A Performance Analysis of $7DS$:
A Peer-to-Peer Data Dissemination and Prefetching Tool for Mobile Users

Maria Papadopouli Henning Schulzrinne
Department of Computer Science
Columbia University
New York, NY 10027
e-mail: {maria,hgs}@cs.columbia.edu

Abstract

This paper presents 7DS, a novel peer-to-peer data sharing system. Peers can be either mobile or stationary 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. It runs as an application complementary to other data access approaches (such as via base stations or info-stations). It anticipates the information needs of users and fulfills them by searching for information among peers. We evaluate via extensive simulations the effectiveness of our system in data dissemination and prefetching for mobile devices. Also, we investigate the effect of the wireless coverage range, network size, query mechanism, cooperation strategy among the mobile hosts and power conservation with a very large number user mobility scenarios.

1 Introduction

New services offered by telecommunication companies and content providers which expose consumers to spatial information have already arrived. Such services provide news, traffic or weather reports, maps, guide books, music and video files, games, notifications about changes in environmental conditions [1] and points of interest. Access to information will become as important as voice communications for wireless roaming through metropolitan areas [2, 3, 4]. We separate mobile information access methods in to three main categories. The first approach provides "continuous", wireless Internet access, such as CDPD, 3G wireless, 802.11 and two-way pagers. The second approach provides information access via fixed (stationary) *information servers* or info-stations in local geographic proximity [5, 6, 7]. The *information servers* are "information kiosks", located at traffic lights, building entrances and airport lounges. These two approaches are using an infrastructure. If the wired infrastructure is low-bandwidth, they can be combined by having caches at the base stations. The third approach is without the support of any infrastructure (i.e., ad hoc) based on peer-to-peer data sharing among the mobile, wireless devices.

Current wireless Internet access either have sparse coverage and low-cost and high speed (802.11) or have major-cities-only coverage, high cost (Metricom [8]) or have wider coverage, but extremely low rates and high costs (CDPD, RIM). Also, info-stations offer high speed but discontinuous coverage. 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 info-stations coupled to the associated high speed wired infrastructure cost is forbidding. For the next few years, continuous connectivity to the Internet will not be universally available at low cost.

We propose $7DS$, a new peer-to-peer solution that complements these existing approaches. Peers can be either mobile or stationary. $7DS$[1] is an architecture, set of protocols and implementation enabling the exchange of data among peers that are not necessarily connected to the Internet. $7DS$ runs as an application and operates in two modes, namely prefetch and on-demand. In the prefetch mode, it anticipates information needs of users. In the on-demand mode, it searches for information among peers when the user searches for information or requests a URL. $7DS$ can also run on the top of location-dependent services and prefetch data on behalf of them, when the host cannot access data via base stations or *information server*. $7DS$ is particular useful for information that do not change rapidly (and therefore do not require continuous Internet connectivity). In the context of these services, spatial locality implies locality in the data that mobile users want to access and in their access patterns. We anticipate that there will be a high probability that the data a mobile host queries, can be found in the cache of another mobile host in close geographic proximity. The system exploits the high spatial locality of information in pervasive computing environments and also the fact that mobile users are likely to be more flexible in their information tastes, (media) quality and information accuracy requirements.

In such a network of mobile hosts, each new device contributes to an ever denser web of communication, where data can move from subway rider to subway rider, among anonymous persons meeting each other in the streets, in the hallways of an office building, at a conference, a public area (such as a train or airport platform), in a battlefield situation or in a disaster recovery area with rescue teams. $7DS$ allows users to browse the content of the cache of peer that has been made accessible to it. Participants

---

[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.

in $7DS$ obtain URLs, web pages, or any application specific data of modest size, cache them and exchange them with others who are interested in them. Each device maintains a cache containing information items received by pervasive computing devices.

The contributions of this paper is the overview of a novel peer-to-peer information sharing system and its performance analysis via simulations. We evaluate via extensive simulations the effectiveness of our system in data dissemination and prefetching for mobile devices. We also investigate the effect of the wireless coverage range, network size, query mechanism, cooperation strategy among the mobile hosts and power conservation with a very large number (300) user mobility scenarios. To assess the efficiency of information dissemination via $7DS$, we performed a simulation-based study of the two data dissemination approaches, namely 7DS-only and FIS-based (fixed information server). In particular, we model several variants of 7DS-only and FIS-based depending on the collaboration strategy of the hosts, their querying scheme and whether or not they use power conservation. We also vary the transmission power of the mobile hosts to simulate different range network interfaces, that cover a range from 55 m to 230 m.

