Mobile Networks and Applications 4 (1999) 117–129

A study on channel allocation for data dissemination in mobile computing environments

Wang-Chien Lee [a], Qinglong Hu [b] and Dik Lun Lee [b]

[a] *GTE Laboratories Incorporated, 40 Sylvan Road, Waltham, MA 02451, USA*
[b] *Department of Computer Science, University of Science and Technology, Clear Water Bay, Hong Kong*

This paper studies channel allocation methods for data dissemination through broadcast and on-demand channels. Analytical models and cost formulae for exclusive broadcast channels and exclusive on-demand channels are provided. Based on the models, we further derive cost models for dynamic channel allocation methods and propose a channel adaptation algorithm for optimizing system performance. The channel adaptation algorithm can be executed in $O(n)$ time, where n is the number of data items in the database. Performance evaluation shows that the channel allocation algorithm produces optimal channel allocation which significantly improves the system performance under various parameter settings.

1. Introduction

Owing to rapid advances of personal computing and wireless communication technologies, research in mobile computing has gained a great momentum in the past few years. Industry, government, and the academia are all anticipating the forthcoming mobile communication and computing era [6,8]. It is envisaged that mobile users with palmtop and laptop computers will be able to easily access Intranet data and various Internet services in the near future. To realize this vision, it is crucial to develop data access methods which efficiently utilize the wireless resources. This paper presents wireless channel allocation methods for data dissemination in mobile computing systems.

Wireless broadcast allows an arbitrary number of users to simultaneously receive information broadcast on the air without any performance downgrade. Thus, research on utilizing and facilitating wireless information broadcast techniques in mobile computing environments has been very active in recent years [1,11,19].

Acharya et al. proposed a broadcast scheduling facility, namely *broadcast disk*, for non-uniform access of data. They load data onto broadcast disks of various speeds based on data access frequencies. Data on different broadcast disks are interleaved into a stream for dissemination on broadcast channels. As a result, data items from high speed broadcast disks are broadcast more frequently than those from lower speed broadcast disks. In [1], the performance of the broadcast disks is investigated by simulative experiments. By taking issues such as prefetching and updating into consideration, the broadcast disk project further extends the study on broadcast scheduling [2,3]. With a separate effort, [9] developed a broadcast scheduling algorithm, also based on data access frequency, for both single-channel and multiple-channel broadcast services.

Broadcast channels are also used to address power conservation issues, because mobile computers consume less power by passively receiving data disseminated on broadcast channels than actively interacting with the base stations. To reduce power consumption of the mobile computers, [12,13,17] studied indexing techniques for data accessing on broadcast channels. By using index delivered along with data, the mobile computers can turn to doze mode and wake up only when the data of interest arrives. Distributed index tree and hashing techniques are used in [13] and [12], respectively, while signature techniques are used in [17]. The above research study on indexing techniques did not consider the factors of data access patterns and the impacts of the on-demand services.

In a previous paper [16], we showed that broadcast technique can effectively complement traditional *on-demand channels* for data dissemination in a mobile computing environment and proposed a hybrid channel allocation method which mixes the broadcast and on-demand channels to utilize system bandwidth more efficiently. In this paper, we extend the idea to propose a *dynamic channel allocation method*, which dynamically assigns channels for broadcast or on-demand services based on system workload. We also provide a channel allocation algorithm which efficiently obtains the optimal channel allocation by approximation calculations. Comparisons of data access efficiency between exclusive on-demand and dynamic channel allocation methods are also conducted.

Research in [4,11,19] is closely relevant to our work presented in this paper. [11] provided an architecture for wireless information dissemination and discussed the broadcast and the on-demand services. However, their model and analysis are different from ours, i.e., they considered only one channel in a cell and the channel is divided into two sub-channels: on-demand and broadcast based on system workload. The on-demand channel is modeled as a M/M/1 queue, while we use a M/M/m/B queuing model. There were no simulation results and comparisons in that paper. [4,19] studied the schemes for allocating data on broadcast and on-demand channels. However, their studies assume a fixed ratio for on-demand to broadcast bandwidth.

The rest of the paper is organized as follows. In section 2, we briefly review the issues related to data dissemination in mobile computing systems. In section 3, we provide analytical models and cost formulae to evaluate the data access efficiency for the on-demand and broadcast channels. In section 4, we further develop the cost formulae for the dynamic channel allocation method and propose an algorithm to approximate the optimal channel allocation based on the formulae. Section 5 gives the performance evaluation and comparison of the exclusive on-demand and dynamic channel allocation methods. Section 6 concludes the paper.

2. Wireless data dissemination

In this section, we briefly introduce the mobile computing model used in this study and the issues related to data dissemination in mobile computing systems, such as channel allocation methods, data access methods and efficiency, and broadcast scheduling strategies.

2.1. Mobile computing model

A mobile computing system consists of a set of mobile computers and stationary hosts connected by a fixed network. The mobile computers communicate with other computers through wireless communication, while the stationary hosts communicate with each other through the wired network. In order to support wireless communication for mobile computers, the geographical area covered by the system is divided into *cells*. The wireless communication in a cell is supported by a *mobile support station* (MSS), which is one of the stationary computers on the fixed network. The MSS controls the wireless communication channels in a cell. In this paper, we assume that the *basic fixed channel assignment strategy* [20] is applied to the system, i.e., there is a fixed set of channels permanently assigned to each cell and a service request at a cell site can only be served by the unoccupied channels at that cell site; otherwise, the request is blocked.

In addition to supporting communication, the MSS may provide various services to the mobile users. As a result, the MSS may be considered as a database server which is responsible for providing and disseminating data to its mobile clients in the cell. Figure 1 shows the architecture for supporting mobile computing and illustrates the terminology such as mobile computers, MSSs, and cells.

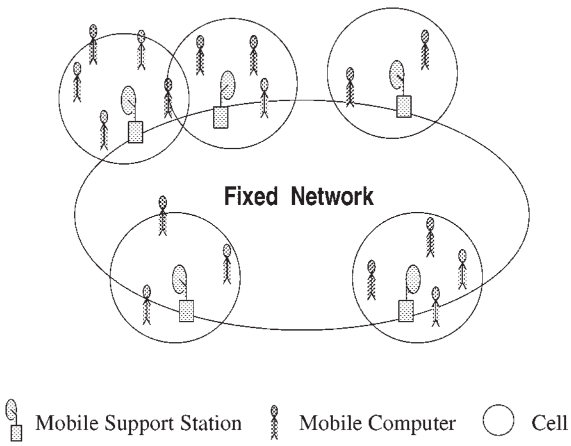

Figure 1. Mobile computing system.

2.2. Channel allocation for data dissemination

In a mobile computing system, the communication load between mobile computers and the MSS varies over time and space because of the mobility of mobile computers. With the limited and fixed bandwidth available in a cell, it is necessary to adjust the allocation of channels for broadcast and on-demand purposes in order to improve the overall communication performance in a cell. In the following, we define four classes of channel allocation methods:

- **Exclusive On-demand**: All of the channels are in on-demand mode. That is, data requests and results are delivered through point-to-point connections. This method is desirable when the number of queries is small compared to the number of channels available and when energy efficiency is not an issue for the mobile computer to transmit uplink requests.
- **Exclusive Broadcast**[1]: Contrary to the exclusive on-demand method, all of the channels are in broadcast mode. Data items are broadcast periodically in broadcast channels. This method is useful when a small number of data items are of interest to a large group of users.
- **Hybrid Allocation** [16]: This method mixes the broadcast and on-demand channels in order to utilize system bandwidth more efficiently. The main idea is to make broadcast and on-demand channels complement each other.
- **Dynamic Allocation** [10]: This method dynamically allocates broadcast and on-demand channels to achieve optimal data access performance. In contrast to the hybrid method, the dynamic method allocates channels based on different workloads at the MSS. When the load is heavy, the broadcast channels may significantly relieve the load on on-demand channels by taking care of frequent accesses to hot data items. When the load is light, on-demand channels can take over to provide instantaneous access to data.

