Performance of data dissemination among mobile devices

Maria Papadopouli[*] Henning Schulzrinne[†]

Abstract

We introduce 7DS a new paradigm of mobile information access that enables mobile devices to share resources in a peer-to-peer and self-organizing manner. There are three facets of cooperation, namely information sharing, message relaying, and bandwidth sharing. 7DS is an architecture, a set of protocols and an implementation enabling the exchange of data among peers that are not necessarily connected to the Internet. Peers can be either mobile or stationary. We present a general framework of mobile data access and classify and model different approaches based on their dependency on an infrastructure of servers and the interaction among the hosts. Via extensive simulations, we find the effect of the wireless coverage range, network size, query mechanism, cooperation strategy among the mobile hosts, and energy conservation with a large number user mobility scenarios. Using theory from random walks and random environments and diffusion of controlled processes, we model one of these data dissemination schemes and show that the analysis confirms the simulation results for this scheme.

1 Introduction

Wireless devices are becoming smaller, more user friendly and more pervasive. They are not only carried by people, but are integrated into physical objects. These devices can be part of data-centric, mobile, ad hoc (without infrastructure) and sensor networks; they collect, measure, process, query, and relay information. The expansion of the Internet and wireless data communications have amplified this trend by making information easier to share and by increasing the amount of information that is shared. People

---

[*] Department of Computer Science, The University of North Carolina at Chapel Hill. This work was done while she was a student at Columbia University.

[†] Department of Computer Science, Columbia University.

access local and general news, traffic or weather reports, sports, maps, guide books, music, video files, and games. Similarly, there is growing interest in the transportation industry to support vehicles with navigation tools and location-based services. We classify mobile information access according to their dependency on an infrastructure into three categories: wireless Internet via base stations, infostations, and peer-to-peer. The first two approaches need an infrastructure. The first approach provides "continuous" wireless Internet access; examples include CDPD, 3G wireless, IEEE 802.11, and two-way pagers [8, 30]. Currently, this wireless access has either sparse coverage, low cost and high speed (IEEE 802.11) or major-cities-only coverage and high cost (Metricom) or wider coverage, but extremely low rates and high costs (CDPD, RIM). In several cities worldwide, nonprofit groups have installed IEEE 802.11b base stations to provide free wireless access to the Internet. The second approach provides information access via infostations. When a wireless device is in close proximity to an infostation, it can query the server and access the information. The infostations can be located at traffic lights, building entrances, cafes and airport lounges. In the simple case, an infostation acts as an autonomous information server with a data repository for wireless devices in its range. The more general case includes an infrastructure of infostations, in which an infostation can be connected to a network of infostations or to the Internet. It can act as a proxy, caching data and forwarding requests to other infostations or to the Internet.

We propose a third approach that does not need the support of any infrastructure (i.e., ad hoc), based on peer-to-peer resource sharing among wireless devices. In peer-to-peer mode, the participants share their resources dynamically based on a user-defined policy. The policy specifies the degree of cooperation, resource sharing, and functionality of peers. The peer-to-peer concept was originally introduced in the context of distributed systems and "reappeared" in 1999 with the widespread popularity of Napster. As we discuss in Section 4, there is substantial peer-to-peer work in the file system and OS literature that is relevant most of them meant for a wide-scale, Internet-based use. As will become clear in the next paragraph, we target a different environment (of mobile wireless data access) and address different research issues. Ad hoc mobile networks are also based on a peer-to-peer mechanism for routing packets among the hosts. This multi-hop routing assumes a relatively high density of devices that are willing to cooperate with each other by routing packets. However, it is not always realistic to assume a connected network of cooperative devices. In this work, we introduce a peer-to-peer system in a new setting. Depending on the density of the mobile peer hosts, their network can be disconnected. We also vary the degree of cooperation of the peer hosts.

Currently, mobile users can access information using the infrastructure of base stations (wireless LAN or WAN). Most wireless data WAN access are only available in major metro areas (such as, Vindigo [36] or RIM [32]). There are situations where a communication infrastructure is not available (such as in emergency situations, disaster relief, rescue teams, inside tunnel or subway). In other situations, there is an infrastructure, but it is overloaded or expensive to access. For instance, on September 11, 2001, after the terrorist attack in New York, it was difficult to access the communication infrastructure and the news web sites. Given the exceedingly expensive license fees attained in recent government auctions of spectrum, the bandwidth expansion route is bound to be expensive. For example, European telecommunications giants spent $100 billion in 2000 for 3G license fees [13]. Similarly, the cost of tessellating a coverage area with a sufficient number of base stations or infostations coupled to the associated high speed wired infrastructure cost is prohibitive. For the next few years, continuous connectivity to the Internet will not be available at low cost for mobile users roaming a metropolitan area. The devices will continue to experience changes in the availability of bandwidth and frequent interruptions of connectivity due to host mobility.

Our main challenges are to accelerate the data availability and enhance the dissemination and discovery of information when hosts face changes in the availability of bandwidth and face the loss of connectivity to the Internet due to host mobility. We aim to investigate ways to enable these devices to share resources to increase their data access considering their power, bandwidth, and memory constraints. We propose $7DS$[1] as a system that complements the three mobile information access approaches we described in the previous paragraph. 7DS is an architecture and set of protocols enabling resource sharing among peers that are not necessarily connected to the Internet. We focus on three different facets of cooperation, namely, *data sharing, message relaying*, and *bandwidth sharing*.

7DS relays, searches and disseminates information, and shares bandwidth. It operates in a self-

---

[1] "7DS" stands for "Seven Degrees of Separation", a variation on the "Six Degrees of Separation" hypothesis, which states that any human knows any other by six acquaintances or relatives. There is an analogy with our system, particularly, with respect to data recipients and the device with the "original" copy. We have not explored if a similar hypothesis is true here.

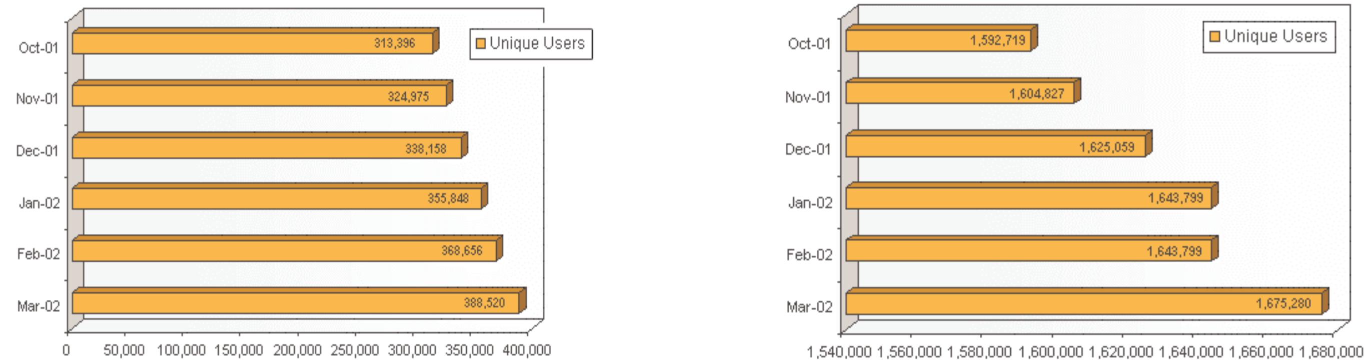


Figure 1: The unique number of Vindigo & Avango users that subscribe to the NYTimes news on-line information, respectively. Source: *The New York Times on the Web* [23].

organizing manner, without the need for an infrastructure and exploits host mobility. It runs as an application on heterogeneous devices (with different capabilities) that are mobile or stationary. A 7DS-enabled device communicates with peers via a wireless LAN. The 7DS-enabled handheld devices are sporadically connected to the Internet and 7DS can coexist with other data access methods (e.g., via wireless modem). 7DS nodes can collaborate by data sharing, forwarding messages (such as, rebroadcasting queries and data or relaying messages to an Internet gateway) or caching popular data objects. In the information sharing facet, peers query, discover, and disseminate information. The system exploits the spatial locality of information that results from the type of services we expect a mobile user will run, namely location-dependent, news services, and collaborative applications.

An environment is characterized by *spatial locality of queries and information*, when users request for location-dependent data and it is likely users in close geographic proximity to query for similar data. It is likely, in an urban environment, such as a part of Manhattan during rush hours, the platform of a train, an airport, a commercial center, a corporation, or a campus we anticipate that the access patterns of the wireless devices will feature high spatial locality of information (such as local and general news, sports, train schedules, weather reports, maps, routes), service discovery queries and also popular information (such as music files or video games). For example, Figure 1 show users of two wireless Internet service providers, Avantgo and Vindigo [2, 37], respectively. We measured the spatial locality of the wireless information via the IEEE 802.11 infrastructure of a campus network at UNC. We observed that over

7% of all requests are for objects which have been requested by a nearby user within the last hour [21]. Furthermore, this proportion varies widely; at some locations on the campus, over 20% of all requests were for such objects. As we discuss in Section 5, we are also in the process of measuring the spatial locality of wireless information in different settings.

7DS takes advantage of the host mobility and periodically queries for data. A host, instead of operating with high transmission power to reach a base station or an infostation that is far away, forwards its messages or requests for data to its peers in close proximity. In that way, the hosts can conserve more power and better utilize the wireless bandwidth. The system uses a simple energy conservation mechanism that periodically enables the network interface. During the *on* interval, 7DS hosts communicate with their peers. In its asynchronous mode, the on and *off* intervals are equal but not synchronized. In synchronous mode, the on and off intervals are synchronized among hosts, although not necessarily equal.

When bandwidth sharing is enabled, the system allows a host to act as an application-layer gateway and share its connection to the Internet with other hosts. When a peer is unable to access the Internet, it may request other peers to act as gateway. Alternatively, hosts can buffer their messages, if they do not have Internet access. When message relaying is enabled, a host forwards its queued messages to another peer. Hosts also relay all their messages when they gain Internet access (via a gateway or base station). 7DS restricts the number of times it forwards a message to a gateway or another relay host.

