulation of P2P MANET

We argue that the state of the art in P2P MANET simulation is far from perfect or even adequate. Although the scalability and robustness of different simulation tools have been enhanced to a remarkable degree, much remains to be done. The issues that require clear and focused attention from the community include validation of simulators and the models inside of them. We believe that publishing venues must raise the bar for acceptance and reject work that relies solely on small-scale simulation without consideration of sensitivity analysis.

Currently the situation is very fragmented and many simulation models are often clearly used (due to ignorance or sheer laziness) outside of their validity areas:

- **Validation of simulation claims at a reasonable level should become the accepted state of the art:** While it is clear that very large-scale systems can only be simulated, we find it worrisome that these models have not been proven or validated, even in the small-scale situations that are clearly achievable with testbeds.
- **Standardized set of experimental scenarios, similar to the database communities' TPC benchmarks:** The practice in the field should aim at towards more robust analysis of reliability and sensitivity of simulations results as a function of simulation parameters. Ideally, such sensitivity analysis functions should be embedded as supporting tools in simulators themselves. A somewhat straightforward, but big step towards a better situation, would be encourage people to run larger sets of different input cases (*e.g.*, user/node mobility, varying topologies, different types of applications) to test the validity and sensitivity of their claims. This change in evaluation methodology also includes researchers making publicly available their synthetic and measured data sets for third-party testing of their assumptions. These stepswould better enable more direct comparisons of results between different researchers.
- **User/node mobility, network topology, and wireless modelling:** Although there has been remarkable progress in mobility modeling and wired topology generation, significant work remains. Many results are not comparable or even reliable, since no common set of measurements based data, or robust simulation "plug-ins" exist. One of the problems is that complex mobility and topology models are rich sub-fields of their own expertise. There should be tools and methods for others to safely use the models of these sub-fields through standard simulators. The groups doing simulations should also be more aware of the limitations, and appropriate warnings should be attached to the data. We refer as an example the fact that although some topology models are more or less validated as topology, the link characteristic such as delay might be completely unrealistic, if taken automatically from those topology generators. Another aspect is that the level of reality in the case of wireless systems simulation is sometimes very bad, and more interdisciplinary work is required to ensure that the state of the art is guaranteed at the appropriate level.

- **Scaling properties of simulators:** The current work has been mostly focused on optimization of simulator tools to provide very large number of nodes, and enough speed (sometimes through parallel simulations) to tackle those large-scale problems. However, this work has not fully addressed the question of scaling. It is not clear that a simple 20 node simulation can be "stretched" to 10,000 node simulations by a "copy-and-paste" methodology. We believe that new avenues of large-scale simulation and modeling should be probed. As examples we mention the agent modeling-based simulations used for highway traffic modeling and the ultra-large scale Monte Carlo simulations used in the area of condensed matter physics.

References

[AC+96] S. Adali, K. S. Candan, Y. Papakonstantinou, and V. S. Subrahmanian, Query Caching and Optimization in Distributed Mediator Systems, *Proc. Int Conf. on Management of Data (ACM SIGMOD)*, Montreal, Quebec, Canada, 1996.

[ABQ02] A. Adya, P. Bahl, and L.Qiu, Characterizing and Analyzing Alert and Browse Services of Mobile Clients, *Proc. USENIX Annual Technical Conference*, Monterrey CA, 2002.

[AB+02] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, Resilient Overlay Networks, *Proc. 5th Int. Symp. on Operating System Principles*, *(ACM SOSP)* 2002.

[AEF+03] F. Anjum, M. Elaoud, D. Famolari, A. Ghosh, R. Vaidyanathan, A. Dutta, P. Agrawa, T. Kodama, and Y. Katsube, Voice Performance in WLAN Networks: An Experimental Study, *Proc. IEEE Globecom*, San Francisco, California, Dec. 2003.

[BV+02] A. Balachandran, G.M. Voelker, P. Bahl, and P.V. Rangan, Characterizing User Behavior and Network Performance in a Public Wireless LAN, *Proc. Int. Conf. on Measurement and Modeling of Computer Systems (ACM SIGMETRICS)*, Marina Del Rey, CA, 2002.

[BC+03] M. Balazinska, and P. Castro, Characterizing Mobility and Network Usage in a Corporate Wireless Local-Area Network, *Proc. 1st Int. Conf. on Mobile Systems, Applications, and Services (ACM MobiSys)*, San Francisco CA, 2003.

[Bet02] C. Bettstetter, On the Minimum Node Degree and Connectivity of a Wireless Multihop Network, *Proc. Int. Int. Symp. on Mobile Ad Hoc Networking and Computing (ACM MobiHoc 02)*, Lausanne Switzerland, 2002.

[BAS04] A. Bharambe, M. Agrawal and S. Seshan, Mercury: Supporting Scalable Multi-Attribute Range Queries, *Proc. Proc. Int. Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM)*, Portland, OR, 2004.

[BM+01] J. Byers, M. Luby, and M. Mitzenmacher, A Digital Fountain Approach to Asynchronous Reliable Multicast, *IEEE Journal Selected Areas in Comm.*, *Special Issue on Network Support for Multicast Communication*, **20**, 1528-1540, 2002.