2.3. Data access methods

With no prior knowledge on how the requested data item is provided by the MSS, a mobile computer has to send a

[1] This method may not be feasible for applications with diversified user access pattern to a large database. It is included here to facilitate our development of cost formulae for the dynamic channel allocation method.

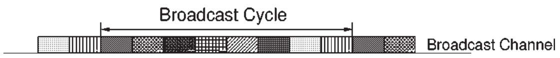

Figure 2. Broadcast channel.

data access request to the MSS. In response, the MSS will either directly deliver the data item, if it is only available through on-demand service, or reply with the *broadcast channel access information*, such as the channel frequencies, data identifier, the decryption key, expiration time, and estimated access time. When the mobile computer receives the broadcast channel access information, it may decide if it wants to get the data through the broadcast channel or right away through on-demand channels. In the first case, the mobile computer will terminate the connection and monitor the broadcast channels, whereas in the second case, it returns a confirmation to the MSS, which will then deliver the data and terminate the connection.

If the mobile computer already has the broadcast channel access information regarding the data item of interest, it does not need to query the MSS unless the information is expired. The mobile computer may use the information to monitor broadcast channels to receive the data item. Since broadcast channels allow simultaneous access by an arbitrary number of users, the charge for accessing data items from the broadcast channels may be based on a different pricing scheme from that of the on-demand channels. For example, the price may be based on the valid period of the broadcast channel access information.

2.4. Broadcast scheduling

A complete broadcast of all of the data items in a channel is called a *broadcast cycle*. Figure 2 illustrates a broadcast channel and a broadcast cycle. In a multiple broadcast channel system, the data items may be distributed to different channels in order to reduce the length of broadcast cycle. Logically, increasing the number of broadcast channels is the same as increasing the bandwidth of a broadcast channel.

The MSS has to schedule the data for broadcast in the allocated broadcast channels. In this paper, we choose to broadcast data items in sequence. An important characteristic of this strategy is that the access time to any broadcast data item is always bounded. The issue of broadcast scheduling is out of the scope of this paper. Readers interested in this topic may refer to [1,5,15].

2.5. Data access efficiency

The goal of our study is to pursue data access efficiency.[2] The criteria used in this paper to evaluate data access efficiency of wireless channels is *access time*,[3] which is the period from the time a mobile computer requests for a data item until the data is received. The broadcast channels and on-demand channels introduce different access overhead for the mobile users. For the broadcast channels, the main overhead is due to *probe time*, defined as the duration between the time the user starts to monitor the channels and the time when the requested data item is located. On the other hand, the overhead for accessing data through the on-demand channels is the waiting time for connecting to the MSS, which takes longer when the service load of the system is high.

3. The analytical models

To facilitate our study on channel allocation methods, we model the on-demand and broadcast channels and derive formulae for their access time performance, respectively. We consider a cell in the mobile computing systems and made the following assumptions:

- The cell consists of a MSS and m mobile computers;
- The MSS maintains n data items, $D_1, D_2, \dots, D_n$, with average size s;
- The mobile computers may issue read requests,[4] which have size r, to the MSS;
- A mobile computer will not issue a new request before the previous request is completed;
- The number of read requests generated from the m mobile computers per unit time to $D_1, D_2, \dots, D_n$ is $\lambda_1, \lambda_2, \dots, \lambda_n$, respectively;
- The MSS supports c wireless channels and each channel has bandwidth b;
- The number of potential mobile computers is greater than the number of channels available in a cell, i.e., $m \gg c$.

3.1. On-demand channels

We first consider the access time for a mobile computer to retrieve a data item from the MSS through the on-demand channels. The mobile computer has to establish a connection to the MSS before submitting a request. Upon receiving the request, the MSS returns the requested data items to the user. Since there are c channels in the cell, there are at most c simultaneous communication sessions. We assume that a mobile computer will not issue a new request before the previous request is completed. Therefore, there are at most m requests to be serviced in the system. To simplify our analysis, we do not consider the time for hand-shaking in order to establish the connections and the process time

[2] Other goals pursuing revenue increase, fair pricing policies and quality of services are out of the scope of this paper.

[3] A criteria frequently used in evaluating wireless data access methods is *tune-in time*, which represents the period of time a mobile computer spent on monitoring the wireless channels. The tune-in time, which may be transformed into CPU time or battery power usage on accessing the data, is mainly used for evaluating power consumption efficiency of data access methods.

[4] The computers may also issue write requests. However, we only consider read requests in this paper.

for MSS to retrieve the requested data items from the disk. Thus, the expected channel use time for each communication session is $1/\mu = (r+s)/b$. As a result, the service rate for an on-demand channel is

$$\mu = \frac{b}{r+s}.$$

The aggregate arrival rate of requests from the mobile computers is

$$\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

We use a M/M/c queuing model (with finite buffers or infinite buffers) to analyze the access time performance of the on-demand channels. Due to the assumptions on input to queuing models, our analysis implicitely assumes that the request arrival rates are stable over time. The traffic intensity, ρ, which indicates the system workload for the quening model, is [14]

$$\rho = \frac{\lambda}{c\mu}. \tag{1}$$

Based on this model, we derive formulae for access time performance of the on-demand channels. However, the formulae are not simple enough for efficient calculation of the access time. Without losing generality, we adopted approximation methods to estimate the on-demand access time performance. We observed from experiments in our previous study [16] that the system performance differs greatly due to changes in several different system parameters, such as the number of channels available in a cell, the aggregate arrival rate of requests and the number of clients served in a cell, which directly affect the system workload. Thus, we develop the formulae to estimate the access time performance of on-demand channels based on different system workloads. We use the traffic intensity of the system, ρ, as the dividing factor of the system workload, i.e., when $\rho > 1$, the cell is under heavy workload and when $\rho < 1$, the cell is under light workload.

3.1.1. On-demand channels under heavy system workload

When $\rho > 1$, the state transitions due to the arrival of data requests and completion of data requests can be modeled as a birth-death system shown in figure 3.

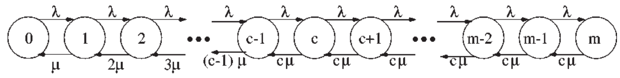

Figure 3. State transition diagram for on-demand channels under heavy system workload.

Each state of the diagram denotes the number of mobile computers requesting a data item during that state. Since there are at most m requests buffered for services, the system can be modeled as a $M/M/c/m$ queue.

The probability of zero data request in the system is [14]

$$p_0 = \left[1 + \frac{(c\rho)^c(1-\rho^{m-c+1})}{c!(1-\rho)} + \sum_{i=1}^{c-1}\frac{(c\rho)^i}{i!}\right]^{-1}.$$

The probability of having i requests in the system is

$$p_i = \begin{cases} p_0\dfrac{(c\rho)^i}{i!}, & 0 \leqslant i < c, \\ p_0\dfrac{c^c\rho^i}{c!}, & c \leqslant i \leqslant m. \end{cases} \tag{2}$$

Since a mobile computer will be blocked until the previous request is completed, the effective arrival rate of requests from the mobile computers is

$$\hat{\lambda} = \sum_{i=0}^{m-1}\lambda p_i. \tag{3}$$

By applying equations (2) and (3), we can derive the expected access time for a mobile computer to retrieve a data item from the on-demand channels as follows:

$$\begin{aligned} E[o] &= \frac{\sum_{i=1}^{m} ip_i}{\hat{\lambda}} \\ &= \frac{c\rho\sum_{i=0}^{c-2}\frac{(c\rho)^i}{i!} + \frac{c^c}{c!}\sum_{i=c}^{m} i\rho^i}{\lambda\sum_{i=0}^{c-1}\frac{(c\rho)^i}{i!} + \lambda\frac{c^c}{c!}\sum_{i=c}^{m-1}\rho^i}. \end{aligned} \tag{4}$$