To evaluate 7DS and the impact of cooperation among hosts, we introduce a general framework with two modes of interaction the peer-to-peer (P-P) and the server-to-client (S-C) modes. 7DS acquires data from other peers (in P-P) or from an infostation (S-C) within its wireless coverage using single-hop broadcast. The motivation of P-P mode is to exploit host mobility and spatial locality of information, better utilize the wireless throughput, and reduce the average delay that a message experiences until it reaches the Internet or acquires some information. In this new framework, we address the effect of wireless coverage range, density of devices, query mechanism, type of cooperation among hosts and their power conservation strategy on data dissemination. For example, we analyze how fast the information spreads in such setting if all nodes are cooperating with each other, and how the performance of data dissemination changes when only a few nodes cooperate (e.g., the 7DS-enabled servers). The performance of data dissemination is defined by the percentage of the nodes that acquire a data item over time, and the average delay that a node experiences until it receives the data. We compare the server-to-client and the peer-to-peer approaches and evaluate how the wireless coverage range, energy conservation, speed, density of devices and servers affect the data dissemination. We also investigate the message relaying, and find the number of times a host should relay a message to another host to reach the Internet. The investigation of these issues can also give insight for the design of a wireless information infrastructure in a metropolitan area.

7DS acquires the data from other peers within its wireless coverage using single-hop broadcast. Due to the highly dynamic environment and the type of information, 7DS does not try to establish permanent caching or service discovery mechanisms. Instead, we explore the *transient aspect of information dissemination*.

We found that the density of the cooperative hosts, their mobility, and the transmission power have great impact on data dissemination. For a region with the same density of hosts, P-P outperforms S-C with no cooperation among the mobile devices. The simulations indicate that the probability a host querying a data object will acquire it by time t follows the function $1 - e^{-a\sqrt{t}}$ when using S-C mode with fixed server and no cooperation among the mobile devices (i.e., fixed information server or FIS). In case of high density of cooperative hosts, the data dissemination using P-P grows even faster.

We also discover two important scaling properties of data dissemination by expanding the area and varying the speed, the density of wireless coverage (i.e., average wireless coverage per space unit) of cooperative hosts, and the density of cooperative hosts (i.e., average number of cooperative hosts per space unit). First, the performance remains the same when we scale the area but keep the density of the cooperative hosts and transmission power fixed. Secondly, for a fixed wireless coverage density, the larger the density of cooperative hosts, the better the performance. In S-C, this implies that for the same wireless coverage density, it is more efficient to have a larger number of cooperative hosts with lower transmission power than fewer with higher transmission power. We can further generalize our simulation results using these properties. These results can also assist in the design of wireless data infrastructures.

A discussion about the design and implementation of 7DS can be found in [28, 29, 26]. In [27], we focused on the network connection sharing. In this paper, we concentrate on the information sharing and message relaying. More specifically, the contributions of this paper are as follows:

1. An evaluation via extensive simulations of 7DS and the effects of the wireless coverage range, 7DS host density, querying mechanism, energy conservation, and cooperation strategy among the mobile hosts as a function of time.

2. Synchronous energy conservation, a mechanism that saves substantial energy, without degrading the efficiency of data dissemination.

3. An analytical model for FIS using theory from random walks and environments, and the kinetics of diffusion-controlled processes. The analytical results on data dissemination are consistent with the simulation results for FIS.

Section 2 describes in more detail the P-P and S-C models and presents simulation results. Section 3 discusses the modeling and analysis of FIS using kinetics of diffusion controlled processes. In Section 4, we discuss related work on peer-to-peer systems and mobile data access. Finally, in Section 5, we summarize our results and discuss directions for future work.

2 Simulation study

The performance analysis of information dissemination does not appear to be amenable to an analytical solution except for simplified settings with respect to the node layout, mobility pattern, and user interaction pattern. Also, there are no real traces available for the access patterns of mobile, wireless users which would be adequate for our purposes. Thus, to investigate these issues and also assess the efficiency of information dissemination via $7DS$, we perform a simulation-based study. In addition to the simulations in Section 3, we present our initial analytical results using diffusion-controlled processes theory. The simulations and analysis are not tied to $7DS$, and provide more general results on data dissemination. Recently, we begun using the actual testbed to measure the performance of the system. Earlier, this was not possible primarily due to cost reasons (e.g., hiring a large number of users to more accurately "approximate" the user's social behavior).

In P-P, 7DS hosts cooperate with each other. S-C schemes operate in a more asymmetric fashion: there are cooperative hosts that respond to queries and non-cooperative, resource constrained clients. 7DS hosts can collaborate by data sharing, forwarding messages (such as, “rebroadcasting” queries and data or relaying messages to an Internet gateway), or by caching popular data objects.

In the simulations, we fix the data object. For simplicity, we refer to the 7DS hosts in these schemes as nodes or peers and the 7DS host that has the data originally in the S-C schemes as the server. At the beginning of each experiment, only one 7DS host has the data item of interest, and the remaining hosts are interested in this data item. We consider a simple energy conservation mechanism that periodically enables the network interface. During the *on* interval, 7DS hosts communicate with their peers. In its asynchronous mode, the on and *off* intervals are equal (but not synchronized). In synchronous mode, the on and off intervals are synchronized among hosts, although not necessarily equal. Servers send advertisements periodically to announce their presence. Power-constrained devices can use a “passive” mode for participating in the system. In particular, they participate only when the expectation for data availability is high, such as when they receive an advertisement. A 7DS host in passive mode sends a query only when it receives an advertisement. We call this “passive” querying, as opposed to active querying that takes place periodically until the host receives the data.

7DS can operate in different modes based on the cooperation strategy among peers (data sharing, forwarding), energy conservation and query mechanism (active or passive querying). To investigate its performance, in particular the effect of transmission power, and the different modes of operation on data distribution, we evaluate P-P and S-C along with their variants. The wireless range of the network interfaces also varies. We evaluate these approaches by measuring the percentage of hosts that acquire the data item as a function of time, and their average delay.

2.1 System models and operation modes

Servers send advertisements periodically to announce their presence. Power-constrained devices can use a “passive” mode for participating in the system. In particular, they participate only when the expectation for data availability is high, such as when they receive an advertisement. A 7DS host in passive mode sends a query only when it receives an advertisement. We call this “passive” querying, as opposed to active querying that takes place periodically until the host receives the data. In the P-P schemes, all nodes are mobile with active querying enabled. We simulate three variations depending on the type of cooperation, namely data sharing (DS), forwarding (FW) and both data sharing and forwarding enabled (DS+FW). When forwarding is enabled, upon the receipt of a query or data, 7DS peers rebroadcast it, if they have not rebroadcast another message during the last 10 s. The last condition is a simple mechanism for preventing flooding in the network. For example, host A queries for some data and host B receives A’s query. B rebroadcasts A’s query, because it does not have any relevant data. Host C receives B’s message and rebroadcasts A’s query, since it does not have any relevant data. Host B receives the query rebroadcast by C, but ignores it.

We separate the S-C schemes into the “straight” S-C scheme without any cooperation among clients (namely, FIS and MIS) and some hybrid ones with cooperative clients. In the FIS (MIS) scheme, there is a fixed (mobile) host with the data that acts as a server. The remaining nodes (clients) are mobile, non cooperative with active querying enabled, and without any energy conservation mechanism. They receive data only from the server. The hybrid schemes are with passive querying enabled and fixed server. In passive querying mode, the server sends an advertisement every 10 sec. Hosts send queries upon the receipt of an advertisement. Let us describe the main motivations for the comparisons made in the remaining section. The P-P vs. straight S-C comparison helps to understand the effect of the cooperation among mobile peers. The P-P and MIS vs. FIS show how mobility affects data dissemination. In particular, the MIS vs. FIS comparison investigates the effect of server mobility on data dissemination.

Nodes move in a 1000 m x 1000 m area according to the random waypoint mobility model [4]. This random walk-based model is frequently used for individual (pedestrian) movement [4, 33, 38]. The random waypoint breaks the movement of a mobile host into alternating motion and rest periods. Each mobile host starts from a different position and moves to a new randomly chosen destination. For each node, the initial and end points for each segment are distributed randomly across the area. Each node moves to its destination with a constant speed selected randomly from the interval (0 m/s, 1.5 m/s). When a mobile host reaches its destination, it pauses for a fixed amount of time, then chooses a new destination and speed

Model	**Cooperation**	**Options**	**Querying**
S-C	only server, server mobile/fixed (FIS/MIS)	DS (only server)	active
P-P	all hosts	DS, FW, DS+FW	active
Hybrid	fixed server, cooperative peers	DS, FW, DS+FW	passive

Table 1: Summary of the schemes with their querying mechanism.

(as in the previous step) and continues moving. Later in the section, we describe two scaling properties that allow us to show that our simulations are robust and to generalize the results when we expand the area or increase the user speed.

The query interval consists of an *on* and *off* interval. The broadcast is scheduled at a random time during the *on* interval. The asynchronous mode is the default energy conservation method. We explicitly denote the schemes having synchronous mode enabled with the word “sync”. In schemes with no energy conservation, the *off* interval is equal to 0 and the *on* and query interval are the same. In the simulations, we concentrate only on the exchange of 7DS queries, data, and advertisements. A cooperative dataholder responds to a query by sending the data item. In this simulation study, we assume one data object, and all hosts in the area are interested in this data item. When a host receives the data item, it becomes a dataholder. This simplification is reasonable in order to investigate the dominant parameters on data dissemination.

Later in this section, we scale the area and vary the density of the hosts and their wireless coverage. We emphasize that this host density does not necessarily represent the total number of hosts in that area, but just indicates the popularity of the defined data object. By varying the density of hosts, we study how data items of different popularity disseminate in such environment. We speculate that in an urban environment such as Lower Manhattan, near the platform of the train or subway stop in a rush hour, there will be from four to 25 wireless devices (carried by humans or integrated into physical objects) that would be interested in getting the local and general news using PDAs or other wireless devices. A density of 25 hosts per km^2 corresponds to a very popular data item whereas a density of five hosts per km^2 corresponds to a more typical data object.

A *scenario* (file) defines the topology and movement of each host that participates in an experiment.

