Performance of information discovery and message relaying in mobile ad hoc networks

Maria Papadopouli Henning Schulzrinne*

Abstract

This paper presents 7DS, a novel peer-to-peer resource sharing system. 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 focus on three different facets of cooperation, namely, data sharing, message relaying and network connection sharing. 7DS enables wireless devices to discover, disseminate, relay information among each other to increase the data access. We evaluate via extensive simulations the effectiveness of our system for data dissemination and message relaying among mobile devices with a large number of user mobility scenarios. We model several general data dissemination approaches and investigate the effect of the wireless coverage range, 7DS host density, and cooperation strategy among the mobile hosts as a function of time. We also present a power conservation mechanism that is beneficial, since it increases the power savings, without degrading the data dissemination. Using theory from random walks, 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 humans, but are integrated into physical objects. People access local and general news, traffic or weather reports, sports, maps, guide books, music and video files, games. We classify mobile information access methods into three main categories. The first approach provides "continuous" wireless Internet access; examples include CDPD, 3G wireless, 802.11 and two-way pagers. Currently, this access either have sparse coverage, low cost and high speed (802.11) or have major-cities-only coverage and high cost (Metricom [1]) or have wider coverage, but extremely low rates and high costs (CDPD, RIM). The second approach provides information access via fixed (stationary) information servers or *infostations*. The *infostations* are "information kiosks", for example, located at traffic lights, building entrances and airport lounges. Typically, an *infostation* is a server attached to a data repository and a wireless LAN. When a wireless device is in close proximity to an infostation, it can query the server and access the information. Both of these approaches need an infrastructure. If the wired infrastructure is low-bandwidth, they can be combined by having caches at the base stations. In that case, when the wireless device is within the wireless range of the servers, it can use the cache or the information server to access the data or directly the Internet via the base station. Given the exceedingly expensive license fees attained in recent government auctions of spectrum, the bandwidth expansion route is bound to be expensive. 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. It is unlikely that the density of base stations and infostations will keep pace, due to regulatory, environmental and cost barriers in deploying them. In addition, 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. Thus, we focus on a third approach that does not need the support of any infrastructure (i.e., ad hoc), based on peer-to-peer data sharing among wireless devices. The devices are autonomous and not necessarily connected to the Internet. For the next few years, continuous connectivity to the Internet will not be available at low cost for mobile users roaming a metropolitan area.

In this paper we focus on the challenge of increasing the data availability to users roaming a metropolitan area that experience intermittent connectivity to the Internet. Two characteristics of the environment motivate our approach, namely the high spatial locality of information and queries and the host mobility. The high spatial locality of information results from the type of services we expect a mobile user will run, namely location-dependent, service discovery, news services and collaborative applications. For example, in an urban environment, such as part of Manhattan during rush hours, a platform of a train or an airport, a commercial center, a corporation or a campus, we anticipate that the access patterns

*The authors are with Department of Computer Science, Columbia University in New York, USA. Email: {maria,hgs}@cs.columbia.edu

of the wireless devices will include 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, video games). Also, a subset of the wireless information access possesses loose delay constraints (on the order of minutes). We propose $7DS$[1] as a system that complements the three mobile information access approaches we describe in the previous paragraph. It is an architecture and set of protocols enabling the 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 network connection sharing. In the information sharing facet, peers query, discover and disseminate information. When the network connection sharing is enabled, the system allows a host to act as an application-based gateway and share its connection to the Internet. Hosts buffer their messages, if they do not have an Internet access. When message relaying is enabled, a host forwards its queued messages to another peer. The motivation is to exploit host mobility, better utilize the wireless throughput and reduce the average delay that a message experiences till it reaches the Internet. The hosts also relay all their messages when they gain Internet access (via a gateway or base station [2]).

$7DS$ runs as an application on heterogeneous devices (with different capabilities) that are mobile or stationary. For example, a $7DS-$enabled server can either be dual-homed device connected to the Internet or to a wired infrastructure of other servers or an autonomous server attached to a cache with an access to a wireless LAN. When $7DS$ runs on handheld devices (e.g., PDAs), it will use power conservation and collaboration methods different from the server's ones. In earlier work [2], we investigated the network connection sharing. Here, we focus on information sharing and message relaying.

We distinguish two principal interaction types, namely peer-to-peer (P-P) and server-to-client (S-C). In P-P, 7DS hosts are cooperative with each other. S-C schemes operate in a more asymmetric fashion: there are some cooperative hosts (e.g., $7DS$ servers) that respond to queries and non-cooperative, resource constrained clients (e.g., PDAs). $7DS$ nodes 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. For example, an autonomous $7DS$ server may monitor for frequently requested data, request them from other peers and cache the data locally to serve future queries. The information is typically web pages or any data elements of modest size.

In this new framework, we address some general questions related to the effect of wireless coverage range, density of devices, query mechanism, type of cooperation among the hosts and their power conservation strategy on data dissemination. For example, how fast does information spread in such setting if all nodes are cooperative with each other? How does it change when only a few nodes cooperate (e.g., the $7DS$-enabled servers)? What is the percentage of the nodes that acquire a data item over time? What is the average delay that a node experiences till it receives the data? How does the server-to-client compare to a peer-to-peer approach? How does the wireless coverage range, power conservation, speed, density of devices and servers affect the data dissemination? The investigation of these issues can also give insight for the design of an wireless information infrastructure in a metropolitan area. It appears to be not amenable to an analytical solution except for simplified settings with respect to the node layout, mobility pattern and user interaction pattern. Also, notice that there are no real traces available for the access patterns of mobile, wireless users which would be adequate for our purposes. 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, we also 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 behaviour).

We would like to emphasize the differences of our approach and the setting we consider from that in related works. Ad hoc and sensor networks typically assume a relative high density of devices that results in a connected network (a host can access other hosts via multi-hop routing) [3, 4, 5, 6]. 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 the time taken for one $7DS$ node to come in close proximity to another can be of the order of minutes. As we mentioned, in our setting, peers have different capabilities and cooperation strategy and they are not necessarily all cooperating with each other. Both in ad hoc and sensor networks the emphasis has been on routing protocols.

$7DS$ acquires the data from other peers within its wireless coverage using single-hop broadcast. Due to

[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.