If we divide both the numerator and the denominator of the above equation by $e^c\rho^m$, we have

$$E[o] = \frac{c\rho\sum_{i=0}^{c-2}\frac{(c\rho)^i}{i!}\frac{1}{e^c\rho^m} + \frac{c^c}{c!}\sum_{i=c}^{m} i\rho^i\frac{1}{e^c\rho^m}}{\lambda\sum_{i=0}^{c-1}\frac{(c\rho)^i}{i!}\frac{1}{e^c\rho^m} + \lambda\frac{c^c}{c!}\sum_{i=c}^{m-1}\rho^i\frac{1}{e^c\rho^m}}. \tag{5}$$

Considering the facts that $\rho > 1$ and $m \gg c$, we get the following approximations:

$$\begin{aligned} 0 &< c\rho\sum_{i=0}^{c-2}\frac{(c\rho)^i}{i!}\frac{1}{e^c\rho^m} < c\rho^{(c-1)}\sum_{i=0}^{c-1}\frac{(c)^i}{i!}\frac{1}{e^c\rho^m} \\ &< c\rho^{(c-1)}e^c\frac{1}{e^c\rho^m} = \frac{c}{\rho^{(m+1-c)}} \approx 0 \end{aligned}$$

and

$$\begin{aligned} 0 &< \lambda\sum_{i=0}^{c-1}\frac{(c\rho)^i}{i!}\frac{1}{e^c\rho^m} < \lambda\rho^{(c-1)}e^c\frac{1}{e^c\rho^m} \\ &= \frac{\lambda}{\rho^{(m+1-c)}} \approx 0. \end{aligned}$$

Based on Stirling's formulae [7] and the facts that $\rho > 1$ and $m \gg c$, we can approximate the above formulae and obtain the expected access time for a mobile computer to retrieve a data item from the on-demand channels as follows:

$$E[o] \approx \frac{m}{c\mu}. \tag{6}$$

Since $\partial^2E(o)/\partial c^2 = 2m/c^3\mu > 0$, we can conclude that $E[o]$ behaves as a convex function.

3.1.2. On-demand channels under light system workload

When the traffic intensity $\rho < 1$, the state transitions due to the arrival of data requests and completion of data requests can be modeled as a birth-death system shown in figure 4.

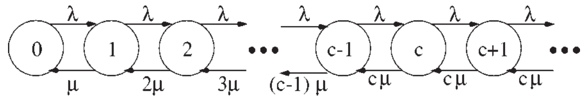

Figure 4. State transition diagram for on-demand channels under light system workload.

Each state of the diagram denotes the number of mobile computers requesting a data item during that state. Since $\rho < 1$ and $m \gg c$, all query requests can be served within bounded time. Therefore, the system can be modeled as a $M/M/c$ queue.

The probability of zero data request in the system is

$$p_0 = \left[1 + \frac{(c\rho)^c}{c!(1-\rho)} + \sum_{i=1}^{c-1} \frac{(c\rho)^i}{i!}\right]^{-1}.$$

The probability of having i requests in the system is

$$p_i = \begin{cases} p_0 \dfrac{(c\rho)^i}{i!}, & 0 \leqslant i < c, \\ p_0 \dfrac{c^c \rho^i}{c!}, & c \leqslant i \leqslant m. \end{cases}$$

Let P_c be the probability that the mobile computers request c or more data items. We can obtain P_c as follows [14]:

$$P_c = \frac{p_0(c\rho)^c}{(1-\rho)c!}. \tag{7}$$

The expected access time for a mobile computer to retrieve a data item is [14]

$$E[o] = \frac{1}{\mu} + \frac{P_c}{c\mu(1-\rho)}. \tag{8}$$

To show the accuracy of our approximation, we conducted a comparison, based on parameter settings used in our evaluation, to compare equation (6) (for $\rho > 1$) and equation (8) (for $\rho < 1$) to equation (4) of the $M/M/c/m$ model. Interested readers are requested to refer to the appendix for the comparison.

3.2. Broadcast channels

Next we develop the cost model for the access time of the broadcast channels. To simplify our analysis, we assume that the mobile computer has the valid broadcast channel access information for the data items of interest. Since there are c channels used for broadcast, the aggregate bandwidth for broadcast channels is bc. Assume that there are k data items broadcast periodically in the c broadcast channels. The size of a broadcast cycle is ks/bc. The average probe time is $ks/2bc$. Thus, the access time for retrieving data through monitoring broadcast channels is

$$E[b] = \frac{ks}{2bc} + \frac{s}{b}. \tag{9}$$

4. Dynamic channel allocation

As we discussed earlier, some of the data items have high request rates due to common interest of a large group of users. Thus, it is logical to disseminate these data items on broadcast channels in order to relieve the system workload at the MSS. When more and more hot data items are moved to the broadcast channels, some of the on-demand channels have to be reallocated to serve as broadcast channels in order to maintain a reasonable access time on the broadcast channels.

There are several factors determining whether on-demand channels should be reallocated as broadcast channels, e.g., the potential number of data requests off-loaded from the on-demand channels to broadcast channels, the number of data items on the broadcast channels, and the number of channels reallocated from on-demand to broadcast. In order to reach optimal data access efficiency in the cell, the system should dynamically reassign channels between the on-demand and broadcast services. In this section, we provide methods to decide optimal allocation of the on-demand and broadcast channels. We first derive the actual service rate of an on-demand channel in dynamic channel allocation environment, which may be used for calculation of the average data access time in the dynamic channel allocation environment. Then, we derive formulae for the average data access time and propose an algorithm to decide the best allocation of the on-demand and broadcast channels.

In the following discussion, we assume that the n data items, $D_1, D_2, \ldots, D_n$, are sorted by their access rates such that $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. Let k ($1 \leqslant k \leqslant n$) be the number of the most frequently accessed data items made available on the broadcast channels and ε denote the probability that a data item is accessed by mobile computers through the on-demand channels, although it is available on the broadcast channels.

4.1. Actual service rate of on-demand channels

In order to calculate the average data access time of the dynamic channel allocation method, we have to derive the actual service rate of the on-demand channels in a dynamic channel allocation environment.

Assume that the size of the message for broadcast channel access information is q. Therefore, the average service rate for each on-demand channel is

$$\mu' = \frac{b}{\omega_o(r+s) + \omega_b(r+q) + \omega_{b|o}(2r+q+s)}, \tag{10}$$

where ω_o, ω_b and $\omega_{b|o}$ are the percentages of communication sessions established between the mobile computers and the MSS for requests of data which are only available through the on-demand channels, requests of broadcast channel access information, and immediate delivery of data available on the broadcast channels, respectively.

4.2. Development of a channel allocation algorithm

Let us assume the total number of the communication channels in a cell is c and each channel has the same bandwidth b. These channels consist of two pools: the broadcast pool and the on-demand pool. Let c_b and c_o denote the number of channels in the broadcast pool and the number of channels in the on-demand pool, respectively. Thus,

$$c = c_b + c_o. \tag{11}$$

To make the discussion simple, we define the arrival rate of the requests for hot data items available on the broadcast channels as $\lambda_{hot} = \sum_{i=1}^{k} \lambda_i$ and the arrival rate of the requests for cold data items that can only be accessed via on-demand channels as $\lambda_{cold} = \sum_{i=k+1}^{n} \lambda_i$.

Since ε percent of the hot data items are retrieved from on-demand channels, the actual arrival rate of the requests for the data items answered via on-demand channels, λ', is as follows:

$$\lambda' = \varepsilon\lambda_{hot} + \lambda_{cold}.$$

The overall expected access time for any data item is

$$E[d] = \frac{1}{\lambda}\big(\lambda_{hot}E[b](1-\varepsilon) + \lambda' E[o]\big), \tag{12}$$

where $E[o]$ denotes the expected access time for a mobile computer to retrieve a data item through an on-demand channel and $E[b]$ denotes the expected access time for retrieving data through monitoring broadcast channels.