We generate 300 different scenarios for different density values. In each of these scenarios, the mobility pattern of each host is created using the mobility pattern we described, except in the FIS-based schemes, where the server is stationary. We run simulations using these scenarios, for the different schemes of Table 1. The wireless LAN is modeled as an IEEE 802.11 network interface. We use the ns-2 simulator [11] with implementation of mobility and wireless extensions contributed from the CMU Monarch project [1]. The raw bandwidth capabilities of the network interface is 2 Mb/s shared by all hosts in the wireless LAN. Hosts in the wireless LAN communicate with each other directly without the need of a base station. We consider transmission powers of 281.8 mW (high), $\frac{281.8}{2^4}$ mW (medium) and $\frac{281.8}{2^8}$ mW (low). Assuming the two-ray ground reflection model these transmission powers correspond to ranges of approximately 230 m, 115 m ($\frac{230}{2}$ m)and 57.5 m ($\frac{230}{4}$ m), respectively. Note that considering the simulation results on these ranges and the scaling properties we describe later, we can generalize the performance of 7DS for different transmission powers.

2.2 Performance evaluation

Parameter	**Value**
Pause time	50 sec
Mobile user speed	(0,1.5) m/sec
Server advertisement interval	10 sec
Forward message interval	10 sec

Table 2: Simulation constants in 7DS.

We evaluate the effectiveness of our approaches by computing the percentage of nodes that acquire the information after a period of time. In the percentage we do not include the node that has the data item at the beginning of the simulation. We also compute the average delay until a mobile host receives the information from the time it sends the first query. For the computation of the average delay, for each simulation set we consider only the queriers that acquired the data by the end of the simulation and average their delay. We ran the 300 generated scenarios for each test and computed the average of the percentage of hosts that *become* dataholders by the end of each test. These are finite-horizon simulations, so that we do not have to deal with initialization bias. The default simulation time is 25 minutes. However, we also investigate data dissemination over time and vary the simulation time from 150 sec to 50 minutes. The 95% confidence interval for the average percentage of dataholders is within 0-11% of the computed average, with the variance tending to be higher for low host density.

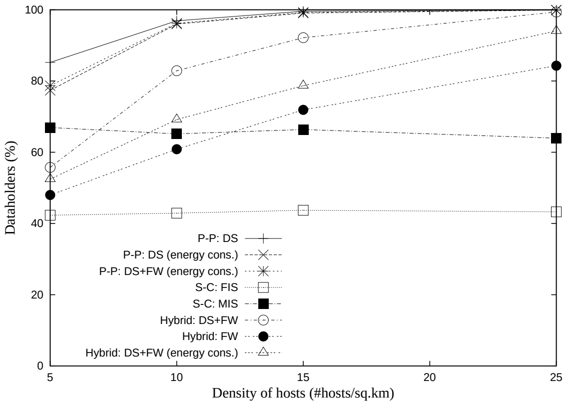


Figure 2: Percentage of dataholders after 25 minutes for high transmission power. The query interval is 15 sec.

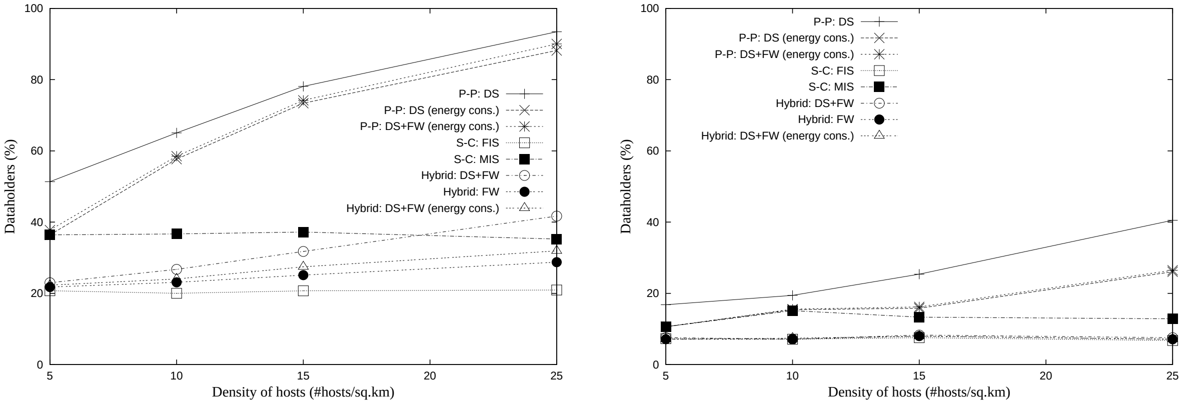


Figure 3: Percentage of dataholders after 25 minutes for medium and low transmission power. The query interval is 15 sec.

Figures 2 and 3 show the percentage of dataholders as a function of the density of hosts for P-P and S-C schemes. In this set of simulations, the query interval is 15 sec. For high transmission power, as in Figure 2, 7DS proves to be an effective data dissemination tool. Even when the network is sparse, 77% of the users will acquire the data during the 25 minutes of the experiment. For networks with ten or more hosts, more than 96% of the users will acquire the data within 25 minutes. For host densities of 25 hosts per km^2, the probability of acquiring the data is very close to 100%. The P-P vs. FIS comparison illustrates the effect of data sharing among mobile peers. In Figure 2, in a setting of 25 hosts, P-P schemes outperform FIS by 55%. In particular, in P-P, 99.9% of hosts will acquire the data after 25 minutes, compare to 42% of the users in the FIS. For lower transmission power, P-P outperforms FIS by 20% to 70% (Figures 2 and 3). The impact of data sharing among peers is also apparent in hybrid schemes. It is more evident in the hybrid vs. S-C schemes for density of ten or more hosts per km^2 and medium or high transmission power.

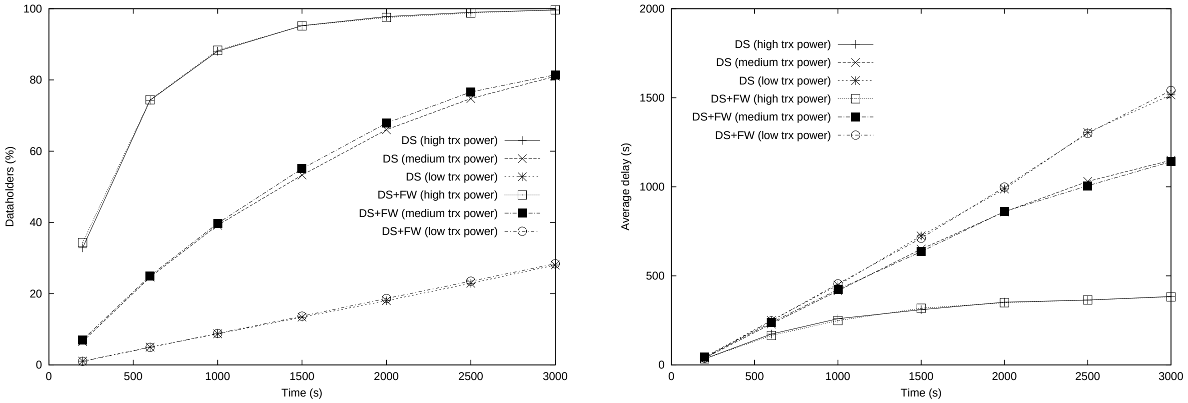


Figure 4: Effect of forwarding on peer-to-peer with data sharing enabled (DS). There are 10 hosts (one initial dataholder) per square kilometer.

Note that forwarding in addition to data sharing does not result in any further performance improvements. This is due to the low probability that a case such as the following occurs: There are a querier A and a dataholder C that cannot listen to each other, and a third host B that can communicate with both and forward data. Moreover, A will not acquire the data directly from a dataholder until the end of the test. We emphasize that this is also true for smaller simulation times, starting from 150 sec (just a few seconds after the hosts start querying). Figures 4 and 5 illustrate the effect of forwarding as a function of time in two settings of 10 hosts per square kilometer and 25 hosts per square kilometer, respectively. In a more dense setting of mobile hosts that forward messages independent of their own data interests, we expect forwarding to have a higher impact. As we mentioned earlier, in order to prevent flooding, when forwarding is enabled, upon the receipt of a query or data, 7DS peers rebroadcast it, if they have not rebroadcast

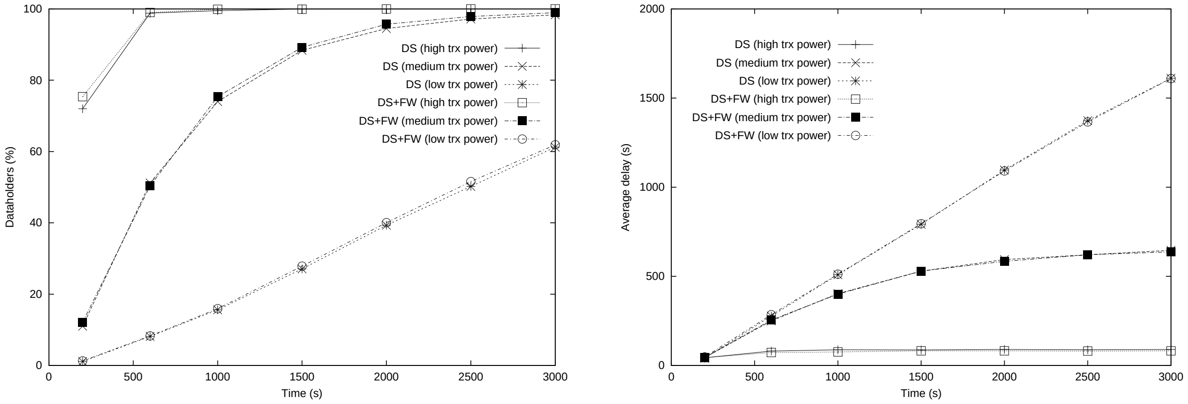


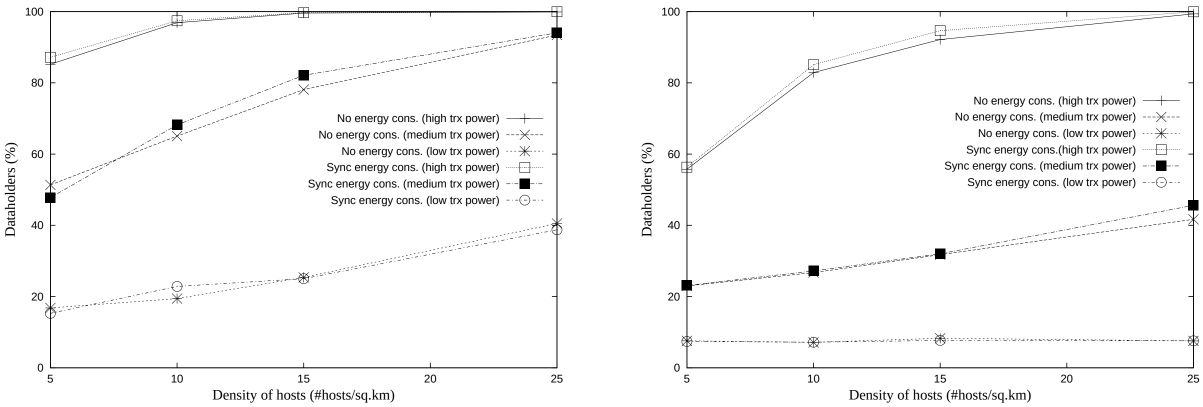
Figure 5: Percentage of dataholders and average delay of peer-to-peer with data sharing and forwarding. Effect of forwarding on peer-to-peer with data sharing enabled (DS). There are 25 hosts (one initial dataholder) per square kilometer.

another query or data during the last 10 s, respectively. This restricts also the effect of forwarding. The use of a routing protocol among the mobile hosts could potentially enhance the impact of forwarding. The impact of forwarding is more apparent in schemes with data sharing among peers disabled. For example, Figures 2 and 3 show that hybrid schemes with forwarding enabled outperform FIS by 4%-40% depending on transmission power.