We evaluate these approaches by measuring the percentage of hosts that acquire the data at the end of the simulation. In all the cases we consider, 7DS-only outperforms FIS-based. Their difference in the percentage of the users that will have the data after 25 minutes becomes prominent. For example, when there are 25 hosts transmitting with high power, in 7DS-only cases, 99.9% of the users will acquire the data compared to just 43% of the users in the FIS (baseline case). For lower transmission powers, 7DS-only cases outperform FIS-based cases by around 20% to 70%. In the case of only five hosts, the two approaches differ by 3% to 43%, depending on the transmission power.

This system raises several interesting research issues; some are theoretical, such as study of the data propagation in this environment using epidemic models, percolation theory and game theory, and others design-related. In [9] we include more results from our performance analysis via simulations as well as a simple analytical study of data propagation. In that paper, we also discuss the design and an implementation of $7DS$. We have implemented a prototype in Linux and are currently in the process of implementing it on Windows CE. The remainder of this paper is organized as follows. Section 2 describes in more detail the 7DS-only and FIS-based models. Section 3 presents simulation results and in Section 4, we discuss briefly related work. Finally, we present conclusions and describe future research directions in Section 5.

2 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). In active querying mode, $7DS$ periodically checks if there are user local requests queued up locally and broadcasts them. In passive mode, an advertisement "triggers" the querying. In data sharing mode, when a $7DS$ peer receives a query, it checks if it has the data in its cache and in the case of cache hit, broadcasts it. In forwarding mode, upon the receipt of a query or data, the $7DS$ peer forwards (rebroadcasts) it to extend the "coverage". To investigate its performance and in particular the effect of transmission power and the different modes of operation on data distribution, we need to introduce two main schemes, 7DS-only, FIS-based with active and passive querying along with their variants.

7DS-only

Mobile hosts are running $7DS$ and there is no fixed (stationary) *information server*. All nodes have the active querying and data sharing modes enabled. Under this scheme, we perform three sets of simulations:

1. $7DS$ nodes with power conservation enabled (P). We describe the power conservation in the following paragraphs.
2. $7DS$ nodes with power conservation and apart from data sharing, forwarding is also enabled (FW). Upon the receipt of a query or data, $7DS$ peers rebroadcast it, if they have not rebroadcasted another message during the last 10 s. FW differs from P in that P supports data sharing, but it does not rebroadcast received data and queries.
3. $7DS$ nodes without any power conservation (NP).

FIS with active querying

There is a fixed (stationary) *information server*. Mobile hosts run $7DS$ and have enabled the active querying mode. However, there is no collaboration among the mobile hosts (neither data sharing nor forwarding). Hosts receive data only from the server. They do not use any power conservation mechanism (FIS).

FIS with passive querying

In passive querying FIS, there is also a fixed *information server*, but $7DS$ on mobile hosts are in passive querying mode. In the passive querying mode, $7DS$ broadcasts a query only upon the receipt of an advertisement. In simulations, we consider that the advertisement is the *information server* beacon or advertisement that is sent every 10 s. $7DS$ broadcasts the locally (queued) queries only upon the receipt of such messages. Generally, the motivation for the advertisement is to notify $7DS$ peers for the presence of a $7DS$-enabled host with information and therefore for higher data availability. Devices with limited power capabilities use the "passive" mode. We consider several variations of this scheme based on the cooperation strategy and power conservation:

1. $7DS$ nodes with both data sharing and forwarding enabled and depending whether or not power conservation is used (FIS-P) and (FIS-NP), respectively.
2. $7DS$ nodes without power conservation and with forwarding enabled (FIS-NDS). In FIS-NDS peers do not share their cache, but only rebroadcasts queries, data and beacons they receive.

Simulation Set	Query	Power Cons.	Cooperation Strategy
P(NP)	Active	yes(no)	data sharing
FW	Active	yes	data sharing, forwarding
FIS	Active	no	no cooperation
FIS-P(NP)	Passive	yes(no)	data sharing, forwarding
FIS-NDS	Passive	no	forwarding

Table 1: Summary of the modes of operation we simulate.

For simplicity of exposition, we fix the data object. All $7DS$ nodes query for the data object till they cache it. In FIS-based schemes, at the beginning only the *information server* has this data object and always responds to queries (i.e., it has data sharing enabled). In 7DS-only schemes, at the beginning, there is only one peer with this data object in its cache. For the purpose of this study we assume that all of the traffic is generated by our application. The 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 or the *information server*. The mobile host broadcasts a query at each *on* interval till it receives the data. The broadcast is scheduled at a random time selected from the *on* interval. We assume that the *information server* has no power constraints, and therefore throughout the simulations, it does not perform any power conservation.