[2] We will use the term *gateway* to refer to a base station or stationary server that provides wireless Internet access.

the highly dynamic environment and the type of information, $7DS$ does not try to establish more permanent caching or service discovery mechanisms. Instead we explore the transient aspect of information dissemination. Note that this setting is orthogonal to the service discovery in the wide area network. In the latter, typically, there is 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 [7]. There is substantial work on hoarding and prefetching [8, 9]. Hoarding allows a mobile device, before disconnecting from the wired network, to prefetch data to increase the user's data availability while she/he remains disconnected and reintegrate upon reconnection. They addressed issues related with the data consistency and targeted in a traditional file system setting. The files exist and the system knows them prior to the disconnection. Finally, an extensive amount of work has been done in the context of infostations. They 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 utilize efficiently the bandwidth of the broadcasting channel [10, 11, 12]. Section 5 discusses related work in greater detail.

In our simulations, we consider variations of the P-P and S-C schemes as well as some hybrid ones. We consider a simple power 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 (and not necessarily equal). We also vary the wireless range of the network interfaces. We evaluate these approaches by measuring the percentage of hosts that acquire the data item as a function of time and their average delay. 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 also evaluate the probability that a message will finally reach the Internet and the impact of message relaying. 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 that a host that queries for 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., FIS). In case of high density of cooperative hosts, the data dissemination using P-P grows even faster (e.g., $1 - e^{-at}$). For example, in P-P, in a setting of 15 hosts with wireless range of 230 m, after 25 minutes, 99% of the users will acquire the data compared to just 42% of the users in the FIS. For the same average delay of 6 minutes, a host using FIS will get the data with a 42% probability, whereas using synchronous P, even in a setting of only five hosts per km^2, this probability is double. For lower transmission power, P-P outperforms FIS by 20% to 70%. In the case of only five hosts, the two approaches differ by 3% to 43%, depending on the transmission power.

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. Second, for 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. Also, these results can assist in the design of wireless data infrastructures.

The contributions of this paper are

1. An overview of the design and implementation of $7DS$, a novel system that enables information dissemination and sharing among mobile hosts in a peer-to-peer fashion.

2. An evaluation via extensive simulations of $7DS$ and the effects of the wireless coverage range, $7DS$ host density, cooperation strategy among the mobile hosts as a function of time.

3. The synchronous power conservation, a beneficial mechanism that increases the power savings substantially, without degrading the data dissemination.

4. 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.

The remainder of this paper is organized as follows. Section 2 gives an overview of the main components of $7DS$. Section 3.1 describes in more detail the P-P and S-C models. Section 3.2 presents simulation results and Section 4 the modeling and analysis of FIS using kinetics of diffusion controlled processes. Finally, in Section 6, we summarize our conclusions and discuss directions for future work.

2 System architecture overview

We assume that the mobile host has a network connection to access the Internet, e.g., via a wireless modem or a base station, and is also capable of communicating with other hosts via a wireless LAN (e.g., IEEE 802.11). $7DS$ runs as an application on mobile hosts and communicates with other $7DS$ participants via a wireless LAN. We focus on information access from the Internet that takes place by retrieving data objects identified by URLs. When such access fails (for example due to the loss of the Internet connection), $7DS$ tries to acquire the data from other $7DS$ peers. Figure 1 (a) illustrates how $7DS$ operates. Mobile host A tries to access a data object (e.g., web page). The local $7DS$ detects that the host has no connection to the Internet and tries to access the page from the peers in close proximity via the wireless LAN. Mobile host D has walked way and cannot listen the query. Both host B and C receive the query. Host C has a copy of the data in its cache and responds to A's query.

$7DS$ uses three types of messages to communicate with other peers: queries, reports and advertisements. A query consists of a set of attributes with their values, such as the URL of the data object and the MAC address of the host that generated it. These two attributes, the URL of the data object and the MAC address, are also used as the query identifier. The system forms queries based on the URL of the data object it tries to acquire. $7DS$ maintains a query list, in which it also includes the URLs the system predicts the user will visit in the next few hours. It multicasts these queries periodically via the wireless LAN to a predefined multicast group. $7DS$ may use different multicast groups for different queries/data items. It determines the appropriate group either by hashing the URL of the requested data item or using some application specific criteria. In order to conserve more power, a host may listen to a subset of these groups depending on the data objects it is willing to share. Both in the prototype and simulations, we consider single-hop multicast, using the "ad-hoc mode" of 802.11. When receiving a query, each $7DS$ peer searches the cache. If a host finds a match, it forms and broadcasts a report. The report is describing the relevant data. After a defined interval, $7DS$ selects from the received reports, the most relevant ones based on application-specific criteria and then it initiates an HTTP GET request.

The advertisements are application-specific messages that announce the presence of $7DS$-enabled server. Power constrained devices use a "passive" mode for participating in the system. In particular, they participate only when the expectation for data availability is high, for example, when they receive an advertisement. A $7DS$-enabled server broadcasts periodically such advertisements with a description of the information or application it supports. A $7DS$ with passive mode enabled host sends the query directly to the server, when it receives an advertisement. We call this "passive" querying, as opposed to active querying that takes place periodically until $7DS$ receives the data.

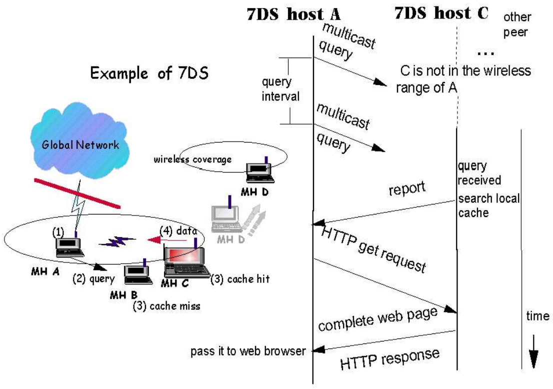


Figure 1: The arrows show the message exchange for the $7DS$ communication. The ellipse denotes the wireless coverage of each host, the shaded signal the wireless LAN and the non-shaded one the (lost) connectivity to the Internet.