As we expect, the performance of both FIS and MIS remains constant as the number of hosts increases, since a data exchange takes place only when a querier is in close proximity to the server. In addition, notice that in Figures 2 and 3, MIS outperforms FIS by approximately 22%, 16%, and 6%, respectively. An intuitive explanation is based on the fact that, in MIS schemes, the relative speed of the server from the clients is larger than in FIS schemes where the server is fixed. Therefore, the mobile information server will meet with more hosts and disseminate the data faster. On the other hand, as we expect, the density of hosts affects the schemes that are based on peer-to-peer cooperation. As the number of hosts increases from five to 25 hosts, in P-P schemes with medium transmission power, the performance improves substantially.

Impact of energy conservation

Measurements of the energy consumption of the wireless network interfaces have shown that they consume substantial power even when they are idle (powered on but not sending or receiving packets). Moreover, receiving packets costs marginally more energy than being idle [35]. Using the asynchronous energy conservation mechanism, there is a 50% energy savings, since the network interface is on only half the time. As Figures 2 and 3 illustrate there is some degradation in data dissemination. This is due to the decrease of the time interval the hosts can communicate. If we keep the query interval constant and reduce the *on* interval, the smaller the *on* interval, the higher the energy savings. However, with smaller intervals, the degradation of data dissemination is larger. To prevent this degradation, we enable synchronous mode. In synchronous mode, the on and off intervals of all hosts are synchronized. As we show in Figure 6, when the synchronous mode is enabled, even with a small *on* interval, energy conservation does not cause any degradation of the data dissemination. More specifically, Figure 6 (a) illustrates P-P schemes with data sharing and Figure 6 (b) hybrid schemes with data sharing and forwarding. The query interval is 15 sec, in which, during the first 1.5 sec the network interface is on, and during the remaining time (13.5 sec), it switches off. In an ideal setting without packet losses and need for retransmission, the number of messages exchanged in the P-P schemes without energy conservation and the ones with synchronous energy conservation are the same. Therefore, the power spending on message receiving/sending is the same, whereas the period in the idle state is reduced (the network interface is on for only 10% of the time). The synchronous mode may result in a 90% reduction in energy dissipation. In general, hosts may query for different data items. In very dense settings retransmissions and packet losses may result in further energy spendings. It



P-P scheme with data sharing Hybrid scheme with data sharing and forwarding

Figure 6: The impact of synchronous mode on data dissemination. Query interval is 15 sec and the on period in “sync” schemes is 1.5 sec. The simulation time is 25 minutes.

is part of our future work to investigate further the synchronous power conservation mode and the impact of retransmissions, packet loss, and on interval in such environment.

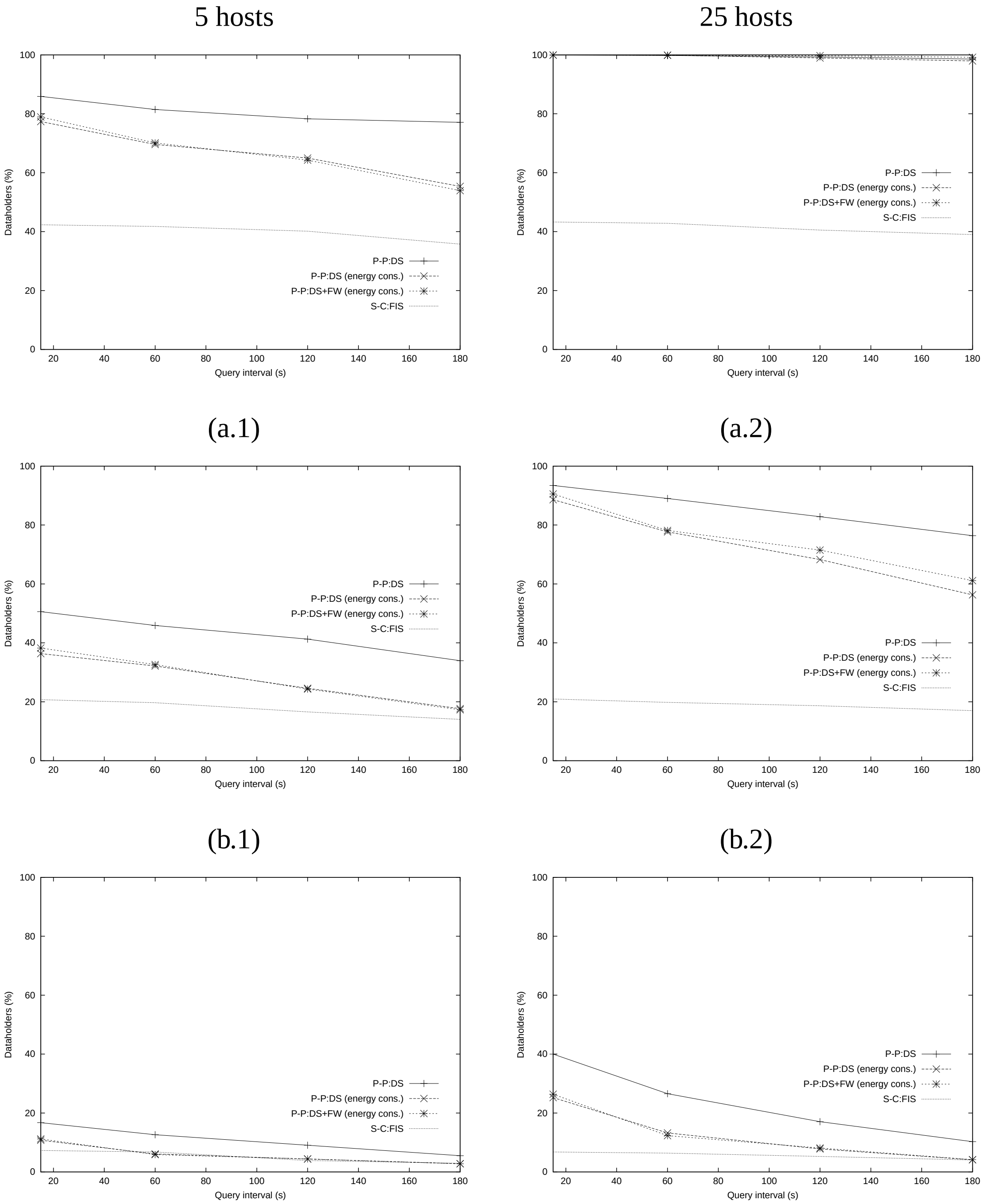
Impact of query interval

We investigate the performance of the system as a function of the query interval using the asynchronous energy conservation method (i.e., the on interval is half the query interval and is not synchronized). The degradation in the FIS performance is relatively small compared to P-P schemes as the query interval increases. This is due to the high probability that a mobile host that gets in close proximity to a server acquires the data (i.e., there is sufficient time to broadcast a query and receive the data).

Figures 7 (a.1), (b.1), and (c.1) correspond to a relatively sparse network of five hosts per square kilometer, whereas Figures 7 (a.2), (b.2), and (c.2) correspond to a more dense network of 25 hosts/km^2. As Figure 7 illustrates, in P-P schemes the impact of the query interval can be more apparent. In a setting of 25 hosts with medium transmission power, data sharing, and no energy conservation, when the query interval increases from 15 sec to 3 minutes the degradation is approximately 30% and for five hosts, it reaches 50%. However, using the synchronous energy conservation, even when we maintain a low ratio of the on-interval (e.g., 5% with on interval to be 6 sec and query interval 2 minutes), we expect the degradation to be much weaker. We need to investigate further what is the optimal on and query interval, and when a mobile host need to switch to passive querying to utilize its battery more efficiently, taking also into consideration the average time that two hosts are in close proximity, and the traffic in the wireless LAN.

Measurement of average delay

As we mentioned earlier, we measure the average delay a host experiences from the first query until it receives the data. For each test, we compute the average delay of the nodes that acquired the data by the end of simulation. Note that for the computation of the average delay, we only consider the hosts that received the data by the end of the simulation. Then, we take the average over all 300 sets, excluding the ones without new dataholders. First, let us fix the simulation time to 25 minutes and compare P-P and FIS schemes in terms of average delay for the same probability of acquiring the data. In P-P with



(a.1) (a.2)

(b.1) (b.2)

(c.1) (c.2)

Figure 7: Percentage of data holders as a function of the query interval. The first and second column correspond to scenarios with 5 hosts per km^2 and 25 hosts per km^2, respectively. Figures (a), (b), and (c) correspond to a high, medium and low transmission power, respectively. Schemes with energy conservation enabled use the sync mode.