3 Simulation Results

Let us first introduce the user mobility pattern. We consider nodes moving in a 1000 m x 1000 m grid according to the random waypoint mobility model [10]. The random waypoint model breaks the movement of a mobile host into alternating motion and rest periods. A mobile host first stays at a location for a certain time, then it moves to a new randomly chosen destination at a speed drawn uniformly from a given interval. Each node starts from a position (x_0, y_0), and is moving towards a destination point (x_1, y_1). For each node, the x_0, y_0, x_1 and y_1 are uniformly selected from $[0,\ 1000]$. Each node is moving to its destination with a 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. We fix the pause time to be 50 s.

We define the query interval as the sum of consecutive *on* and an *off* intervals. In all the *scenarios* with power conservation enabled, the *on* and *off* intervals are equal to half the query interval. In the *scenarios* with no power conservation, the *off* interval is equal to 0 and the *on* and query interval are the same. A *scenario* is a Tcl file that “defines” the topology and movement of each host that participates in an experiment. A host is modeled as a ns-2 *mobilenode* with $7DS$ running as an *agent* of the *mobilenode* [11, 12]. We generate 300 different *scenarios* for each case of network size of 5, 10, 15 and 25 hosts. For the FIS-based schemes, one of these hosts is the fixed *information server* and the rest are $7DS$ peers. For the 7DS-only schemes, all hosts are $7DS$ peers. As mentioned before, we consider a single data item, with the $7DS$ peers trying to get a copy of it. So, the network size indicates its popularity. In each of these *scenarios*, the mobility pattern of each host is created using the mobility pattern we described. We run simulations using these *scenarios*, for the different schemes of Table 1. We use the ns-2 simulator [11] with the mobility and wireless extensions [12]. The wireless LAN is modeled as a WaveLAN network interface[2].
We consider transmission powers of 281.8 mW (high), $\frac{281.8}{2^4}$ mW (medium) and $\frac{281.8}{2^8}$ mW (low) [3].

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 at the beginning of the simulation. We run the 300 generated *scenarios* for each test and computed the average of the percentage of hosts that *became* data holders (at the end of each test). Figures 1 (a), (b), and (c) show the percentage of dataholders as a function of the network size for 7DS-only and FIS-based schemes. In this set of simulations, the query interval is 15 s. 7DS-only schemes outperform FIS-based ones. The difference between the 7DS-only schemes and the FIS is prominent, independent of the network density. For network size of 25 hosts, the $7DS$-only schemes reach their maximum effectiveness of data sharing. Also, notice that the network size and the power transmission have the most dominant effect on the data propagation. For example, for large network size and high power transmissions, P, NP and FW perform the same. In Figure 1 (a), the data sharing is effective. On the other hand, the forwarding does not result in any performance improvements, due to the low probability that another querier will be close to a host that just acquired the data. Another interesting remark is that for the query interval we consider, the power conservation has small effect on the data propagation especially when either the power transmission is high and/or the network size is large (and therefore, in those cases, it can be used to save energy). Also, the network size has a small effect on the data dissemination in FIS-based schemes. This is because the probability that a mobile host receives data from the server is higher from the probability that it receives it from another mobile host. Also, the lower the transmission power, the lower is the probability that the mobile host will be within the coverage from the *information server* making the collaboration less effective. As we expect, the performance of FIS remains constant as the network size increases, since the data sharing takes place only when a host is in proximity with the information server.

4 Related work

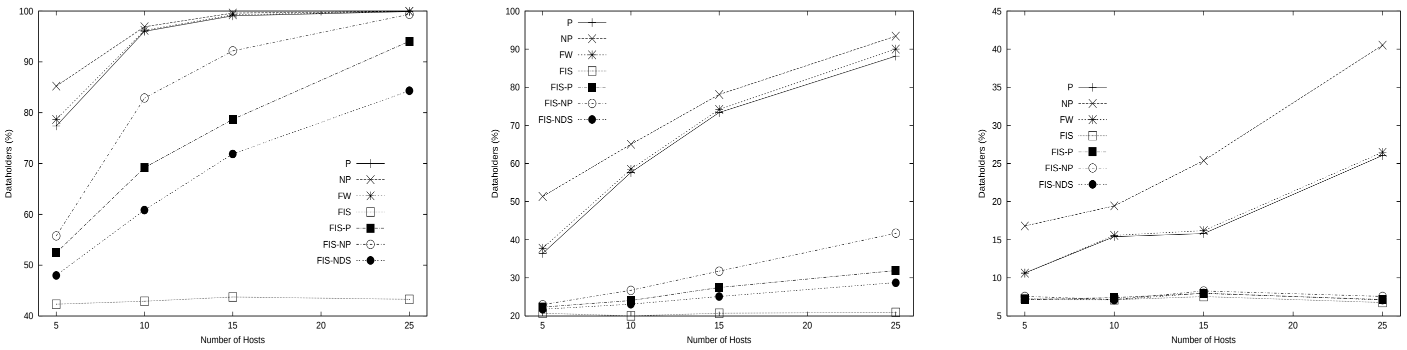
Napster [13] and Gnutella [14] are two systems that explore the cooperation among hosts and enable data sharing among users in a fixed wired network. The first is focus in sharing music files, whereas the latter for any type of files. In the case of Gnutella, the hosts need to maintain fix connections with each other. This would not be easy or feasible in the very dynamic setting we are considering. As opposed 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. As opposed to Napster that requires a centralized server for indexing the music files to be shared, $7DS$ operates in a distributed fashion without the need of any central server. Moreover, Napster requires user intervention and effort for uploading files, whereas $7DS$ does this automatically.