We use XML to describe $7DS$ messages. $7DS$ extracts the metadata from the queries received from other peers and performs an attribute-matching search in its local cache. A report includes an identifier that matches the identifier of its corresponding query, and a data description field that contains the relevant information in the local cache of the peer that responds. The report message also consists of some optional attributes with their values, such as the original URL, the time the object (of the data description field) was cached locally, time the original copy was created, its HTML title, size and format, the quality of the wireless transmission (using the signal-to-noise ratio value), the author, language, size and content type of the object. Some of this information is inherently provided by web objects, while others require adding additional (application-specific) meta information. $7DS$ either passes the received reports for display or issues an HTTP GET request automatically (via the web client) using the local URL of the selected report to receive the complete object. A miniature web server is running as a part of $7DS$ which services the HTTP GET requests. The primary information propagation is through the use of caching rather than reliable state maintenance. It is not a goal of the current prototype to resolve inconsistency among copies of a data object. 7DS peers may have several objects matching a single query.

$7DS$ organizes and indexes the cache. Through a GUI, it provides the capability to the user to view, browse and manage the cache. In the current prototype, the content of the cache is displayed in a tree-like structure. We are extending it to support grouping of the cache content to predefined categories and searching tools using the meta-data attributes of the stored objects. The user can set the access permissions for files and directories in the cache and specify the objects to share with other peers. To protect the user privacy, the system only transmits reports or pages that corresponds to public available objects. 7DS can encrypt a private object before its transmission.

$7DS$ trades power for data only when the battery level is above a threshold. Via a battery monitor and a power management protocol, $7DS$ aims to adapt communication to reduce energy consumption during idle periods, when there is low expectation for data or collaboration and when the battery life is below a threshold. Generally, the prediction for data is a hard problem. In order to decide for the data availability, we currently use advertisements from the servers and the traffic in the network. When power conservation is enabled, the mobile host periodically turns off its wireless LAN interface. The system can also adapt its communication with other $7DS$ peers by tuning several "thresholds" in the battery level. Generally, the degree of participation depends on the querying (active or passive, frequency interval) and type of collaboration (data sharing and forwarding support). We also consider a power conservation method we consider alternates from the *on* state of the network interface to the *off* state. During the interval that the network interface is on, $7DS$ communicates with the other hosts by sending queries, forwarding or receiving reports or data. The mobile host broadcasts a query at each *on* interval till it receives the data.

An important feature of our architecture is its easy deployment. The system displays both the reports as well as the complete data information using any browser. It is transparent to wired and wireless networks as well as to different information providers that participate in the system. Also, $7DS$ is flexible to support different applications and able to form queries and application-specific criteria for the selection of the appropriate cached copies as long these applications access the data using URLs. Users only need to install the software and $7DS$ configures itself with minimal manual intervention; the system does not require any registration for data distribution. The system is resilient to failures and inconsistencies that occur in this dynamic environment. $7DS$ is resource aware and tries to utilize the constrained resources efficiently.

The communication and caching components are simple. However, the paradigm is powerful and it triggers many challenging issues. It allows us to address some general questions on the performance of data dissemination among mobile devices which is the focus of this paper. Also, we currently investigate mechanisms to provide security and incentives to cooperate using a token-based approach [13], lightweight offline micropayments methods [14] and a rendezvous-based approach. The main idea of the rendezvous-based approach is that a group of cooperative hosts decides on a time interval to communicate (potentially with encrypted messages). They can turn their network interface on and start participating in 7DS only during the decided time interval. This is also related to the synchronization power conservation we describe in Section 3. The prototype is written in Java and uses the Glimpse search engine [15]. We are in the process of implementing it on Windows CE. Details on the implementation and discussion can be found at [16].

3 Performance Evaluation

3.1 System models and operation modes

$7DS$ can operate in different modes that depend on the cooperation strategy among peers (data sharing, forwarding), power conservation and query mechanism (active, passive querying). To investigate its performance and 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. 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 server. 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 rebroadcasted another message during the last 10 s. The last condition is a simple mechanism for preventing flooding in the network. We separate the S-C schemes into the "straight" S-C without any cooperation among clients (namely, FIS and MIS) and some hybrid ones with cooperative clients. In 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 power conservation mechanism. They receive data only from the server. The hybrid schemes are with passive querying enabled and fixed server. As we mentioned in Section 2, in passive querying mode, the server sends an advertisement every 10 sec. Hosts send queries upon the receipt of an advertisement.

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.

Let us describe the main motivations for the comparisons we make in the remaining section. The P-P vs. straight S-C comparison is to understand the effect of the cooperation among mobile peers. The P-P and MIS vs. FIS shows how mobility affects data dissemination. In particular, the MIS vs. FIS comparison focus exactly on the effect of server mobility on data dissemination.

3.2 Model assumptions

Nodes move in a 1000 m x 1000 m area according to the random waypoint mobility model [3]. This random walk-based model is frequently used for individual (pedestrian) movement [3, 4, 5]. 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 point for each segment are distributed randomly across the area. Each node is moving to its destination with a constant speed uniformly selected from (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 (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 selected from the *on* interval. The asynchronous mode is the default power conservation method. We explicitly denote the schemes with synchronous mode enabled with the word "sync". In schemes with no power conservation, the *off* interval is equal to 0 and the *on* and query interval are the same. The exchange of queries, reports and advertisements takes place during the on interval. Generally, the transmission of the complete data object (for example, web page) is scheduled separately. For example, the dataholder may select a time ("rendezvous point" in which the HTTP transmission takes place) and include it in the report message. At that time, both the querier and dataholder set their network interface on and the querier initiates the HTTP get request (as described in Section 2). In the simulations, we concentrate only on the exchange of $7DS$ queries, reports and advertisements. A cooperative dataholder responds to a query by sending the data item in the report. 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 a report for this data item, it becomes dataholder. This simplification is reasonable in order to investigate the dominant parameters on data dissemination.

A *scenario* (file) "defines" the topology and movement of each host that participates in an experiment. We consider different number of hosts in the area. Later in this section, we scale the area and vary the density of the hosts and their wireless coverage. We would like to 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 could be interested to get the local and general news using PDAs or other wireless devices. A density of 25 hosts per km^2 corresponds to very popular data whereas a density of five hosts per km^2 corresponds to a more typical data object [17]. 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, that the server is stationary. We run simulations using these scenarios, for the different schemes of Table 1. The wireless LAN is modeled as an 802.11 network interface. We use the ns-2 simulator [18] with the mobility and wireless extensions [19]. 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 and 57.5 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 power.