[CB+02] T Camp, J. Boleng, V Davies, and C. Golden, A Survey of Mobility Models for Ad Hoc Network Research, *Wireless Communication and Mobile Computing*, **2**, 2002.

[Car96] B. Carpenter, (Ed.) Architectural Principles of the Internet, *IETF Request for Comments RFC 1958*, June 1996.

[CD+02] M Castro, P Druschel, A Ganesh, and A Rowstron, Security for Structured Peer-to-peer Overlay Networks, *Proc. 5th Symp. on Operating Systems Principles* (*ACM SOSP*), 2002.

[CR+03] Y. Chawathe, S. Ratnasamy, L. Breslau, N. Lanham, and S. Shenker, Making Gnutella-like P2P Systems Scalable, *Proc. Proc. Int. Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM)*, Karlsruhe Germany, 2003.

[CM02] B. Chen and R. Morris, L+: Scalable Landmark Routing and Address Lookup for Multi-hop Wireless Networks, *Technical Report MIT-LCS-TR-837, Massechusetts Institute for Technology, 2002.*

[CLP+04] F. Chinchilla, M. Lindsey, and M. Papadopouli, Analysis of Wireless Information Locality and Association Patterns in a Campus, *Proc. IEEE Infocom,* Hong Kong, 2004.

[CJ+01] T. Clausen and P. Jacquet, Optimized Link State Routing Protocol, IETF RFC 3626, http://www.ietf.org/rfc/rfc3626.txt

[Fal03] K. Fall, A Delay Tolerant Networking Architecture for Challenged Internets, *Proc. Int. Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM),* Karlsruhe Germany, 2003.

[FR+05] R. Fonseca, S. Ratnasamy, J. Zhao, C.T. Ee, D. Culler, S. Shenker, and I. Stoica, Beacon Vector Routing: Scalable Point-to-point Routing in Wireless Sensor Nets, *Proc. 2nd Symp. on Networked Systems Design and Implementation (ACM NSDI)*, Boston, MA, 2005.

[GT02] M. Grossglauser and D. Tse, Mobility Increases the Capacity of Ad-hoc Wireless Networks, *IEEE/ACM Trans. on Networking* **10**, 477-486, 2002.

[JL+99] P. Johansson, T. Larsson, N. Hedman B. Mielczarek, and M. Degermark, Scenario-based Performance Analysis of Routing Protocols for Mobile Ad Hoc Networks, *Proc. 5th Int. Conf. on Mobile Computing and Networking (ACM Mobicom)*, Seattle, WA, USA, 1999.

[KRR+01] J. Kangasharju, J. Roberts and K. Ross, Object Replication Strategies in Content Distribution Networks, *Proc. Int. Web Caching and Content Distribution Workshop*, Boston, MA, June 2001.

[KL+04] A. Klemm, C. Lindemann, M. Vernon, and O.P. Waldhorst, Characterizing the Query Behavior in Peer-to-Peer File Sharing Systems, *Proc. 4th ACM Internet Measurement Conference (IMC 2004)*, Taormina, Italy, 2004.

[Gnu] Gnutella Developer Forum, Gnutella – A Protocol for a Revolution, 2003. http://rfc-gnutella.sourceforge.net.

[GM+00] T. Goff, J. Moronski, D. Phatak, and V. Gupta, Freeze-TCP: A True End-to-End TCP Enhancement Mechanism for Mobile Environments, *Proc. Infocom* 2000.

[HP+01] Z.J. Haas and M.R. Pearlman, The Performance of Query Control Schemes for the Zone Routing Protocol, *IEEE/ACM Trans. Networking*, **9**, 2001.

[HK+04] T. Henderson, D. Kotz, and I. Abyzov, The Changing Usage of a Mature Campus Wide Wireless Network, Proc. *Int. Conf. on Mobile Computing and Networking (ACM Mobicom)*, Philadelphia PA, 2004.

[HC+02] T.D. Hodes, S.E. Czerwinski, B.Y. Zhao, A.D. Joseph and R.H. Katz, An Architecture for Secure Wide-Area Service Discovery, *Wireless Networks, Selected Papers from MobiCom 1999*, **8**, 213-230, 2002.

[HG+99] X. Hong, M. Gerla, G. Pei, and C. Chiang, A Group Mobility Model for Ad Hoc Wireless Networks, *Proc ACM/IEEE MSWiM*, Seattle, WA, pp.53-60, 1999.

[IRD02] S. Iyer, A. Rowstron, and P. Druschel, Squirrel: A Decentralized Peer-to-peer Web Cache, *Proc. Symp. on Principles of Distributed Computing (PODC)*, Monterey, CA, 2002.

[JD+05] S. Jain, M. Demmer, R. Patra, and K. Fall, Using Redundancy to Cope with Failures in a Delay Tolerant Network, *Proc. Int. Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM)*, Philadelphia, PA, 2005.

[KK00] B. Karp and H. Kung, GPSR: Greedy Perimeter Stateless Routing for Wireless Networks, *Mobile Computing and Networking*, 243–254, 2000.