4.2.1. Heavy system workload

In the following, we adopt approximation methods for a faster estimation of the optimal access time and channel allocation. First, we explore the situation when the system is under a heavy system workload ($\rho > 1$). Based on equations (6), (9) and (11), we substitute $E[o]$, $E[b]$ and c_b in equation (12) to obtain

$$E[d] = \frac{1}{\lambda}\left(\lambda_{hot}(1-\varepsilon)\left(\frac{ks}{2b(c-c_o)} + \frac{s}{b}\right) + \lambda' \frac{m}{c_o\mu'}\right), \tag{13}$$

where μ' is the actual service rate of on-demand channels in the dynamic channel allocation environment (see equation (10)).

To simply further development of the formula, we define the following constants:

$$A = \frac{1}{\lambda}\lambda_{hot}(1-\varepsilon)\frac{ks}{2b},$$
$$B = \frac{1}{\lambda}(\lambda_{hot}\varepsilon + \lambda_{cold}),$$
$$C = \frac{1}{\lambda}\lambda_{hot}(1-\varepsilon)\frac{s}{b}.$$

Thus, equation (13) can be rewritten as

$$E[d] = \frac{A}{c-c_o} + \frac{Bm}{c_o\mu'} + C. \tag{14}$$

We differentiate the function $E[d]$ on c_o and make it equal to zero (i.e., $\partial E[d]/\partial c_o = 0$) in order to evaluate the optimal point for c_o. As a result, the optimal allocation of on-demand channels, $\hat{c}_o$, can be determined as follows:

$$\hat{c}_o = \frac{c\sqrt{mB}}{\sqrt{mB} + \sqrt{\mu' A}}.$$

It is important to note that the condition for the above derivation to be valid is that the system must be under heavy workload. Unfortunately, the actual traffic intensity, ρ', cannot be decided until the optimal allocation for the on-demand channels is obtained. Thus, there is a chance that $\hat{c}_o$ obtained above is not valid because it causes the resulting system to operate under light load (i.e., $\rho' = \lambda'/(\hat{c}_o\mu') < 1$). In other words, (14) cannot be used under such a circumstance. To confirm the validity or otherwise obtain the real optimal value, we can make use of the property that (14) is a concave function and that if $\hat{c}_o$ is not the real optimum the real one must be smaller than $\hat{c}_o$:

$$\bar{c}_o = \max\big\{c_o | \lambda'/(c_o\mu') > 1 \text{ and } 0 \leqslant c_o \leqslant \hat{c}_o\big\}. \tag{15}$$

In the next section, we derive formulae for light-loaded systems. Thus, the real optimal point can be decided by comparing the performance obtained under both circumstances.

4.2.2. Light system workload

Next, we look at the situation when the system is under light workload ($\rho < 1$) . Based on equations (8), (9) and (11), we substitute $E[o]$, $E[b]$ and c_b in equation (12) to obtain

$$E[d] = \frac{1}{\lambda}\left(\lambda_{hot}(1-\varepsilon)\left(\frac{ks}{2b(c-c_o)} + \frac{s}{b}\right) + \lambda'\left(\frac{1}{\mu'} + \frac{P_{c_o}}{c_o\mu'(1-\rho')}\right)\right), \tag{16}$$

where P_{c_o} is the probability that the mobile computers are requesting c_o or more data items from the on-demand channels and ρ' is the actual traffic intensity (i.e., $\rho' = \lambda'/(c_o\mu')$).

We continue to use the constants A, B and C defined above and further define a constant $D = \frac{\lambda'}{\lambda\mu'}$. In order to simplify the complexity of differential derivation in the next steps, we set P_{c_o} to the upper bound (i.e., 1) to simplify (16) into

$$E[d] \leqslant \frac{A}{c-c_o} + \frac{B}{c_o\mu'(1-\rho')} + C + D.$$

By differentiating the above upper boundary to evaluate the optimal point for c_o, we find the minimum point for the light-loaded system:

$$\hat{c}_o = \frac{\rho''\sqrt{\mu' A} + \sqrt{B}}{\sqrt{\mu' A} + \sqrt{B}}c. \tag{17}$$

If the actual traffic intensity is greater than 1, with the same argument as in the previous section, the minimum point where the on-demand channels work in the light-loaded mode can be recalculated as follows:

$$\bar{c}_o = \min\left\{c_o | \lambda'/(\mu' c_o) < 1 \text{ and } \hat{c}_o \leqslant c_o \leqslant c\right\}. \quad (18)$$

This gives us an approximate method for allocating broadcast channels and on-demand channels in a way that the access time is no more than $A/(c - \bar{c}_o) + B/(\bar{c}_o\mu' - \lambda') + C + D$ for any data item.

4.2.3. Channel allocation algorithm

Based on the formulae derived above, we present the following algorithm to approximate the optimal channel allocation for optimizing the system performance.

Algorithm Approx-Optimal Channel Allocation.
Given n data items, $D_1, D_2, \ldots, D_n$, sorted in descending order by data access frequency, and c wireless channels in a cell.
Decide the (approximate) optimal allocation of channels for broadcast and on-demand services and the set of data items to be disseminated on the broadcast channels.

begin

1. Initialize the allocation to be exclusive on-demand (i.e., $C_O = c$, $C_B = 0$, and the set of broadcast data is empty, i.e., $i = 0$).
2. If $\rho > 1$, initialize the approximate lowest access time T using equation (6)
   else initialize the approximate lowest access time T using equation (8).
3. For $i = 1$ to n do
   begin
   - Identify $D_1, \ldots, D_i$ as the data items to be broadcast.
   - Compute c_o^h using equation (15) and c_o^l using equation (18).
   - Compute the approximate lowest access time t^h using equation (13) and t^l using equation (16) based on c_o^h and c_o^l, respectively.
   - Let t be min(t^h, t^l) and c_o be the corresponding assignment of on-demand channels.
   - Compute c_b using equation (11).
   - Keep track of the optimal access time $T = \min(T, t)$, the corresponding allocation, C_O and C_B, and the corresponding number of broadcast data items, I.

   end;
4. Return the corresponding I, C_O, and C_B values with the optimal access time T.

end

It is obvious that the average time complexity of the algorithm is $O(n)$.

5. Performance evaluation

Our model consists of a single server, a set of clients and a fixed number of channels. The server can either continuously broadcast top hot data items in broadcast channels or reply requests from the clients for other data items in on-demand channels.

We use a M/M/m/B queuing system to model the activity on the on-demand channels. The simulation is written in C and CSIM, which is a process-oriented simulation toolkit [18]. Each run of the simulations lasts 4,000 seconds. The performance of the broadcast channels is obtained by using the formula we developed in section 3.2, because the formulae are simple to evaluate.

We make the following assumptions in our experiments. In a cell, 50 channels have been allocated for 500 mobile users. The MSS has maintained a database of 2000 data items, each of which is 1000 bytes in size. The size of a read request and the size of a returned message for broadcast channel access information are both 10 bytes. The average bandwidth of the channels is 1000 bytes/second. Finally, we assume that every user makes one data request per second. Thus, the average arrival rate of the requests is 500/second.

To facilitate our study, we also assume the following access pattern from mobile users: the number of accesses to the data items is exponentially distributed in the simulation. About 25% of the data access requests is on the top 50 most frequently accessed data items and 50% of the requests is on the top 174 most frequently accessed data items.

Unless otherwise specified, the parameter values used in the simulation and the analytic models are provided in table 1. The primary performance criterion employed in this study is the client access time.