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.

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

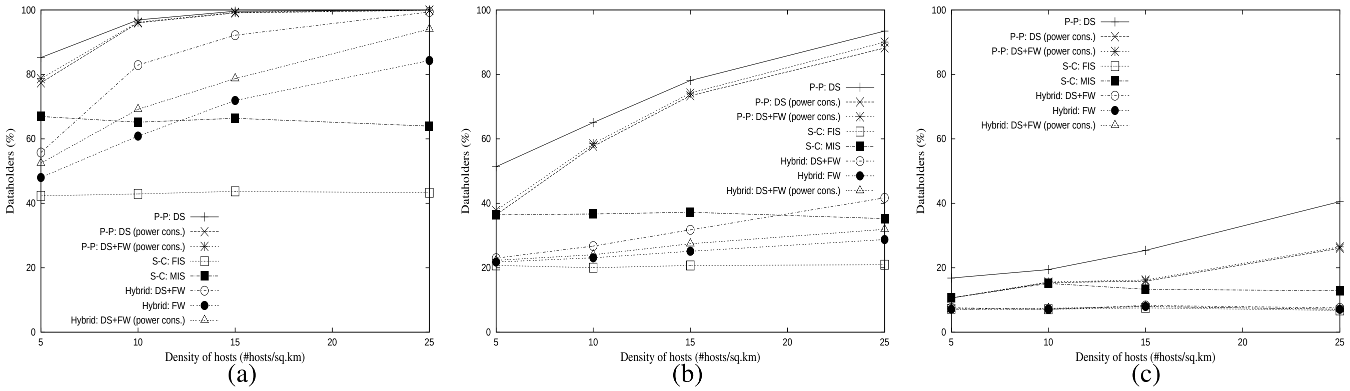


Figure 2: Percentage of dataholders after 25 minutes for high, medium and low transmission power, respectively. The query interval is 15 sec.

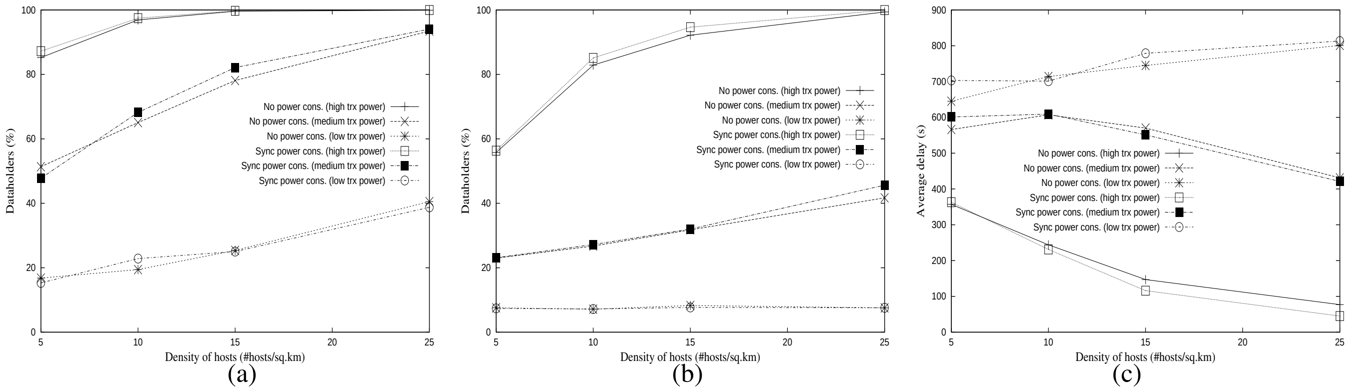
not include the node that has the data at the beginning of the simulation. We also compute the average delay until a mobile host receives the information from the time it sends the first query. We run the 300 generated scenarios for each test and computed the average of the percentage of hosts that *become* dataholders by the end of each test. 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 hosts start querying in average after 90 sec from the beginning of the simulation. 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.

Figures 2 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 (a), $7DS$ 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 during the 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 (a), 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). The impact of data sharing among peers is also apparent in hybrid schemes. In particular, the hybrid vs. S-C schemes for density of ten or more hosts per km^2 and medium or high transmission power.

Notice 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 as the following occurs: There is 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 till the end of the test. We would like to emphasize that this is true also for smaller simulation times, starting from 150 sec (just a few seconds after the hosts start querying). Due to the lack of space, we do not include the graphs that illustrate it; the reader is referred to [17] for more information. In a more dense setting of mobile hosts that forward messages independent on their data interests, we expect forwarding to have a higher impact. Forwarding without data sharing results in a performance improvement. For example, in Figures 2 hybrid schemes with forwarding enabled outperforms FIS by 4%-40% depending on transmission power.

As we expect for both FIS and MIS, their performance remains constant as the number of hosts increases, since a data exchange takes place only when a querier is in proximity to the server. In addition, notice that in Figures 2 (a) ,(b) and (c) MIS outperforms FIS by approximately 22%, 16%, 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.

Measurements of the power 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 [20]. Using the asynchronous power conservation mech-



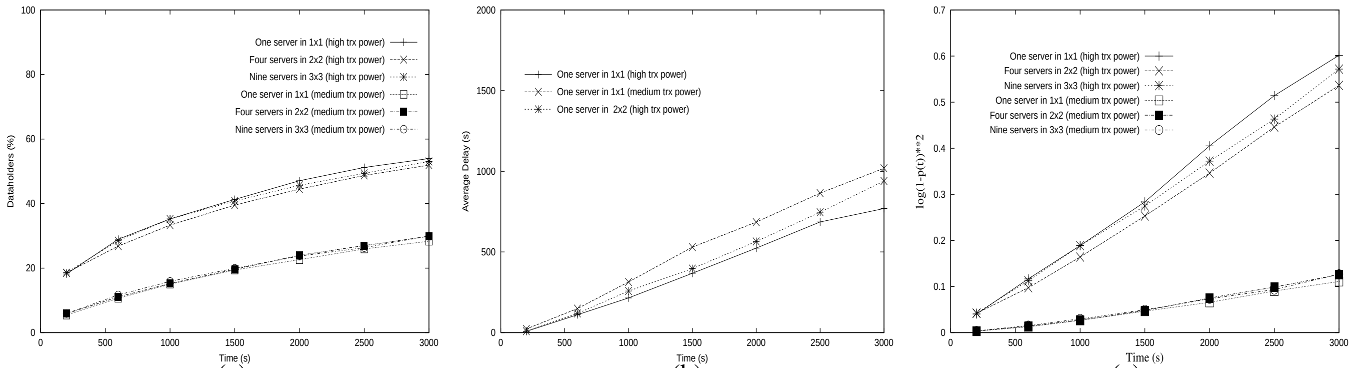
(a) (b) (c)