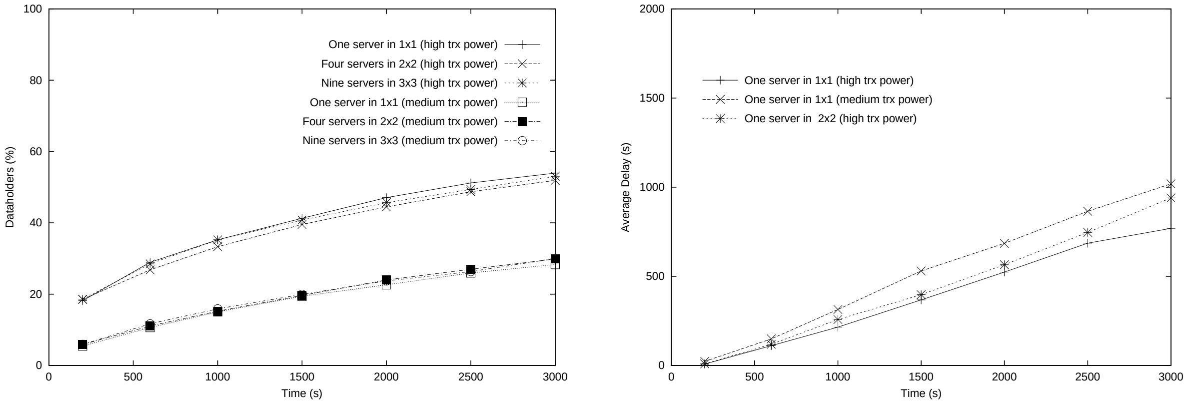


Figure 8: Percentage of dataholders and average delay of the fixed information server (FIS) schemes as a function of the simulation time. “AxA” indicates the size of the area in square kilometers. For example, the curve with the circle corresponds to a FIS scheme with nine servers per 3 km x 3 km (all hosts transmitting with medium power).

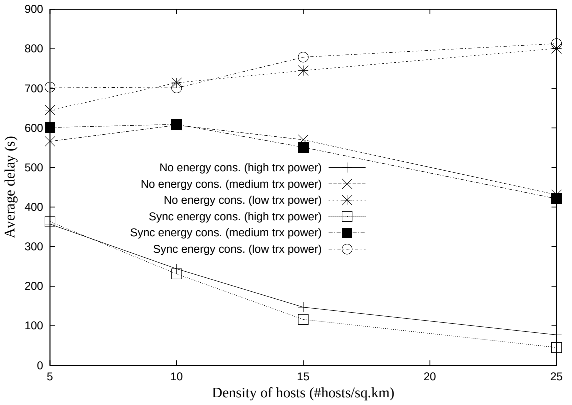


Figure 9: The average delay for the P-P with data sharing. Query interval is 15 sec and the on period in sync schemes is 1.5 sec. The simulation time is 25 minutes.

data sharing and no energy conservation, for high transmission power (Figure 9), the average delay is as high as 6 minutes for sparse networks and drops to 77 sec for dense networks (Figure 9). In the case of low transmission power, it reaches 13 minutes. Using FIS, the average delay is constant over the number of hosts in the area. For high transmission power, it is 6 minutes, while for low transmission power it reaches 9 minutes. So, (sync) P-P with data sharing, even in the case of low host density, performs better than FIS. For the same average delay to acquire the data (6 minutes), the probability to acquire the data in the P-P doubles. This becomes clear when we compare P-P in Figure 6 (a) and 9 (five peers with high

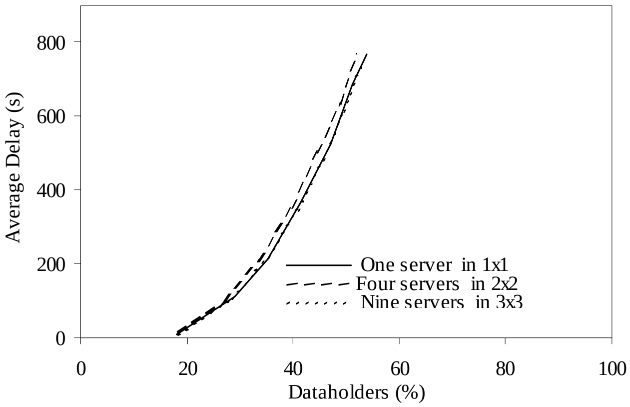
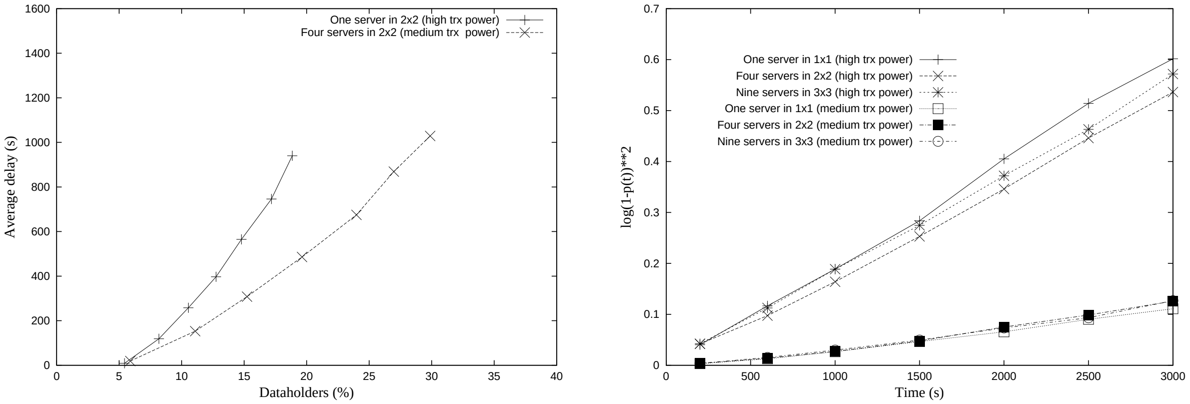


Figure 10: The average delay to receive the data as a function of the probability to acquire it in FIS for high transmission power. The "AxA" indicates the size of the area where the server and mobile clients are placed (in square kilometers).

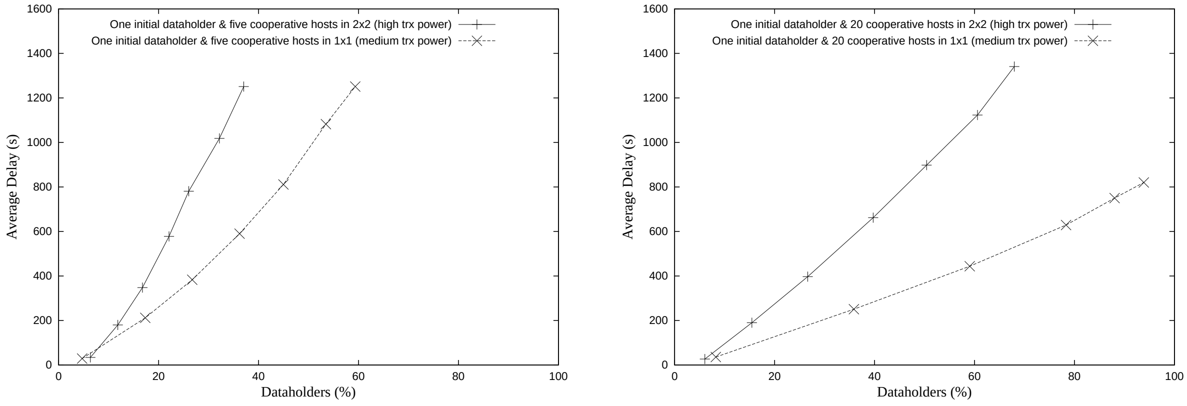


(a) (b)

Figure 11: Figure (a) shows the average delay to receive the data as a function of the probability to acquire it in FIS. Figure (b) shows the performance as a function of the simulation time. The "AxA" indicates the size of the area where the server and mobile clients are placed (in square kilometers).

transmission power) and FIS in Figure 10 (e.g., the one server in 1x1 with high transmission power).

For the Figures 10, 11 and 12, we have combined the simulation results for the probability a host acquires the data, and the average delay it experiences as functions of time. For example, in the case of one server in a 1 km x 1 km area with high transmission power in Figure 10, we use the results of Figures 8 that correspond to one server in a 1 km x 1 km area with high transmission power. For that, we find in Figure 8 (a) at which time t a given percentage of dataholders p is reached, and then in Figure 8 (b), the



Five cooperative hosts

Twenty cooperative hosts

Figure 12: The average delay to receive the data as a function of the probability to acquire it in P-P with data sharing schemes. The “AxA” indicates the size of the area (in square kilometers).

average delay d that hosts who received the data by that time t have experienced (since their first query was sent). The graph in Figure 10 consists of these pairs (p,d). The percentage of hosts that acquire the data in P-P with high transmission power reaches 40% with an average delay of 135 sec. With the same delay and using FIS, 30% of hosts will acquire the data (Figure 10). With FIS, a 40% probability of acquiring data corresponds to an average delay of 6 minutes (Figure 10) whereas using (sync) P-P this probability doubles, even for a low density cooperative host setting (Figure 9 and 6 (a)). For a higher average delay of 10 minutes, 85% hosts will acquire the data using P-P (Figure 15 for five hosts at simulation time equal to 2000 sec), and 50% using FIS (Figure 10). In the case of medium transmission power, with an average delay of 315 sec, a host will get the data with a probability of 15% and 22% using FIS (one server in 1x1 or four servers in 2x2 scheme in Figure 11) and P-P (one initial dataholder and five cooperative hosts in 1x1 scheme in Figure 12) (left), respectively.

Scaling properties of data dissemination

Let us now discuss the scaling properties and generalize our performance results. First, we focus on expanding the area but keeping the movement pattern the same. In both P-P and FIS schemes, when we expand the area but keep the density of hosts and their transmission power fixed, the performance of data dissemination remains the same. This indicates that our simulation results are robust. For example, Figure

10 showes this scaling property in FIS for high transmission power. Specifically in FIS, it is sufficient to fix only the density of the servers, since only the servers cooperate. Let $p(t)$ denote the probability a host will acquire the data by time t. Figure 11 (b) shows the probability that a host will *not* acquire the data by time t, i.e., $1 - p(t)$, or *survival probability* on a logarithmic scale. This figure shows the percentage of data holders at time t using the transformation $(\log(1 - p(t)))^2$. Their shape indicates that $p(t)$ in FIS follows the $1 - e^{-\sqrt{at}}$. In P-P settings (e.g., P-P with data sharing and energy conservation) $p(t)$ grows faster than in FIS. Our simulation results indicate that the P-P with data sharing and energy conservation can be approximated by the function $1 - e^{-at}$, especially for less dense settings, such as those with fewer than 20 hosts per km^2 transmitting with high or medium power. In the above function, the constant a depends on the density of FIS servers. For very dense settings, this probability grows even faster.