5.1. The approx-optimal channel allocation algorithm

We use both simulation and the analytical formulae specified in the approx-optimal channel allocation algorithm to find the best channel assignment for a given number of data items broadcast on the broadcast channels. The simulation sequentially moves the top n hottest data items to the broadcast channels and for each number the simulation is run to find the best channel assignment. The case when the number of hot data items to be broadcast is zero corresponds to the exclusive on-demand case.

Table 1
System parameter settings.

Parameters	Values
Number of clients	500
Data item size	1000 bytes
Number of data items	2000
Access information size	10 bytes
Number of channels	50
Channel bandwidth	1000 bps
Request arrival rate	500/s

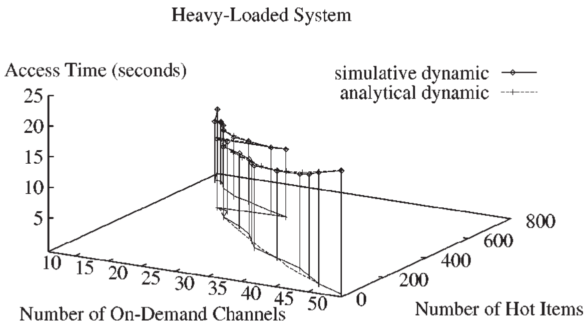

Figure 5. Optimal access time for a heavy-loaded system.

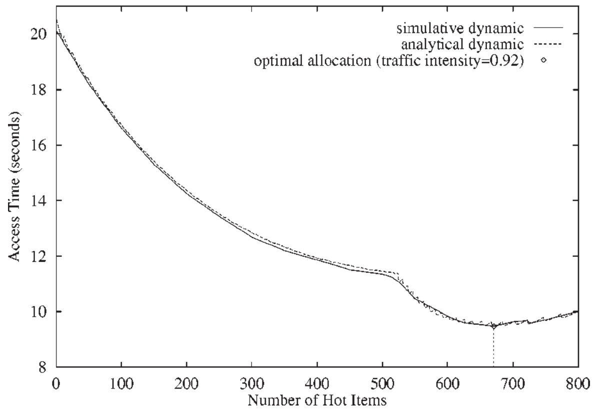

Figure 6. Optimal access time vs. number of hot items for a heavy-loaded system.

We control the request arrival rate and number of mobile computers to set up *heavy-loaded* and *light-loaded* environments, such that under heavy-loaded, the number of clients is 1000 and the request arrival rate is 500/s, while under light-loaded, the number of clients is 100 and the request arrival rate is 45 per second.

Figures 5 and 8 illustrate the best access time corresponding to the number of channels assigned to on-demand service[5] and the number of hot data items on broadcast channels under heavy-loaded and light-loaded environments, respectively. Figures 6 and 9 depict in two dimensions the optimal access time corresponding to the number of data items on broadcast channels for the heavy-loaded and light-loaded environments, respectively. For the heavy-loaded environment, we also include figure 7 to illustrate in two dimensions the optimal number of on-demand channels allocated corresponding to the number of data items on broadcast channels. In these figures, ‘analytical dynamic’ and ‘simulative dynamic’ denote the data access time of the dynamic channel allocation method obtained by the analytical formulae and by simulation, respectively. The optimal channel allocation and the number of data items which should be disseminated through broadcast channels correspond to the lowest access time in the figures. Also, from the figures, we can easily observe that the access time obtained from analytical formulae are consistently close to the simulative results.

[5] The rest of the channels are assigned to broadcast service.

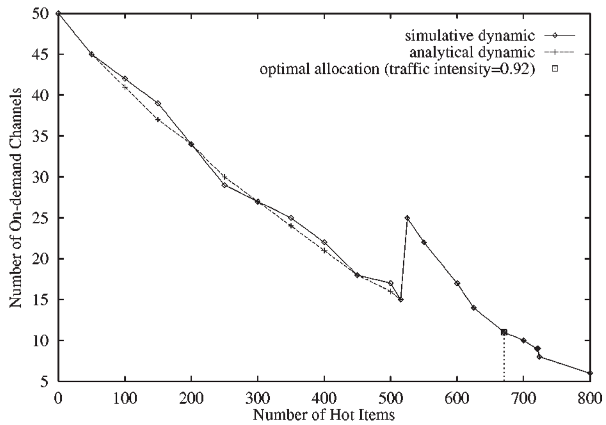

Figure 7. Number of hot items vs. number of channels for a heavy-loaded system.

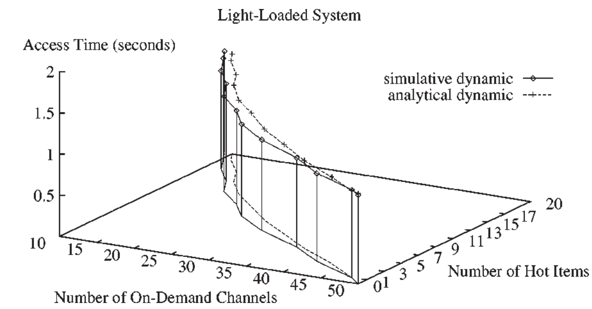

Figure 8. Optimal access time for a light-loaded system.

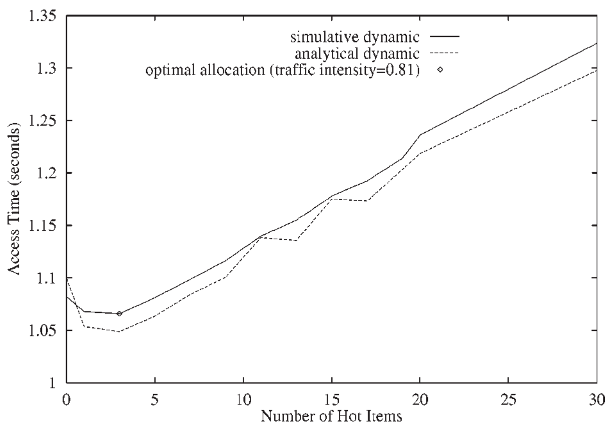

Figure 9. Optimal access time vs. number of hot items for a light-loaded system.

As shown in figures 5 and 6, the system access time of a heavy-loaded system may be significantly improved by moving some of the hot data items to the broadcast channels. The difference between the access time for the dynamic allocation method and exclusive on-demand method (i.e., points where the number of hot data items is equal to 0 in figure 6) indicates the improvement we can get when certain number of the hottest items are broadcast. In this experiment, the best data access time, 9.47 seconds, is achieved when 671 data items are broadcast on 39 broadcast channels with the remaining 11 channels providing on-demand services to the rest of data items. This is a 50.34% im-

Table 2
Optimal channels and data item allocation.

	Analytical optimal		Simulated optimal	
	Heavy load	Light load	Heavy load	Light load
Optimal allocation (on-demand/broadcast)	11/39	43/7	11/39	41/9
Number of hot data items	671	3	671	3
Optimal access time	9.489	1.049	9.470	1.066
Access time for exclusive on-demand	20.20	1.100	20.14	1.082

provement, comparing to the access time of 20.14 seconds for the exclusive on-demand services, which corresponds to the data point when the number of on-demand channels is 50 on-demand channels and the number of hot data items is zero.

Also observed in figures 6 and 7, the access time for the dynamic channel allocation method decreases as the number of hot data items moving to the broadcast channels increases (figure 6) and the corresponding number of channels assigned to on-demand channels decreases (figure 7). However, this trend changes when the number of hot items reaches 525 (figure 7), when the number of channels assigned to on-demand service jumps, while the access time continues to drop. From then on, the number of on-demand channels decreases again to let more broadcast channels handle the increased broadcast data items. This turning behavior can be explained as follows. As the system moves more hot data items to broadcast channels, the workload for on-demand channels will gradually become light-loaded. In a light-loaded system, the waiting time for on-demand channels drops greatly, the broadcast channels do not have as significant an effect on alleviating on-demand workload as they do in a heavy-loaded system. Thus, the dynamic channel allocation method reassigns some channels back to the on-demand service in order to improve the overall system performance. In our experiments, when about 525 data items are disseminated via broadcast channels, the workload of on-demand services becomes light-loaded, i.e., $\rho = 0.92$.