Figure 3: Figures (a) and (b) illustrate the impact of synchronous mode on data dissemination in a P-P with data sharing and a Hybrid with data sharing and forwarding scheme, respectively. Figure (c) illustrates 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.

anism we described with the *on* and *off* intervals to be equal there is 50% power savings, since the network interface is on only half the time. As Figures 2 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 power savings. However, with smaller intervals, the degradation of data dissemination is larger. To prevent this degradation, we enable the synchronous mode. As we show in Figure 3, when the synchronous mode is enabled, even with a small *on* interval, the power conservation does not cause any degradation of the data dissemination. More specifically, Figure 3 (a) illustrates P-P schemes with data sharing and Figure 3 (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 at the remaining time (13.5 sec) switches *off*. In an ideal setting without packet losses and need for retransmission, the number of messages exchanged in the P-P schemes without power conservation and the ones with synchronous power conservation are the same. Therefore, the power spendings due to message receiving/sending is the same, whereas the period in the idle state is reduced (the network interface is on only during 10% of the time). The synchronous mode may result in a 90% reduction in power dissipation. In general, hosts may query for different data items. In very dense settings retransmissions and packet losses may result in further power spendings. It 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.

In [21], we investigated the performance of the system as a function of the query interval using the asynchronous power conservation method (i.e., the on interval is half the query interval and not synchronized). The degradation in the FIS performance is relative 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). 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 power 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 power 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.

As we mentioned earlier we evaluate the average delay a host experiences till it receives the data since its first query. For each test, we compute the average delay of the nodes that acquired the data by the end of 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 data sharing and no power conservation, for high transmission power (Figure 3 (c)), the average delay is as high as 6 minutes for sparse networks and reaches 77 sec for dense networks (Figure 3 (c)). 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) and for high transmission power is 6 minutes, whereas for low transmission power it reaches 9 minutes. So, (sync) P-P with



(a) (b) (c)

Figure 5: Performance of the fixed information server (FIS) schemes as a function of the simulation time.

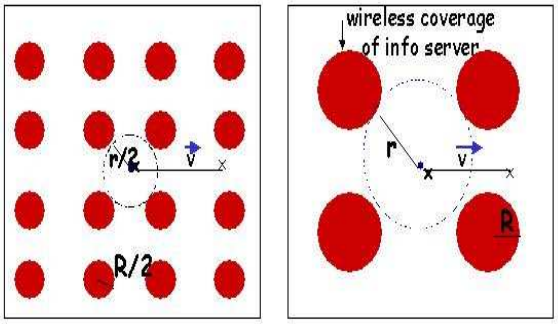


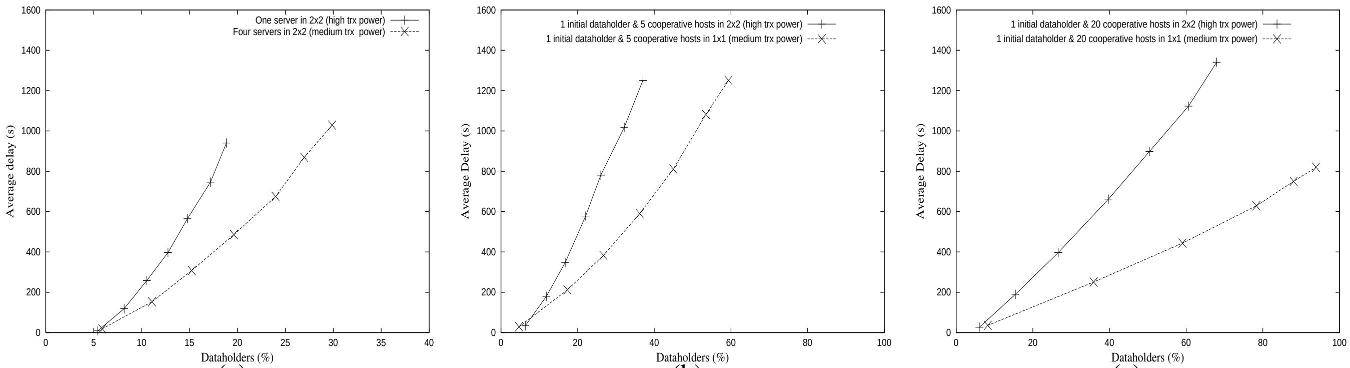
Figure 4: Scalability property. The dark disks depict the wireless coverage of a host. For fixed wireless coverage, the larger the density of cooperative hosts, the more efficient the data dissemination.

data sharing, even in the case of low density of hosts, 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 compare P-P in Figure 3 (a) and (c) and FIS in Figures 5 (a) and (b).

Figures 6 compare FIS and P-P with data sharing and no power enabled. Note that for these Figures we have combined the simulation results on the probability a host acquires the data and the average delay it experiences as functions of time. For example, for the case of one server in a 2km x 2km area with high transmission power (Figure 6 (a)), we use the results of Figures 5 (a) and (b). 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. With FIS a 40% probability of acquiring data corresponds to an average delay of 6 minutes whereas using (sync) P-P this probability doubles (even for a low density cooperative host setting). For a higher average delay of 10 minutes, 85% hosts will acquire the data using P-P and 50% using FIS. 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 and P-P, respectively.