Another important scaling property is related to the effect of density of cooperative hosts vs. their wireless coverage density. Assuming the same total area of wireless coverage, we investigate the impact of host density for both the P-P and FIS schemes. Particularly in FIS, this can be viewed as a design decision. Figure 13 illustrates two possible deployments of servers with the same wireless data coverage, assuming an ideal transmission model with the power inverse to the square of distance. The density of servers in the left is higher than that in the right, but they have lower transmission power. For both settings, we assume the same mobility pattern. Figure 13 depicts a host moving with fixed speed *v* and traveling on a line segment during an interval. The setting of the larger number of servers with lower transmission power is more effective in terms of power spendings and wireless throughput utilization. We found that for fixed total wireless coverage, the higher the cooperative host density, the better the performance. Simulations indicate that this is true with both the FIS and the P-P schemes. An intuitive explanation is that, in Figure 13, the two deployments become equivalent by “scaling down” the left scheme (to match the right one). But after this “scaling”, it is as if the speed of the hosts at the left scheme doubles. That is, the left setting is the same as the right one, in terms of area, transmission power of the servers, and server density, but with the hosts moving faster. Therefore, the probability a host will get into the coverage area of a server increases.

Figure 11 (a) compares two FIS settings with the same total wireless coverage density of cooperative

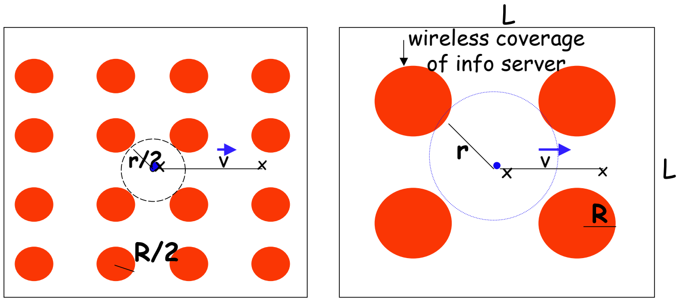


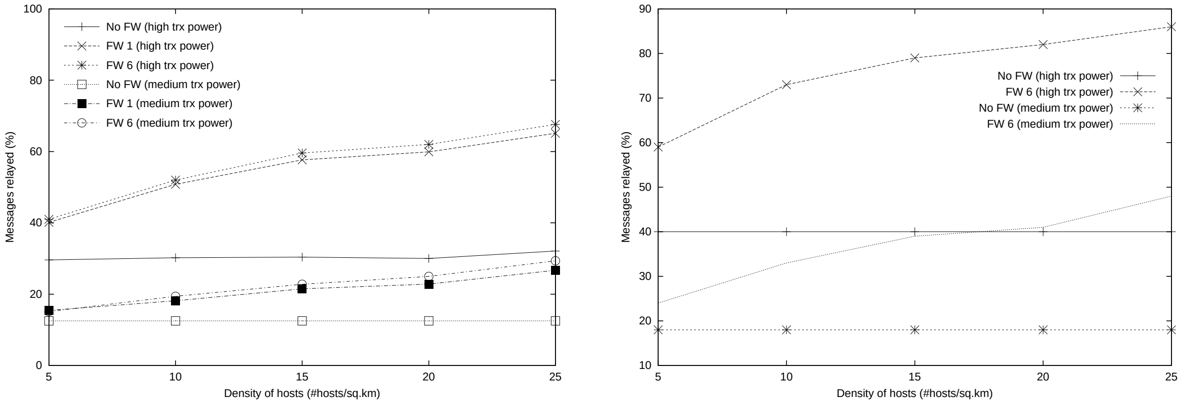
Figure 13: Investigating the scaling properties of data dissemination. The dark disks depict the wireless coverage of a host. For fixed total wireless coverage, the larger the density of cooperative hosts, the more efficient the data dissemination.

hosts (servers). The first includes one server in a 2 km x 2 km area with high transmission power and the latter four servers in a 2 km x 2 km area with medium transmission power. The case with higher density of servers performs better. For example, for a 20% probability of acquiring the data, the FIS scheme with higher density of servers produces an average delay of 500 s. For the same total wireless coverage, but lower density of servers, the average doubles. Figures 12 illustrate similar results in P-P schemes for different host densities. For a 40% probability of acquiring the data, the average delay is 600 sec in the higher density of hosts setting (5 hosts/km^2) whereas in a lower density setting, it doubles. Note that when we scale the speed of the mobile hosts and fix the mobility pattern and host density, we can compute the performance of data dissemination from the previous setting.

2.3 Message relaying

As we discussed in the Section 1, message relaying is another facet of cooperation among mobile hosts. We assume hosts generate messages and buffer them locally when there is no Internet access. When a host gains access by reaching the wireless coverage area of a gateway, it relays these messages to the gateway. A host may relay its own messages to a peer when forwarding is enabled. We investigate the impact of message relaying on the probability that a message will reach a gateway and on the average delay from the time the message was created until it reaches a gateway.

To avoid a message explosion, we impose two restrictions. First, a host relays all queued messages to



(a) (b)

Figure 14: Figure (a) shows the percentage of the messages generated at each host that reach the Internet after 25 minutes. Figure (b) the probability that a message will reach the Internet within 25 minutes from the time it was generated on the source host. We assume an area with one gateway per km^2. We use the notation "FWa" to indicate the maximum number of copies for each message, a, that a 7DS host relays to other nodes.

a gateway, but only its own messages to another peer. That is, a given message reaches the Internet via at most two hops. Secondly, we restrict the number of times a host may relay a given message. When a host has queued messages for relaying, it queries for a gateway or a relay host. A host that receives these queries may respond. Upon the receipt of such response, the querier forwards the queued messages to that host. Those messages need to satisfy the above two restrictions. In addition, a host transmits the same message only once to another host. The gateways periodically advertise their presence. Upon the receipt of such advertisement, a host forwards all the queued messages to the gateway.

We assume one gateway per km^2 area. Figure 14 (a) shows the percentage of the messages generated at each host during an interval (here is 25 minutes) that reach the Internet and the impact of forwarding to relay nodes. We assume that the hosts generate messages with a constant rate of one message every three minutes. The average number of buffered messages at each host is five. For high host density, forwarding doubles the percentage of messages that reach the Internet. Note in Figure 14 (a) that forwarding a message to more than one relay node does not substantially improve the performance (FW6 vs. FW1 schemes). Figure 14 (b) illustrates the probability that a message will reach the Internet within 25 minutes from the time the message was created on the source host. Note that when there is no forwarding, the probability that a message will reach a gateway is the same as the probability that the host will reach a gateway. Essentially, this probability is the same as the probability that a host will acquire the data in FIS for a gateway density equal to the server density in FIS. As in Figure 14 (a), this probability increases when forwarding is enabled.

In a setting with a very low transmission power corresponding to a range of 8 meters (e.g., Bluetooth), and with high density of hosts (such as 100 hosts/km^2), after 2.5 hours, 5% of the messages generated during the first 25 minutes will reach a gateway directly and 38% of them via another relay host. This corresponds to a setting with forwarding enabled and forwarding number equal to 20, so that a given message can be relayed to at most 20 hosts. For a forwarding number equal to six, the percentage becomes 21%. In this section, we discuss the performance of 7DS via simulations and the impact of the transmission power, host density, cooperation, query interval, and querying mechanism on the effectiveness of information dissemination and message relaying.

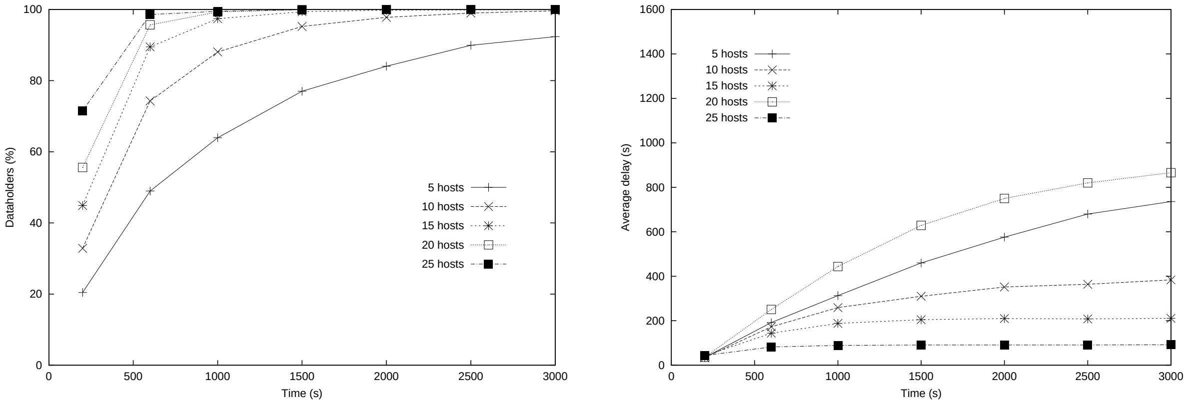


Figure 15: Performance of the peer-to-peer with data sharing and energy conservation enabled (DS) as a function of the simulation time for high transmission power. Each scheme has N cooperative hosts ($N = 5, \ldots, 25$) in a square kilometer area. Initially, one of them is dataholder.

3 Data dissemination as a diffusion-controlled process

This section discusses our initial efforts to study the data dissemination analytically and to further generalize our results. It contributes a novel approach to model data dissemination. We also address the main theoretical results and challenges. The models are based on diffusion-controlled process that uses theory from random walks and environment [17], and the kinetics of diffusion-controlled chemical processes [24]. In particular, we use the diffusion in a medium with randomly distributed static traps to model the FIS scheme.

Let us first define the static trapping model. Particles of type C perform diffusive motion in d-dimension space. Particles of type S ("sinks" or traps) are static and randomly distributed in space. Particles C are absorbed on particles S when they step onto them. The basic trapping model assumes traps of infinite capacity. The diffusion controlled processes focus on the survival probability, that is the probability that a particle will not get trapped as a function of time.