For a light-loaded system, we assume that the cell has only 100 mobile users and the request arrival rate is only 45 per second. Figures 8 and 9 show that the dynamic channel allocation method reaches the optimal performance when it assigns 9 channels to broadcast 3 data items[6] and 41 on-demand channels to serve the rest of the data items. With this optimal channel assignment, $\rho = 0.81$. The optimal data access time of the dynamic allocation method obtained by simulation is 1.066 seconds, which represents 0.016 second of improvement over the access time of exclusive on-demand channels, 1.082 seconds. Figure 9 also illustrates that the exclusive on-demand channel is better than most combinations in which broadcast channels are used. Even the optimal dynamic allocation has only 1.51% of improvement over the exclusive on-demand method. We can conclude that exclusive on-demand channels is a very good choice for a light-loaded system. Generally speaking, the dynamic channel allocation method does not improve much the data access time for light-loaded systems and the improvement only happens when the probe time of the broadcast channels is shorter than the transmission time of the data request message for on-demand services. In our analysis and experiments, we have neglected the setup time for on-demand services. In a realistic environment, the dynamic channel allocation method may still significantly improve the performance of light-loaded systems.

[6] We assume that channel bandwidth can be aggregated to deliver the data items.

Table 2 summarizes the access time and optimal channel allocations obtained in our experiments. From the data collected from analysis and simulation, we observed that analytical and simulative results are very close, which demonstrates the feasibility of utilizing the approx-optimal channel allocation algorithm to decide the best assignment of on-demand/broadcast channels and the number of data items to broadcast. Thus, this algorithm can be implemented into efficient system tuning tools for the MSSs.

5.2. Comparison of channel allocation methods

There are several factors affecting the data access performance of the mobile computing systems:

- Sizes of data items and messages.
- Frequency of data requests.
- Number of channels.
- Number of users.
- Number of data items.
- Bandwidth of the channels.

In order to observe the characteristics of data dissemination in mobile computing systems, we vary some of the above mentioned factors to compare the change of access time performance among the exclusive on-demand, exclusive broadcast, hybrid channel allocation,[7] and dynamic channel allocation methods. The performances are obtained by evaluating the analytical formulae developed above. The frequency of data access is controlled such that 25% of access is on the top 50 data items, 50% of the access is on top 174 data items, and almost all of the access is on the top 500 data items. Thus, unless otherwise specified, we conveniently assume the database contains 500 data items. This

[7] The hybrid channel allocation method is based on [16]. We only show experimental results for 25 on-demand channels and 25 broadcast channels due to space constraint. Other hybrid configurations of the channels have consistent results.

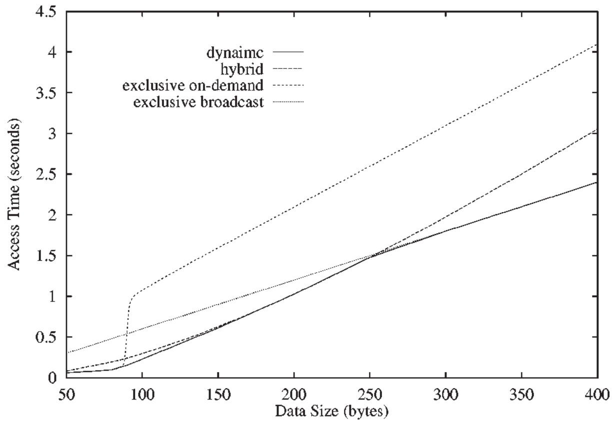

Figure 10. Access time vs. data size.

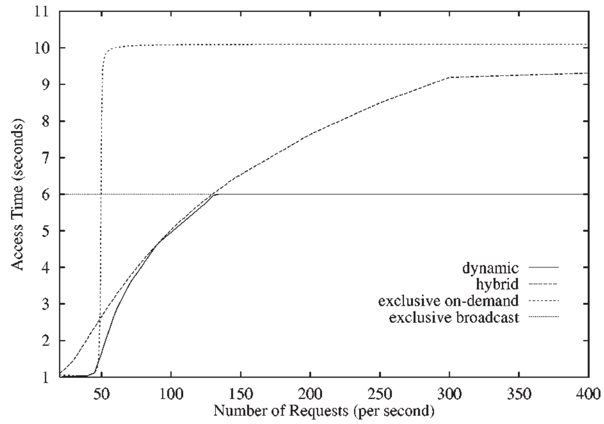

Figure 11. Access time vs. request arrival rate.

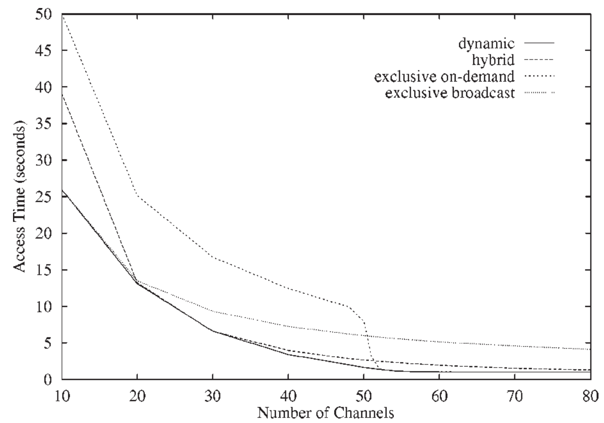

Figure 12. Access time vs. number of channels.

assumption is to facilitate the delivery of exclusive broadcast service, which is not feasible for serving the rarely accessed data items in the database.

First, we increase the size of data items from 50 bytes to 400 bytes, while fixing the other parameters. Figure 10 shows the access time of the exclusive on-demand, exclusive broadcast, hybrid and dynamic channel allocation methods.

The access time of exclusive on-demand service shows that the system is rapidly loaded after the data size is greater than 90 bytes. The leap on the access time of on-demand service is due to the increase of system workload which in turn is caused by the increase of waiting time for mobile computers to connect with the MSS. On the other hand, the access time of broadcast services is proportional to the size of the data items broadcast. With 50 channels, the access time of the broadcast method is worse than that of the on-demand method when the data access load is light, i.e., when the data item size is less than 90 bytes, and is better than that of the on-demand method when the load is heavy. The hybrid channels allocation method is realized by broadcasting some of the hottest data items on broadcast channels. Thus, the broadcast channels relieve a significant part of the system workload off the on-demand channels and improve the overall system performance. While the hybrid method outperforms the exclusive methods in most circumstances, the dynamic method has the optimal performance all the time.

The data request frequency is another factor which influences the system workload, and thus the access time of the on-demand channels. Figure 11 shows that, for the exclusive on-demand method, when the request arrival rate is over 50 data items per second, the system is overloaded. When the request arrival rate is less than 50 data items per second, the access time for on-demand channels is close to the transmission time for a data item. Since we assume that the mobile computers will not attempt to issue new requests before their previous requests are completed. Therefore, the access time for an overloaded system is constantly close to 10.01 seconds. On the other hand, the broadcast channels have the advantage of not being affected by the frequency of data access. No matter how frequently the broadcast data items are requested, the access time for broadcast channels remains constant. The hybrid method has a pretty good performance in general. When the request arrival rate is between 50 and 150 per second, its performance is better than both of the exclusive on-demand and exclusive broadcast methods. Wisely, the dynamic channel allocation method chooses exclusive on-demand when the system workload is light (i.e., less than 50 requests per second), adapts itself to the hybrid method (i.e., 25 broadcast channels and 25 on-demand channels at around 90 requests per second) and switches to the exclusive broadcast method when the system workload is heavy (i.e., greater than 150 requests per second). By adjusting its channel allocation dynamically, the dynamic method maintains the best access time performance. Note that if we vary the number of mobile computers, we obtain similar results, since the number of mobile computers has a direct impact on the request arrival rate.[8]

Next, we vary the number of channels available in the system to observe the change in access time for the channel allocation methods. Figure 12 shows that the access time of broadcast channels drops rapidly when the number of broadcast channels increases from 1 to 20. From then on, however, increasing the number of broadcast channels has

[8] Assuming the request arrival rate for each computer is fixed.

a diminishing impact on the broadcast channel access time. This observation suggests not to invest too many broadcast channels in order to lower the access time on broadcast channels. For the on-demand services, the number of channels used has a dramatic impact on the access time performance. In other words, with a small increase of the channels at the dividing point (number of channels is around 50) between a heavy-loaded system and a light-loaded system, the system access time may improve significantly. The hybrid method, though not better than the exclusive methods under some circumstances, has a very good performance in general. Finally, as expected, the dynamic channel allocation method has the best access time all the time.