Let us now discuss the scaling properties and generalize our performance results. First, we focus on expanding the area but keep the movement pattern the same. Both in 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 is an indication that our simulation results are robust. For example, Figure 5 (a) shows this scaling property in FIS. 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 5 (c) 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 power conservation) $p(t)$ grows faster than in FIS and our simulation results indicate that the P-P with data sharing and power conservation can be approximated by the function $1 - e^{-at}$, especially, for not very dense settings (e.g., with less than 20 hosts per km^2 transmitting with high or medium power). 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 design decision. Figure 4 illustrates two possible deployments of servers with the same wireless data coverage (assuming an ideal transmission



(a) (b) (c)

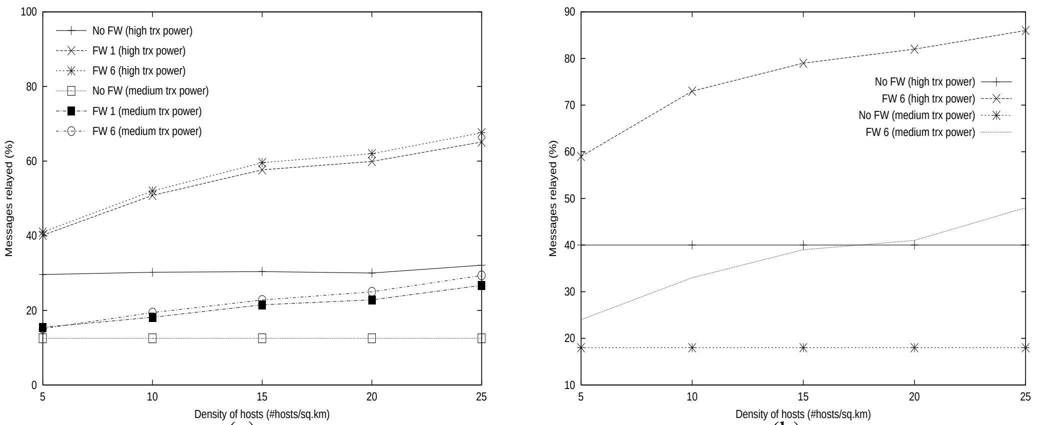
Figure 6: The average delay to receive the data as a function of the probability to acquire it. Figure (a) illustrates FIS whereas Figures (b) and (c) P-P with data sharing schemes.

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 the two settings, we assume the same mobility pattern. Figures 4 depict 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 4 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 with the right one (in terms of area, transmission power of the servers and servers density), but with the hosts moving faster. Therefore, the probability a host will get into the coverage of a server increases.

Figure 6 (a) compares two FIS settings. The first includes one server in a 2km x 2km area with high transmission power and the latter four servers in a 2km x 2km area with medium transmission power. The case with higher density of servers perform better. For example, a 20% probability of acquiring the data, the FIS scheme with higher density of servers corresponds to an average delay of 500 s. For the same wireless coverage, but lower density of servers, this probability corresponds to a double average delay. Figures 6 (b) and (c) 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.

As we discussed in the Section 1, the message relaying is another facet of cooperation among the mobile hosts. We assume hosts generate messages and buffer them locally, when there is no Internet access. When a host gains access (i.e., it is in the wireless coverage of a gateway), it relays these messages to the gateway. A host may relay its own messages to another 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 till it reaches a gateway. To control the message explosion, we impose two restrictions. First, a host relays all queued messages to a gateway, but only its own messages to another peer. That is, a given message reaches the Internet via at most two hops. Second, 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 only once the same message to a host. The gateways advertize periodically their presence. Upon the receipt of such advertisement, a host forwards all the queued up messages to the gateway.

Figures 7 (a) and (b) illustrate the probability that a host will reach the Internet as a function of the host density. We assume one gateway per km^2 area. Figure 7 (a) shows the percentage of the generated messages at each host that reach the Internet and the impact of forwarding to relay nodes. We assume that the hosts generate messages with constant rate (a message at 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 reaches the Internet. Notice that forwarding a message to more than one relay nodes does not improve substantially the performance. Figure



(a) (b)

Figure 7: Figures illustrate the performance of message relaying after 25 minutes in an area with one gateway/km^2.

7 (b) illustrates the probability that a message will reach the Internet after 25 minutes. We would like to mention that when there is no forwarding, the probability that a message will reach a gateway is the same with the probability that the host will reach a gateway. Essentially, this probability is the same with 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 7 (a) this probability increases when forwarding is enabled. In a different setting with a very low transmission power, that corresponds to 8 meters, and with high density of hosts (such as 100 hosts/km^2), after 2.5 hours, 5% of the total generated messages will reach a gateway directly and 38% of them via another relay host.

To summarize the simulation results,

- P-P schemes outperform S-C schemes. The 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, the data dissemination using P-P grows even faster. Generally, their difference becomes more prominent in cases of lower transmission power (medium and low) 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 three times higher than in FIS. In some of the cases, their difference in average delay is negligible, in other cases FIS has lower average delay (with a maximum difference of 100 sec).

- Forwarding (i.e., rebroadcasting $7DS$ messages upon their receipt) in addition to data sharing does not result in any performance improvements.

- The query interval has negligible effect on S-C schemes.

- The synchronous power conservation method is beneficial. It increases the power savings without degrading the data dissemination.

- The performance remains the same when we scale up the area, but keep the density of cooperative dataholders and transmission range fixed.

- Dominant parameters are the density of cooperative hosts and their wireless coverage density. Also, their mobility can contribute to higher data dissemination. For a given wireless coverage density, the larger the density of cooperative hosts, the better the performance. 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.

- Message relaying can increase the data access by exploiting the host mobility.

4 Data dissemination as a diffusion controlled process

This section discusses our initial efforts to study analytically the data dissemination and generalize further our results. This section 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 [22] and the kinetics of diffusion-controlled chemical processes [23]. 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 is showed that the large-n behaviour of the survival probability

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

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

As we mentioned in Section 3.1, in FIS the 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 with respect to the model is considered "trapped". Figure 8 illustrates the the analytical and simulation results on data dissemination as a function of time. The analytical results on Trap model are derived from Eq. 1 (i.e., Rosenstock's trapping model) for high and medium wireless coverage.
We define q as $\pi R^2 N_{servers}/A$ and use an R equal to 230 m and 115 m for high and medium wireless coverage, respectively. For the simulation results on FIS in

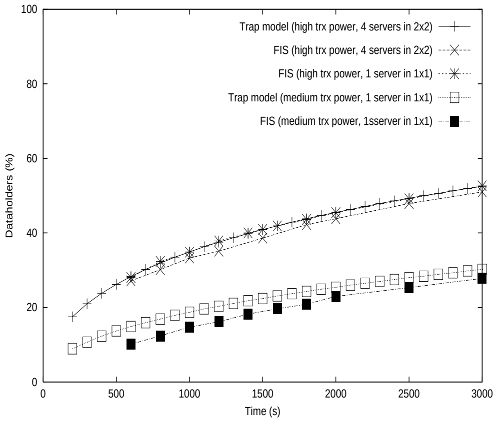


Figure 8: Simulation and analytical Trap model results on FIS.