[KLW03] A. Klemm, C. Lindemann, and O.P. Waldhorst, A Special-Purpose Peer-to-Peer File Sharing System for Mobile Ad Hoc Networks, *Proc. IEEE Vehicular Technology Conf. (VTC)*, Orlando, FL, 220-226, October 2003.

[LLW05] C. Lindemann. C. Lambert, and O. Waldhorst, Effective Dissemination of Presence Information in Mobile Ad Hoc Networks, *Technical Report, CS Department, University of Dortmund*, 2005.

[LW03] C. Lindemann and O. Waldhorst, Exploiting Epidemic Data Dissemination for Consistent Lookup Operations in Mobile Applications, *ACM Mobile Computing and Communication Review (MC2R), Special Issue on Mobile Data Management*, **8**, 44-56, 2004.

[LW05] C. Lindemann and O. Waldhorst, Modeling Epidemic Information Dissemination on Mobile Devices with Finite Buffers, *Proc. Int. Conf. on Measurement and Modeling of Computer Systems (ACM SIGMETRICS)*, Banff, CA, 2005.

[MW+04] X. Meng, S. Wong, Y. Yuan, and S. Lu, Characterizing Flows in Large Wireless Data Networks, *Proc. Int. Conf. on Mobile Computing and Networking (ACM Mobicom)*, Philadelphia PA, 174–186, 2004.

[MP+03] A. Mislove, A. Post, C. Reis, P. Willmann, P. Druschel, D.S. Wallach, X. Bonnaire, P. Sens, J.-M. Busca, L. Arantes-Bezerra, POST: A Secure, Resilient, Cooperative Messaging System, *Proc. 9th Workshop on Hot Topics in Operating Systems (HotOS)*, May Lihue, HI, 2003.

[NS03] J. Newsome and D. Song, GEM: Graph eMbedding for Routing and Data-centric Storage in Sensor Networks without Geographic Information, *Proc. 1st Int. Symp. on Networked Sensor Systems (ACM SenSys)*, 2003.

[PGH00] G. Pei, M. Gerla, and X. Hong, LANMAR: Landmark Routing for large scale Wireless Ad Hoc Networks with Group Mobility, *Proc. 6th Int. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Stanford, CA, 2000.

[PG+05] J.A. Pouwelse, P. Garbacki, D.H.J. Epema, H.J. Sips, The BitTorrent P2P File-Sharing System: Measurements and Analysis, *Proc. 4rd Int. Workshop on Peer-to-Peer Systems (IPTPS)*, Ithaca NY, LNCS Springer 2005.

[RD01] A. Rowstron and P. Druschel, Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems, *Proc. IFIP/ACM Int. Conf. on Distributed Systems Platforms, LNCS 2218*, Heidelberg Germany, Springer 2001.

[RLS98] R. Raman, M. Livny and M. Solomon, Matchmaking: Distributed Resource Management for High Throughput Computing, *Proc. Int Symp. on High Performance Distributed Computing, (HPDC)*, 1998.

[RF+01] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, A Scalable Content-Addressable Network, *Proc. Int. Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM)*, San Diego, CA., 2001.

[RSS01] J. Rosenberg, H. Schulzrinne and B. Suter, Wide Area Network Service Location, *IETF Request for Comments (RFC) (December 1997)* draft-ietf-svrloc-wasrv-01.txt

[SG+04] O.D. Sahin, A. Gupta, D. Agrawal, A. El Abbadi, A Peer-To-Peer Framework for Caching Range Queries. *Proc. Int. Conf. on Data Engineering (ICDE)*, 165-176, Boston, 2004.

[Sha] Sharman Networks Ltd., http://www.kazaa.org

[SM+02] E Sit and R Morris, Security Considerations for Peer-to-Peer Distributed Hash Tables, *Proc. 1st Int. Workshop on Peer-to-Peer Systems (IPTPS)*, 2002.

[SM+01] I. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan, Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications, *Proc. Int. Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM)*, San Diego, CA, 2001.

[TXD03] C. Tang, Z. Xu and S. Dwarkadas, Peer-to-Peer Information Retrieval Using Self-Organizing Semantic Overlay Networks, *Proc. Int. Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communication (ACM SIGCOMM)*, Karlsruhe Germany, 2003.

[TV04] J.B. Tchakarov and N.H. Vaidya, Efficient Content Location in Wireless Ad Hoc Networks, *Proc. IEEE Conf. on Mobile Data Management*, Berkeley, CA, 2004.

[ZD+02] B. Zhao, Y. Duan, L. Huang, A. Joseph, and J. Kubiatowicz, Brocade: Landmark Routing on Overlay Networks, *Proc. 1st Int. Workshop on Peer-to-Peer Systems (IPTPS)*, 2002.

[ZM+04] W. Zhao, M. Ammar, and E. Zegura, A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks, *Proc. Int. Symp. on Mobile Ad Hoc Networking and Computing (ACM MobiHoc)*, Philadelphia, PA, 2004.

[ZH01] L Zhou and Z.J. Haas, Securing Ad Hoc Networks, *IEEE Network Magazine*, 2001.