One of the most important differences between the on-demand and broadcast channel services is that the broadcast channels may service any number of computers, while the access time of the on-demand channels gets worse when more computers are in the system. In contrast to the request frequency and the number of mobile computers in the systems, which have no impact on the access time of broadcast channels, the number of data items accessible through the on-demand services has no impact on the access time of on-demand services. Therefore, on-demand channels are good for providing user access to a large database, though scaling up with the number of the mobile computers is a problem. The broadcast channels are good at providing access for an arbitrary number of users. However, the access time may be unacceptable with a large database to broadcast.

Figure 13 shows the access time for various services corresponding to the number of data items accessed through the services. We observe that, with less than 900 data items, the exclusive broadcast channel does better than the exclusive on-demand channel. With more than 900 data items, the exclusive broadcast method has poorer access performance than the exclusive on-demand method which always gives the same performance independent of the number of data items. If the user accesses focus on a small portion, off-loading this portion of the on-demand data items to the broadcast channels will improve the overall system performance. This is evident by the performance of both hybrid and dynamic channel allocation methods shown in the figure, i.e., when the number of data items is smaller than 1300, the hybrid method is better than the exclusive methods, while the dynamic channel allocation always gives the best choice. In our model, we assume the user accesses areas will be enlarged as the number of data items increases. Therefore, if the number of data items is more than 1300, the user accesses will be distributed to a large number of data items, shifting on-demand channels to broadcast channels is not going to help the system, the exclusive on-demand channel allocation is the best choice.

In all of the above experiments, except for the one in which the number of data items is varied, the performance of on-demand channels has certain turning point due to the change of system workload. The turning point is corresponding to the system traffic intensity, $\rho = 1$. The factors varied in the experiments have an impact on the traffic intensity ρ, thus causing a dramatic change in the data access time.

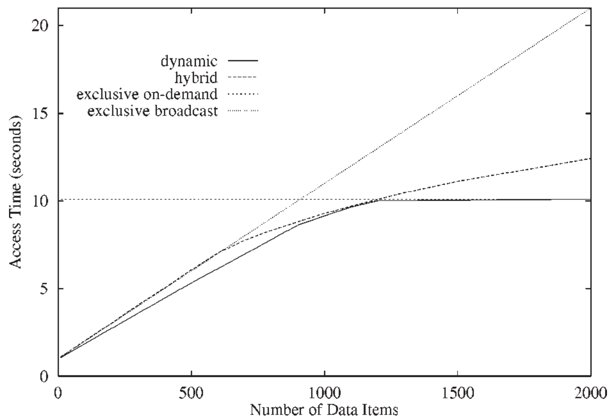

Figure 13. Access time vs. number of data items.

On the other hand, the performance of broadcast channels has no such abrupt change. In the studies, we have deliberately chosen the database size to be small, i.e., 500 data items, so that it is feasible to disseminate them on broadcast channels. In a realistic environment, the database usually is quite large, and the exclusive broadcast mode may never be adopted. Under certain circumstances, when many users are accessing the same set of data items, e.g., during the business hours, the downtown office building areas will be loaded with requests for business information such as stock prices, changing some of channels into broadcast mode for disseminating these hot data items can avoid jamming the MSSs services.

6. Conclusion

Mobile computing has received a lot of attention from the industry and academia in the past few years. Due to the limited wireless communication bandwidth allocated for mobile computing, researchers are seeking techniques to make efficient use of the available bandwidth. Wireless broadcast is an important technique for disseminating data to mobile users, because it scales up to an arbitrary number of users. Thus, broadcast channels are often used for broadcasting commonly requested data items to complement the traditional approach of data access through the on-demand channels. The issue of wireless channel allocation for data dissemination is to monitor the user access pattern and dynamically assign channels for broadcast and on-demand services to achieve optimal system performance.

The scheduling strategy we proposed for data dissemination on the broadcast channels is different from the strategies proposed in the literature in that we choose to broadcast every hot data item with equal frequency in order to prevent unlimited access delay for less popular items and that unpopular data items are provided by on-demand channels only.

Analytical models have been provided for broadcast and on-demand channels. Cost formulae for the access time of the exclusive broadcast, exclusive on-demand, and dynamic channel allocation methods are developed based on the analytical models. Based on the analysis, an optimal channel allocation algorithm is proposed. Experimental studies show that our analytical models match well with simulative results and that the optimal channel allocations obtained via our algorithms conform with the results obtained through simulation. The result shows that the dynamic channel allocation method improves the system performance significantly when the system is heavy-loaded. Thus, it shows the importance of dynamic channel allocation to efficient usage of the wireless bandwidth.

We compare the access time performance of the exclusive on-demand, exclusive broadcast, hybrid, and dynamic channel allocation methods to observe their access time behavior corresponding to system parameters such as the size of data items, frequency of data requests, number of channels used for the services, bandwidth of the channels in the system, and number of data items accessed through the services. The results show that the broadcast channels outperform the on-demand channels when the data size is large, the data access frequency is high, and when the bandwidth of the channels is low. In other words, the broadcast method performs very well when the system is overloaded, while the on-demand method performs very well when the system is light-loaded. Under all circumstances, the hybrid method has a very good access time because it mixes the on-demand and broadcast channels to complement each other. The dynamic channel allocation method performs the best, since it has the flexibility to assign channels to either broadcast or on-demand mode to optimize performance.

We observed that the exclusive broadcast method does not scale up with the number of data items broadcasted. Therefore, for uniformly distributed data accesses, on-demand channels are more efficient than broadcast channels. Both broadcast and on-demand channels are sensitive to the number of channels used. The broadcast channel access efficiency may be significantly improved by adding the first few more channels. However, after some point, increasing the number of channels does not improve the efficiency much. On the other hand, the access time of on-demand channels might be dramatically improved by adding a few channels to the on-demand service.

As for the future work, we will study the global channels management issues for neighboring cells, in which channel borrowing is allowed. Moreover, use of the broadcast channels as shared cache for the mobile users is an interesting idea. A coherent approach for the caching, invalidation and maintenance of data items through a combination of the broadcast and on-demand channels needs further research.

Appendix

In this paper, we use a M/M/c/m queue to model on-demand service of the mobile computing systems. In section 3.1.1 we simplify the access time formulae of the M/M/c/m queue by approximation (equation (6)). Then in section 3.1.2, when the system workload is light, we use a M/M/c queue to replace the M/M/c/m and further simplify the access time formulae (equation (8)). In this appendix, we conduct simulations to show that the approximations are acceptable.

Table 3
Approximation of different queuing and analytical models.