Figure 8, we use the FIS simulations we described in Section 3. Note that, using Eq. 1 the $1 - \phi_n$ expresses the fraction of hosts that acquire the data at time n. As Figure 8 illustrates, our simulations are consistent with Eq. 1 (in two dimensions) for α equal to 0.021. That is, using the 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.

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 currently explore 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.

Generally, the dominant contribution to the survival probability comes from particles which happen to be in large trap-free regions. Thus, as long as the trap radius is finite, it can be neglected compared to the radius of the optimal trap-free region, and its dependence on the survival probability is asymptotically negligible [24]. However, there are settings and experimentally relevant time scales, the effect of the trap radius is important, but also not easy to calculate. These calculations are general much harder than the asymptotic estimate. As we describe in Section 3, the simulation results indicate the transient behaviour of the data dissemination ("survival probability") and the impact of the wireless coverage ("trap radius").

5 Related work

Napster [25] and Gnutella [26] are two systems that explore the cooperation among hosts and enable data sharing among users in a fixed wired network. The first is focus in sharing music files, whereas the latter for any type of files. In contrast to Gnutella, $7DS$ does not need to discover its neighbors or maintain connections with them, but only multicast its queries to a well known group. Also, note that 7DS (in the default mode) restricts the query propagation to the wireless LAN. Unlike Napster $7DS$ operates in a distributed fashion without the need of any central indexing server. Moreover, Napster requires user intervention and effort for uploading files, whereas $7DS$ does this automatically.
Infostations have first been mentioned by Imielinski in the DataMan project [11]. 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. Imielinski *et al* [10] 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 index or hashing based access to the data transmitted over the

wireless can result in significant improvement in battery utilization. Barbara *et al* [27] propose and study a taxonomy of difference cache invalidation strategies and study the impact of client's disconnection times on their performance.

Caching and prefetching have been successfully employed to alleviate user perceived latencies and there has been extensive research. In the context of mobile users, hoarding is a similar technique to prefetching to improve the data availability (for users that experience intermittent connectivity) [8, 9]. In a similar context to ours, prefetching targeted for mobile users in a wide-area wireless network has been used in [28]. Tao Ye, *et al* [28] assume an infostation deployment. 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 an infostation coverage 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 wireless network. Devices are using 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 as compared to a traditional wide-area wireless network. For that, 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 (that corresponds to low infostation density). As we explained in the Section 1, the focus of this paper 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 (like various mobility patterns, power conservation methods and querying schemes) that have not been investigated in [28]. We should mention that their qualitative result, that having many infostations that cover small ranges is a more optimal topology than having few infostations that cover large ranges is consistent with ours.

Kravets, *et al* [29] present an innovative transport level protocol that achieves power savings by selectively choosing short periods of time to suspend communication and shut down the communication device. It queues data for future delivery during periods of communication suspension, and decides predicting when to restart communication. This work motivated us to consider schemes for predicting high data availability in our setting to power on the communication device and start $7DS$. In Section 3, we discuss this in more detail.

There is substantial peer-to-peer work in the file system and OS literature that is relevant, including the Ficus [30], JetFile [31], Bayou [32] projects. All of them are replicated storage systems based on the peer-to-peer architecture. Ficus is a distributed file system aiming to 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, but 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. Our system is targeted in 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.

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, [33] presents 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 the dissipate. It 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. A more theoretical work [34] assumes a system where the nodes are placed on a line. They present an optimal algorithm for broadcasting and compute the expected number of time steps required for it to complete. More theoretical studies on information dissemination have used percolation theory [35] or epidemic models. In percolation theory, the nodes are typically fixed on a lattice and therefore these models are not adequate for our setting.

Grossglauser *et al* [36] show how the mobility can increase the capacity of mobile ad-hoc wireless networks. They evaluate the average per session throughput and asymptotic performance. In our work we focus on the transient behaviour of the message relaying and the impact of various parameters.

6 Conclusions and future work

In this paper, we presented $7DS$, a new peer-to-peer data sharing system. $7DS$ is an architecture enabling the exchange of data among peers that are not necessarily connected to the Internet. It anticipates the information needs of users and fulfills them by searching for information among peers. To assess the efficiency of information dissemination via $7DS$, and also investigate the effect of the wireless coverage range, network size, query mechanism, cooperation strategy among the mobile hosts and power conservation over time, we performed an extensive simulation-based study. This study involves two main data dissemination approaches, namely peer-to-peer (P-P) and server-to-client (S-C). We measured the percentage of data holders and also the average delay until a querier gets the data item since it sent the first query. We found that the density of cooperative hosts and their mobility are dominant parameters for data dissemination. We studied the scaling effect in our simulations and found that the performance remains the same when scale the area, but keep the density of cooperative dataholders and power transmission fixed. Also, the simulation results indicate that for fixed wireless coverage density, the larger the density of cooperative hosts, the better the performance. For example, both in FIS and P-P, 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. The combination of power conservation and synchronization of the network interface *on* periods is beneficial for $7DS$ users, since it increases substantially the power savings without degrading data dissemination. For example, for an average delay of 6 minutes, a host using FIS gets the data with a 40% probability. Whereas with (sync) P-P with data sharing, even in the case of low density (five hosts per km^2 interested in the same data), this probability is double.

It is important to understand the impact of mobility pattern on data dissemination. For the same mobility pattern, when only the speed changes, we can compute the new probability of acquiring the data as a function of time using the previous results. In cases of group mobility, we also expect the impact of cooperation and the difference of FIS and P-P schemes to increase. For user mobility, most of the studies on ad hoc routing protocols use random-walk based models. The randway model appears reasonable for a conference setting [37]. We do not expect substantial changes in the performance of FIS vs. P-P using other random-walk based models. For example, the analytical results in Section 4 use random walk (that differs from the randway model in the time scale of the displacement step). However, it is difficult to speculate the performance of the system for other mobility patterns. Unfortunately, there are not many traces available of actual data access patterns of mobile, wireless users or realistic models for different scenarios. In earlier work [38], we simulated a baseline scenario of information sharing in a subway, where users enter the platform, ride a subway car, reach their destination stop and leave. However, the mobility model was oversimplified and the performance of data dissemination tied to that setting. The development of realistic and general models for the performance analysis of mobile, ad hoc networking systems remains imperative.