For Rosenstock's trapping model in d dimensions (with a genuinely d-dimensional, unbiased walk of finite mean-square displacement per step), it has been shown that the large-n behavior of the survival probability

$$\log(\phi_n) \approx -\alpha[\log(\frac{1}{1-q})]^{2/(d+2)} n^{d/(d+2)} \tag{1}$$

In Eq. 1, α is a lattice-dependent constant, and q denotes the concentration of the independently distributed, irreversible traps.

One question is when Eq. 1 is a useful approximation. All previous analyses of this question have relied on some form of simulation, but so far there was no information available on the range of validity of Eq. 1. In the Letter [16], Havlin *et al* present evidence suggesting that Eq. 1 is a useful approximation when

$$\rho > 10 \tag{2}$$

where ρ is the scaling function

$$\rho = (ln\frac{1}{1-q})^{\frac{2}{d+2}} n^{\frac{d}{d+2}}. \tag{3}$$

This value of ρ corresponds to a survival probability equal to 10^{-13} in both $d = 2$ and $d = 3$ dimensions. They argue that pure simulation techniques will always lead to an exponential decay at sufficiently long times, rather than to the correct decay given by the theoretically-proven Eq. 1. Their evidence for the new lower value of ρ (or higher value of $P(n)$) is based on two numerical techniques that they have developed. One of them is practical for high trap concentrations only ($0.9 \leq q$). This case of high trap concentrations is similar to our case.

In FIS information sharing takes place among the server and the querier. When a $7DS$ querier comes in close proximity to the server, it acquires the data. It is easy to draw the analogy: the traps are the stationary information servers, the particles C are the queriers, and the trapping is essentially receiving the data. We model the stationary information servers as traps and the mobile peers as particles C. When a host acquires the data, it stops participating in the system, and is considered “trapped” with respect to the model. Figure 16 illustrates the analytical and simulation results for data dissemination as a function of time. The analytical results for trapping model are derived from Eq. 1 (Rosenstock’s trapping model) for high and medium transmission power.

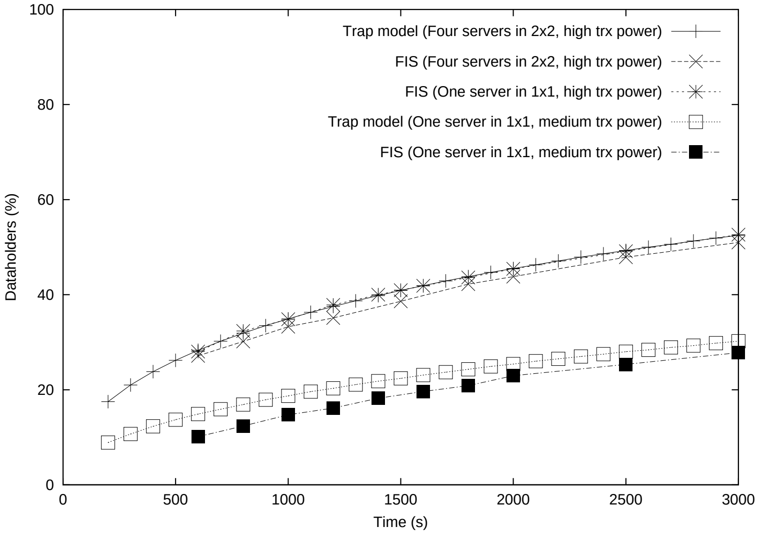


Figure 16: Simulation (FIS) and analytical trapping model (Trap) results. The “AxA” indicates the size of the area in which the servers are placed (in square kilometers).

We define q as $\pi\ R^2 N_{servers}/A^2$, where $N_{servers}$ is the number of servers placed in an area of size AxA and R is the wireless coverage equal to 230 m and 115 m for high and medium wireless coverage, respectively. For the simulation results on FIS in Figure 16, we use the FIS simulations we described in Section 2.2. Note that, using Eq. 1, the term $1 - \phi_n$ expresses the fraction of hosts that acquire the data at time n. As Figure 16 illustrates, our simulations are consistent with Eq. 1 (in two dimensions) for α equal to 0.021. That is, using Eq. 1, the $(1 - \phi_n) * 100\%$ match our simulation results for the percentage of dataholders at time n for the FIS scheme we described. In a setting of one stationary host per square kilometer with high transmission power that corresponds to a range of 230 m, the server concentration is $c = 0.1661$ and the criteria $\rho > 10$ is equivalent to $n > 550$.

An attractive feature of the diffusion-controlled processes in the context of our research is that it can provide elegant tools and methodology to investigate data dissemination for different server distributions. Also, we are currently exploring how we can use it to model other types of interaction (S-C and P-P schemes) and incorporate parameters such as the expiration of data objects.

4 Related work

There is substantial peer-to-peer work in the file system and OS literature that is relevant, including the Ficus [25], JetFile [14], and Bayou [34] projects. All of them are replicated storage systems based on the peer-to-peer architecture. Ficus is a distributed file system meant for a wide-scale, Internet-based use. It supports replication using a single-copy availability optimistic update policy. Its main focus is on the consistency among the different copies and reconciliation algorithms to reliably detect concurrent updates and automatically restore consistency. Like Ficus, Bayou provides support for application-dependent resolution of conflicts. Unlike Ficus, it does not attempt to provide transparent conflict detection. JetFile requires file managers to join a multicast group for each file they actively use or serve. Napster [22] and Gnutella [12] are two systems that explore the cooperation among hosts and enable data sharing among users in a fixed wired network. The first focuses on sharing music files; the latter on any type of file. In contrast to Gnutella, a $7DS$ host does not need to discover its neighbors or maintain connections with them, but only multicasts its queries to a well-known multicast group. In addition, 7DS (in the default mode) restricts the query propagation to the wireless LAN. Unlike Napster, $7DS$ operates in a distributed fashion without the need for a central indexing server. Moreover, Napster requires user intervention for uploading files, whereas $7DS$ does this automatically. Furthermore, our setting is orthogonal to the information discovery in the wide area network. In information discovery, there is typically an infrastructure of cooperative servers that create indices to locate data based on the queries and the content of the underlying data sources of their local domain (e.g., [6, 19]). Our system targets a different environment and addresses different research issues. The primary concern of our work is the effect of the wireless coverage, collaboration strategy, and power conservation method in the data dissemination across mobile hosts, rather than consistent replication.

Ad hoc and sensor networks typically assume a relatively high density of devices that results in a connected network, a host can access other hosts via multi-hop routing [4, 33, 38, 10]. They also assume cooperative nodes, part of the same infrastructure, that relay packets on behalf of other nodes. On the other hand, a $7DS$ network is rarely connected, and it can take minutes for one $7DS$ node to come in close proximity to another. As we mentioned, in our setting, peers have different capabilities and cooperation strategies and they are not necessarily all cooperating with each other. Both in ad hoc and sensor networks, the emphasis is on routing protocols.

Infostations were first mentioned by Imielinski in the DataMan project [31]. Badrinath was among the first to propose an infrastructure for supplying information services, such as e-mail, fax, and web access by placing infostations at traffic lights and airport entrances. Infostations use a single server/multiple clients model in which the server broadcasts data items based on received queries. They mostly address issues related to efficient scheduling algorithms for the server broadcast that minimize the response delay and power consumption of mobile devices and efficiently utilize the bandwidth of the broadcasting channel [18, 31, 3]. Imielinski *et al* [18] investigate methods for accessing broadcast data in such a way that running time (which affects battery life) and access time (waiting time for data) are minimized. They demonstrate that providing an index or hashing based access to the data transmitted over the wireless can result in significant improvement in battery utilization. Barbara *et al* [3] propose and study a taxonomy of cache invalidation strategies and study the impact of clients' disconnection times on their performance.

In a context similar to ours, prefetching targeted for mobile users in a wide-area wireless network has been used in [39]. Tao Ye *et al* [39] consider an infostation deployment. They consider data representation in different levels of detail. Their prefetching algorithm uses location, route, and speed information to predict future data access. Their emphasis is on devising and evaluating techniques for building network-aware applications. They describe an intelligent prefetching algorithm for a map-on-the-move application that delivers maps, at the appropriate level of detail, on demand for instantaneous route planning. When a mobile user enters the infostation coverage area, it prefetches a fixed amount of k bytes that corresponds to a map with a certain level of detail, where k depends on user speed. They investigate the effectiveness of infostations as compared to a traditional wide-area wireless network. There are two main differences of their setting with our FIS based schemes. First, in their environment, mobile clients are constantly connected to a low-speed wireless network. Devices use a high bandwidth link when they are within infostation coverage. Outside these regions, their requests are passed to the server via a conventional cellular base-station. In our case, the mobile hosts have no wide-area network access. Second, they investigate the effectiveness of (fixed) infostations compared to a traditional wide-area wireless network. For that, in their simulation study, they vary the infostation density and its coverage. In our case, we consider a fixed infostation (i.e., FIS) in the region of 1 km x 1 km, corresponding to low infostation density. Here, our focus is to investigate a different data access method, namely, peer-to-peer data sharing among mobile users. For its evaluation, we compare it to the access via an infostation. Also, we vary several parameters that have not been investigated in [39], including various mobility patterns, power conservation methods, and querying schemes. Their qualitative result, that having many infostations covering small ranges is a better topology than having few infostations covering large ranges, is consistent with ours.

Another project with similar goals to ours is Portolano [9]. They also aim to provide service discovery to mobile clients with intermittent connections. Their research exists in a hybrid world where they plan to leverage a wired infrastructure in addition to wireless links. It appears that their emphasis is on user interfaces that allow mobile clients to discover the semantics of any service, and present an interface suited to the client's needs and resource limitations.

Flooding and gossiping (a variant on flooding that sends messages only to some neighboring nodes instead of all) protocols have been also studied extensively. For example, Kulik *et al* [20] present a protocol for information dissemination in sensor networks. In their setting, the sensors are fixed and the network fully connected. They measure both the amount of data these protocols disseminate over time and the amount of energy they dissipate. Their system features meta-data negotiation prior to data exchange to ensure that the latter is necessary and desired, eliminating duplicate data transmissions, and with power resource awareness. They compare their work with more conventional gossiping and flooding approaches.