ρ	Simulated results	Analytic results	Heavy approx. error ratio	Light approx. error ratio
0.20	1.01	1.26	–	1.01 (0.0%)
0.60	1.01	1.10	–	1.01 (0.0%)
0.80	1.07	1.13	–	1.08 (1.0%)
0.95	1.80	1.82	–	1.77 (1.7%)
0.96	2.01	2.07	–	2.02 (0.5%)
0.97	2.48	2.48	–	2.44 (1.6%)
0.98	3.23	3.32	–	3.28 (1.6%)
0.99	5.45	5.65	–	5.83 (7.0%)
1.01	20.42	20.40	25.25 (23.8%)	–
1.05	24.26	24.24	25.25 (4.1%)	–
1.10	24.72	24.75	25.25 (2.1%)	–
1.11	24.80	24.79	25.25 (1.8%)	–
1.13	24.86	24.86	25.25 (1.6%)	–
1.15	24.90	24.91	25.25 (1.4%)	–
1.20	24.89	25.00	25.25 (0.4%)	–
2.00	25.25	25.20	25.25 (0.0%)	–

Intuitively, when the system is lightly loaded, we can assume that there is no request lost (i.e., all arrival requests can get serviced within bounded time). Therefore, the M/M/c queue has similar mean response time as the M/M/c/m queue which consists of service time $1/\mu$ and waiting time. For a light-loaded system, the waiting time is very small. According to the comparison studies made in sections 5.1 and 5.2, the similarity between the simulated results and analytical results suggested that the M/M/c queue is a good approximation of the M/M/c/m queue when $\rho < 1$ and $m \gg c$.

To clearly illustrate how good the approximation of the M/M/c queue and the M/M/c/m queue is, we performed an experiment where we assumed that there are 20 on-demand channels in the cell and varied the traffic intensity ρ. We assume that there are on average 500 mobile computers in the cell.

Table 3 lists the access time obtained via experiments and analytical study. The first column gives the different values of ρ, the second column shows the results obtained through simulation experiments,[9] the third column shows the analytical results obtained via equation (4), and the last two columns show the results obtained by using equations (6) and (8) for $\rho > 1$ and $\rho < 1$, respectively. The error ratios with respect to the simulative results are also shown in the last two columns of the table to illustrate how good the approximations are. It is obvious that the approximated results are acceptable when $\rho \leqslant 0.98$ and

[9] The simulation program is the same as what we used in sections 5.1 and 5.2.

$\rho \geqslant 1.11$. The error ratios are less than 2.0%. Thus, we can conclude that the approximations match well with the M/M/c/m queue.

References

[1] S. Acharya, R. Alonso, M. Franklin and S. Zdonik, Broadcast disks: Data management for asymmetric communications environments, in: *Proc. ACM SIGMOD Conf. on Management of Data*, San Jose, California (1995).

[2] S. Acharya, M. Franklin and S. Zdonik, Dissemination updates on broadcast disks, in: *Proc. 22nd VLDB Conference*, Mumbai (Bombay), India (1996).

[3] S. Acharya, M. Franklin and S. Zdonik, Prefetching from a brodcast disk, in: *Proc. Int. Conf. on Data Engineering*, New Orleans, LA (February 1996).

[4] S. Acharya, M. Franklin and S. Zdonik, Balancing push and pull for data broadcast, in: *Proc. ACM SIGMOD Conf. on Management of Data*, Tuscon, Arizona (May 1997) pp. 183–194.

[5] T.C. Chuieh, Scheduling for broadcast-based file system, in: *NSF MOBIDATA Workshop at Rutegers University* (1995).

[6] N.J. Colmenares, The FCC on personal wireless, IEEE Spectrum 31(5) (May 1994) 39–46.

[7] T.H. Cormen, C.E. Leiserson and R.L. Rivest, *Introduction to Algorithms* (McGraw-Hill, New York, 1990).

[8] D.J. Goodman and T. Imielinski, Future directions for wireless data, *Digest of Papers, Spring COMPCON '94* (March 1994) 464–466.

[9] S. Hameed and N.H. Vaidya, Efficient algorithms for a scheduling single and multiple channel data broadcast, Technical Report 97-002, Computer Science, Texas A&M University (February 1997).

[10] Q.L. Hu, D.L. Lee and W.-C. Lee, Optimal channel allocation for data dissemination in mobile computing environments, in: *Proc. 18th Int. Conf. on Distributed Computing Systems (ICDCS '98)*, Amsterdam, The Netherlands (May 1998) pp. 480–487.

[11] T. Imielinski and S. Viswanathan, Adaptive wireless information systems, in: *Proc. SIGDBS (Special Interest Group in DataBase Systems) Conf.*, Tokyo, Japan (October 1994).

[12] T. Imielinski, S. Viswanathan and B.R. Badrinath, Energy efficiency indexing on air, in: *Proc. Int. Conf. on SIGMOD* (1994) pp. 25–36.

[13] T. Imielinski, S. Viswanathan and B.R. Badrinath, Power efficiency filtering of data on air, in: *Proc. Int. Conf. on Extending Database Technology* (1994) pp. 245–258.

[14] R. Jain, *The Art of Computer Systems Performance Analysis* (Wiley, New York, 1991).

[15] S. Jiang and N.H. Vaidya, Scheduling algorithms for a data broadcast system: Minimizing variance of the response time, Technical Report 98-005, Computer Science, Texas A&M University (February 1998).

[16] W.-C. Lee, Q.L. Hu and D.L. Lee, Channel allocation methods for data dissemination in mobile computing environments, in: *Proc. 6th IEEE Int. Symposium on High Performance Distributed Computing* (August 1997) pp. 274–281.

[17] W.-C. Lee and D.L. Lee, Using signature techniques for information filtering in wireless and mobile environments, Special Issue on Database and Mobile Computing, Journal on Distributed and Parallel Databases 4(3) (July 1996) 205–227.

[18] H. Schwetman, *Csim User's Guide (Version 17)* (MCC Corporation, 1992).

[19] K. Stathatos, N. Roussopoulos and J.S. Baras, Adaptive data broadcast in hybrid networks, in: *Proc. 23rd VLDB Conf.*, Athens, Greece (1997) pp. 326–335.

[20] S. Tekinay and B. Jabbari, Handover and channel assignment in mobile cellular networks, *IEEE Communications Magazine* (November 1991) 42–46.

Wang-Chien Lee received the B.S. degree in information science from National Chiao Tung University, Hsin-Chu, Taiwan, in 1985; the M.S. degree in computer science from Indiana University, in 1989, and the Ph.D. degree in computer and information science from the Ohio State University in 1996. He has been a senior member and then a principal member of technical staff in the Advanced Systems Laboratory of GTE Laboratories in Waltham, Massachusetts, since June 1996. In 1995, he worked as a research assistant in the Computer Science Department at the Hong Kong University of Science and Technology. His current interests are in the areas of mobile computing, object-oriented database systems, data warehousing, Internet information retrieval and integration, and telecommunications management network. Dr. Lee is a member of the Association for Computing Machinery, the IEEE Computer and Communications Societies, and the Upsilon Pi Epsilon Association.
E-mail: wlee@gte.com

Qinglong Hu was born in Hefei, China, in 1964. He received the B.S. and the M.S. degree in computer science from East China Normal University, Shanghai, China, in 1986 and 1989 respectively, he is currently studying toward a Ph.D. degree at the Hong Kong University of Science & Technology. His research interests include mobile computing and distributed system.
E-mail: ginglong@cs.ust.hk

Dik Lun Lee is Reader of the Computer Science Department at Hong Kong University of Science and Technology. Prior to joining HKUST, he was an Associate Professor of Computer and Information Science at the Ohio State University. In 1992, he was a Distinguished Visiting Scholar at the Online Computer Library Centre (OCLC, Dublin, USA) and a consultant to the Information Dimension Incorporated (IDI, Dublin, USA) on the BasisPlus document management system. Dr. Lee's research interest includes document management, access techniques and query processing for text, indexing and search on World Wide Web, object-oriented databases, and mobile computing. He received his Ph.D. in computer science from the University of Toronto. He is currently the Chairman of the ACM Hong Kong Chapter.
E-mail: dlee@cs.ust.hk