An attractive feature for $7DS$ is a mechanism that would indicate the appropriate interaction (P-P, S-C with active or passive querying) based on several parameters, such as data availability prediction, cooperative users in close proximity and battery level. Advertisement messages from servers or other cooperative hosts, traffic measurements and location information (of gateways or servers) can provide hints. We plan to investigate how they can be used to improve the power utilization and the performance of data dissemination and message relaying. It is also part of our future work to investigate the impact of on interval of the rendezvous-based power conservation, considering packet losses and retransmissions.

Currently, we deploy the 7DS on the campus and integrate it with a tour guide and an academic news notification system. We would like to get a better understanding of the applications for mobile devices with respect to delay and privacy requirements, information and query locality, percentage of users that actually read/use the data and adequate replacement policies. Design issues related to the security and user privacy are also part of the ongoing effort.

References

[1] Metricom, “The Ricochet wireless network overview,” 1999.

[2] M. Papadopouli and H. 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)*, (Basking Ridge, New Jersey), pp. 169–185, June 1999.

[3] J. Broch, D. Maltz, D. Johnson, Y.-C. Hu, and J. 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), Oct. 1998.

[4] E. M. Royer and C. E. Perkins, "Multicast operation of the ad-hoc on-demand distance vector routing protocol," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)* (ACM, ed.), (Seattle, Washington), pp. 207–218, Aug. 1999.

[5] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed energy conservation for ad-hoc routing," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, (Rome, Italy), Aug. 2001.

[6] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)* (ACM, ed.), (Seattle, Washington), pp. 263–270, Aug. 1999.

[7] P. Castro, B. Greenstein, R. Muntz, and M. Papadopouli, "Locating application data across service discovery domain," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, (Rome, Italy), Aug. 2001.

[8] J. J. Kistler and M. Satyanarayanan, "Disconnected operation in the coda file system," *Thirteenth ACM Symposium on Operating Systems Principles*, vol. 25, no. 5, pp. 213–225, 1991.

[9] H. Lei and D. Duchamp, "An analytical approach to file prefetching," in *USENIX Annual Technical Conference*, (Anaheim, CA), Jan. 1997.

[10] T. Imielinski, S. Viswanathan, and B. R. Badrinath, "Energy efficient indexing on air," in *International conference on Management of Data, ACM SIGMOD*, (Minneapolis), 1994.

[11] DATAMAN. http://www.cs.rutgers.edu/dataman/.

[12] D. Barbara and T. Imielinski, "Sleepers and workaholics: Caching strategies in mobile environments," 1994.

[13] J. P. H. L. Buttyan, "Nuglets: a virtual currency to stimulate cooperation in self-organized mobile ad hoc networks," Tech. Rep. DSC/2001/001, Swiss Federal Institute of Technology, Lausanne, Jan. 2001.

[14] M. Blaze, J. Ioannidis, and A. D. Keromytis, "Offline micropayments without trusted hardware," in *Proceedings of Financial Cryptography*, (Cayman Islands), Feb. 2001.

[15] Glimpse. http://www.webglimpse.org.

[16] 7DS. http://www.cs.columbia.edu/˜maria/7ds.

[17] M. Papadopouli and H. Schulzrinne, "Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices," Tech. Rep. CUCS-005-01, Columbia University, 2001.

[18] K. Fall and K. Varadhan, "ns: Notes and documentation," tech. rep., Berkeley University, 1998. Technical Report.

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

[20] 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*, vol. E80-B, no. 8, pp. 1125–31, 1997.

[21] M. Papadopouli and H. Schulzrinne, "Effects of power conservation, wireless coverage and cooperation on dat a dissemination among mobile devices," in *ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, (Long Beach, California), Oct. 2001.

[22] B. D. Hughes, *Random Walks and Random Environments*. Oxford Science Publications, 1995.

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

[24] S. Redner, *A guide to first-passage processes*. Cambridge University Press, 2001.

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

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

[27] D. Barbara and T. Imielinski, "Sleepers and workaholics: Caching strategies in mobile environments," 1994.

[28] T. Ye, H.-A. Jacobsen, and R. Katz, "Mobile awareness in a wide area wireless network of infostations," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, (Dallas, Texas), Oct. 1998.

[29] R. Kravets and P. Krishnan, "Power management techniques for mobile communication," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, (Dallas, Texas), pp. 157–168, Oct. 1998.

[30] 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*, vol. 28, pp. 155–180, February 1998.

[31] B. Gronvall, A. Westerlund, and S. Pink, "The design of a multicast-based distributed file system," in *Operating Systems Design and Implementation*, pp. 251–264, 1999.

[32] M. Spreitzer, M. Theimer, K. Petersen, A. Demers, and D. Terry, "Dealing with server corruption in weakly consistent, replicated data systems," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, (Budapest, Hungary), pp. 234–240, Sept. 1997.

[33] J. Kulik, W. Rabiner, and H. Balakrishnan, "Adaptive protocols for information dissemination in wireless sensor networks," in *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)* (ACM, ed.), (Seattle, Washington), pp. 24–35, Aug. 1999.

[34] K. Ravishankar and S. Singh, "Broadcasting on [0,L]," 1994.

[35] R. Durrett, *Lecture notes on particle systems and percolation*. Pacific Grove, CA, 1988.

[36] M. Grossglauser and D. Tse, "Mobility increases the capacity of mobile ad-hoc wireless networks," in *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, (Anchorage, Alaska), Apr. 2001.

[37] T. Camp, J. Boleng, and V. Davies, "Mobility models for ad hoc network simulations," 2001.

[38] M. Papadopouli and H. Schulzrinne, "Seven degrees of separation in mobile ad hoc networks," in *Proceedings of the IEEE Conference on Global Communications (GLOBECOM)*, (San Francisco, California), Nov. 2000.