Grossglauser *et al* [15] show how the mobility can increase the capacity of mobile ad hoc wireless networks. They evaluate the average per-session throughput and asymptotic performance. On the other hand, the main focus of this thesis is on the transient behavior of the message relaying, and the impact of various parameters.

Davis *et al* [7] investigate the message relaying. Their main focus is on the additional storage at nodes as packets are stored, carried, and forwarded to the destination. They impose finite buffer sizes on hosts and investigate different packet dropping strategies. They show that the two drop strategies that perform best (among the ones they consider) are the Drop-Oldest and Drop-Least-Encountered. In the first, the packet that has been in the network longest is dropped. In the latter, the packet is dropped based on the estimated likelihood of delivery. For that, they use information about host location and movement.

Buttyan *et al* [5] consider a geodesic packet forwarding algorithm in order to evaluate micropayment mechanisms for message relaying in an ad hoc network. The geodesic algorithm assumes that the source of a packet knows the geographic position of the destination, its own geographic position, and that of its neighbors. Before sending the packet, the source puts the coordinates of the destination in the packet's header. It forwards the packet to the neighbor closest to the destination. Each forwarding host performs the same operation. If the forwarding host does not have any neighbor that is closer to the destination than the host itself, then the packet is dropped. In their setting, the hosts are stationary.

5 Conclusions and future work

The main challenges in our research is to accelerate the data availability and enhance the dissemination and discovery of information when hosts face changes in the bandwidth availability and loss of connectivity to the Internet due to host mobility. 7DS that addresses this challenge by providing a novel mechanism that enables wireless devices to share resources in a self-organizing manner, without the need of an infrastructure. We measured the percentage of hosts that acquire the data item as a function of time, and their average delay via simulations.

Our results lead to the following conclusions: P-P schemes outperform S-C schemes. The simulation results indicate that the probability that a host that queries for a data object will acquire it by time t using FIS and P-P, follows the $1 - e^{-a\sqrt{t}}$ and $1 - e^{-at}$, respectively. In case of high density of cooperative hosts, data dissemination using P-P grows even faster. Generally, the difference becomes more prominent in cases of medium or low transmission power, with more than ten hosts. In our setting with ten or more hosts per km^2, P-P provides 60% or higher probability for acquiring the data item to hosts that move in the area for 25 minutes and transmit with medium or high power. This probability is two to three times higher than in FIS. In some of the cases, difference in their average delays is negligible, in other cases FIS has lower average delay (with a maximum difference of 100 sec). Also, we found that forwarding (i.e., rebroadcasting 7DS messages upon their receipt) in addition to data sharing does not result in any performance improvements. Beneficial to the data dissemination is the the synchronous energy conservation method. It increases the power savings without degrading the data dissemination. Finally, we investigated some scaling properties of the data dissemination. For example, in both the P-P and FIS schemes, for the same wireless coverage, it is more efficient to have a larger number of servers with lower transmission power than fewer with higher transmission power.

There are several research issues that we plan to investigate further. One main research direction is to extend the synchronized energy conservation method and use the "rendezvous" idea for preventing denial of service attacks by malicious users who jam the wireless LAN. The other research effort continues the measurements of the spatial locality of mobile information access. As mentioned in the Section 1, we are in the process of measuring the spatial locality of the wireless information in the IEEE 802.11b infrastructure of the UNC campus. We expect that several environments and situations (e.g., in a corporation, a classroom) are characterized by spatial locality of information. In those situations and environments, our system can facilitate as a new paradigm of data access for mobile users and potentially improve the performance of information access and collaborative applications. We are integrating a number of location-dependent applications, such as an interactive map editing application and a note-sharing application with 7DS. 7DS acts as the underlying information discovery mechanism. We will run actual experiments with the testbed in different settings such as during a class, in a seminar, in a local coffee store. We will evaluate the effectiveness of the information sharing on improving the response delay and investigate the spatial locality of information property.

References

[1] Wireless and mobility extensions to ns-2. http://www.monarch.cs.cmu.edu/cmu-ns.html.

[2] Avantgo. http://www.avantgo.com/.

[3] Daniel Barbara and Tomasz Imielinski. Sleepers and workaholics: Caching strategies in mobile environments. In *ACM SIGMOD International Conference on Management of Data*, pages 1–12, 1994.

[4] Josh Broch, David Maltz, David Johnson, Yih-Chun Hu, and Jorjeta Jetcheva. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, Dallas, Texas, October 1998.

[5] L. Buttyan and J. P. Hubaux. Nuglets: a virtual currency to stimulate cooperation in self-organized mobile ad hoc networks. Technical Report DSC/2001/001, Swiss Federal Institute of Technology, Lausanne, January 2001.

[6] Paul Castro, Benjamin Greenstein, Richard Muntz, Parviz Kermani, Chatschik Bisdikian, and Maria Papadopouli. Locating application data across service discovery domains. In *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 28–42, Rome, Italy, July 2001.

[7] James Davis, Andy Fagg, and Brian Neil Levine. Wearable computers as packet transport mechanisms in highly partitioned ad-hoc networks. In *Proc. International Symposium on Wearable Computers*, October 2001.

[8] Chip Elliott. Building the wireless internet. *IEEE Spectrum*, 38(1), January 2001.

[9] Mike Esler, Jeffrey Hightower, Tom Anderson, and Gaetano Borriello. Next century challenges: data-centric networking for invisible computing – the Portolano project at the university of washington. In *Proceedings of the Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pages 256–262, Seattle, Washington, August 1999.

[10] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar. Next century challenges: Scalable coordination in sensor networks. In ACM, editor, *Proceedings of the Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pages 263–270, Seattle, Washington, August 1999.

[11] Kevin Fall and Kannan Varadhan. ns: Notes and documentation. Technical report, University of California at Berkeley, LBL, USC/ISI, and Xerox PARC. Technical Report.

[12] Gnutella. http://gnutella.wego.com.

[13] Boris Grndahl. Wireless world meets to lick its wounds. http://www.thestandard.com/article/display/0,1151,22322,00.html?nl=dnt.

[14] Bjorn Gronvall, Assar Westerlund, and Stephen Pink. The design of a multicast-based distributed file system. In *Operating Systems Design and Implementation*, pages 251–264, 1999.

[15] Matthias Grossglauser and David Tse. Mobility increases the capacity of mobile ad-hoc wireless networks. In *IEEE INFOCOM*, Anchorage, Alaska, April 2001.

[16] Shlomo Havlin, Menachem Dishon, James E. Kiefer, and George H. Weiss. Trapping of random walk in two and three dimensions. In *Physical Review Letter*, page 407, 1984.

[17] Barry D. Hughes. *Random Walks and Random Environments*. Oxford Science Publications, 1995.

[18] Tomasz Imielinski, S. Viswanathan, and B. R. Badrinath. Energy efficient indexing on air. In *International conference on Management of Data, ACM SIGMOD*, Minneapolis, 1994.

[19] David Karger M. Frans Kaashoek Hari Balakrishnan Ion Stoica, Robert Morris. Chord: A scalable peer-to-peer lookup service for internet applications. San Diego, California, August 2001.

[20] Joanna Kulik, Wendi Rabiner, and Hari Balakrishnan. Adaptive protocols for information dissemination in wireless sensor networks. pages 24–35, Seattle, Washington, August 1999.

[21] Mark Lindsey, Maria Papadopouli, Francisco Chinchila, and Abshishek Sing. Measurement and analysis of the spatial locality of wireless information and mobility patterns in a canmpus. Technical report TR03-006, Department of Computer Science, The University of North Carolina at Chapel Hill, Chapel Hill, March 2003.

[22] Napster. http://www.napster.com.

[23] The New York Times on the Web. http://www.nytimes.com/adinfo/wireless_audience.html.

[24] A. Ovchinnikov, S. Timashev, and A. Belyy. *Kinetics of Diffusion controlled chemical processes*. Nova Science Publishers, 1989.

[25] T. W. Page, R. G. Guy, J. S. Heidemann, D. Ratner, P. Reiher, A. Goel, G. H. Kuenning, and G. J. Popek. Perspectives on optimistically replicated peer-to-peer filing. *Software—Practice and Experience*, 28(2):155–180, February 1998.

[26] Maria Papadopouli. *Resouce sharing in mobile wireless networks*. PhD thesis, Department of Computer Science, Columbia University, New York,USA, October 2002.

[27] Maria Papadopouli and Henning Schulzrinne. Connection sharing in an ad hoc wireless network among collaborating hosts. In *Proc. International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV)*, pages 169–185, Basking Ridge, New Jersey, June 1999.

[28] Maria Papadopouli and Henning Schulzrinne. Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices. In *ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc)*, Long Beach, California, October 2001.

[29] Maria Papadopouli and Henning Schulzrinne. A performance analysis of 7ds a peer-to-peer data dissemination and prefetching tool for mobile users. *Advances in wired and wireless communications, IEEE Sarnoff Symposium Digest*, March 2001. Best student paper & poster award.

[30] Roman Pichna, Tero Ojanpera, Harri Posti, and Jouni Karppinen. Wireless internet - IMT-2000/wireless LAN interworking. *Journal of Communications and Networking*, 2(1):46–57, March 2000.

[31] DATAMAN project. http://www.cs.rutgers.edu/dataman/, 2001.

[32] RIM. http://www.goamerica.net/coverage/cingular.html.

[33] Elizabeth M. Royer and Charles E. Perkins. Multicast operation of the ad-hoc on-demand distance vector routing protocol. pages 207–218, Seattle, Washington, August 1999.

[34] M. Spreitzer, M. Theimer, K. Petersen, A. Demers, and D. Terry. Dealing with server corruption in weakly consistent, replicated data systems. pages 234–240, Budapest, Hungary, September 1997.

[35] M. Stemm and R. H. Katz. Measuring and reducing energy consumption of network interfaces in hand-held devices. *IEICE Transactions on Communications, vol.E80-B, no.8, p. 1125-31*, E80-B(8):1125–31, 1997.

[36] Vindigo. http://www.vindigo.com/learn_more.html.

[37] Vindigo. http://www.vindigo.com.

[38] Ya Xu, John Heidemann, and Deborah Estrin. Geography-informed energy conservation for ad-hoc routing. Rome, Italy, August 2001.

[39] Tao Ye, H.-Arno Jacobsen, and Randy Katz. Mobile awareness in a wide area wireless network of info-stations. Dallas, Texas, October 1998.