Info-stations have first been mentioned by Imielinski in the DataMan project [5]. Badrinath was among the first to propose infrastructure for supplying information services, such as e-mail, fax and web access by placing info-stations at traffic lights and airport entrances. Prefetching targeted for mobile users in a wide area wireless network has been used in [15] in a context more similar to ours. Tao Ye *et al* assume an info-station deployment.

---

[2] We used the WaveLAN interface simulation available with the ns-2 software.

[3] Assuming the two-ray ground reflection model these transmission powers correspond to ranges of approximately 230 m, 115 m and 57.5 m, respectively.



(a) (b) (c)

Figure 1: Figures (a), (b) and (c) correspond to a high, medium and low transmission power, respectively. The query interval is 15 s.

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. They investigate the effectiveness of info-stations 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 info-station 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 via any base-station. Second, they investigate the effectiveness of info-stations as compared to a traditional wide-area wireless network. For that, they vary the info-station density and its coverage. In our case, we consider a fixed info-station (i.e., FIS) in the region of 1 km x 1 km (that corresponds to low info-station 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 info-station. Also, we vary several parameters (like various mobility patterns, power conservation methods and querying schemes) that have not been investigated in [15]. Due to the space limitations, we will not discuss further the related work. We include a more detailed discussion in [9].

5 Conclusions and Future Work

In this paper, we presented $7DS$, a new peer-to-peer data 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. It runs as an application complementary to other data access approaches. It anticipates the information needs of users and fulfills them by searching for information among peers. To assess the efficiency of information dissemination via $7DS$, we performed an extensive simulation-based study of the two data dissemination approaches, namely 7DS-only and FIS-based by measuring both the percentage of data holders at the end of the simulation. Also, we investigated the effect of the wireless coverage range, network size, query mechanism, cooperation strategy among the mobile hosts and power conservation with a very large number user mobility scenarios. For example, we found that the network size and the power transmission are dominant parameters for the data dissemination. Also, for a large network (of 25 hosts), the pure peer-to-peer ($7DS$-only) approach reaches its maximum effectiveness. Our current research direction includes the intergration of $7DS$ with a tour guide and an academic news notification system and its deployment in the campus. Ideally, $7DS$ should be *on* when there is high probability for data availability. The challenge is to predict when the data availability increases. The data availability depends on the $7DS$ density and whether the mobile host is within the coverage of an *information server*. It is part of future work to investigate how these observations can be used to improve the power utilization as well as some theoretical issues in respect to the data propagation in this environment.

References

[1] ALERT. http://www.alertsystems.org.

[2] Airflash. http://www.airflash.com.

[3] Cell-Loc. http://www.cell-loc.com.

[4] "NTT DoCoMo." http://www.nttdocomo-inpaku.com.

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

[6] R. H. Frenkiel, B. R. Badrinath, J. Borrs, and R. D. Yates, "The infostations challenge: balancing cost and ubiquity in delivering wireless data," *IEEE Personal Communications Magazine*, vol. 7, pp. 66–71, Apr. 2000.

[7] N. Davies, K. Mitchell, K. Cheverst, and G. Blair, "Developing a context sensitive tourist guide," 1998.

[8] Metricom, "The Ricochet wireless network overview," 1999.

[9] M. Papadopouli and H. Schulzrinne, "$7ds$: A peer-to-peer data dissemination and prefetching tool for mobile users," in *submission for publication*.

[10] 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 *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom'98)*, (Dallas, Texas), Oct. 1998.

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

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

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

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

[15] T. Ye, H.-A. Jacobsen, and R. Katz, "Mobile awareness in a wide area wireless network of info-stations," in *Proc. of Mobicom*, (Dallas, Texas), Oct. 1998.