社団法人　情報処理学会　研究報告
IPSJ SIG Technical Report

2006－DBS－138（13）
2006－GN－58（13）
2006／1／27

放送・通信融合環境における放送スケジューリングと基地局キャッシングの連携手法

蔡 菁†　寺田 努†　原 隆浩†　西尾章治郎†

†大阪大学大学院情報科学研究科マルチメディア工学専攻
〒565–0871 大阪府吹田市山田丘 1–5
E-mail: {cai, tsutomu, hara, nishio}@ist.osaka-u.ac.jp

本稿では，放送・通信融合環境における放送スケジューリングおよび基地局キャッシングを効率よく連携させる手法を提案する．放送・通信融合環境とはプッシュ型放送通信，プル型放送通信，プル型無線通信の三つの通信形態を統合的に利用することで効率的な問合せ処理を行う環境である．提案手法では，放送スケジューリングの作成と基地局におけるキャッシュ管理を連動させることで，効率のよい問合せ処理を実現する．また，提案方式を用いることでシステムの性能が向上することをシミュレーション評価により確認した．

A Joint Strategy of Broadcast Scheduling and Base Station Caching in the Hybrid Wireless Broadcast Environment

Jing CAI†, Tsutomu TERADA†, Takahiro HARA†, and Shojiro NISHIO†

† Dept. of Multimedia Eng., Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita-shi, Osaka, 565-0871 Japan

Hybrid Wireless Broadcast (HWB) model, which combines push- and pull- based broadcast with pull-based point-to-point wireless communication, has been confirmed to be able to enhance the system performance. In this paper, based on the HWB model, we investigate a joint control strategy, which makes effective use of the HWB data dissemination and cooperate broadcast scheduling with cache management of the base station. Simulation study demonstrates that our proposed mechanism can further improve the system performance.

1 Introduction

1.1 Background

Various techniques have been developed to improve the performance of wireless information services, such as point-to-point data dissemination, information broadcasting, and data caching. Many studies have been conducted on the individual technique. Recent advances in computer and wireless communication technologies have increased interest on the combination of these techniques.

There are some researches augmenting the pure push-based broadcast, integrated with the pull-based broadcast or with the pull-based wireless data dissemination[2, 5, 6]. Moreover, our previous study [3] proposed a novel wireless communication model, namely Hybrid Wireless Broadcast (HWB) model, which combines push- and pull- based broadcast with pull-based point-to-point wireless communication to provide a flexible and complementary information service in different bandwidths and service ranges.

Furthermore, data caching has been widely used in many broadcasting systems. Some of studies developed the joint control methods by integrating cache management and broadcast scheduling. Broadcast Disks (BD) was well known on integrating multi-disks scheduling and cost-based PIX caching[1]. Additionally, Su et al. proposed an integrated method to produce the broadcast schedule and the client's prefetching scheme[8]; Ercetion et al. addressed the joint control in two stage satellite-terrestrial wireless broadcast system[4]. In these studies, it has been confirmed that the joint control can provide more efficient data delivery, by integrating broadcast scheduling and data caching.

However, most of researches investigated the joint control methods based on the push-based broadcast, and normally conducted cache management at the client side. So far, there has been no study discussing the integrated control based on the hybrid broadcasting environments. Therefore, this paper aims at constructing the efficient joint control under the combined HWB environment by integrating broadcast scheduling and cache management of base station.

1.2 HWB Model

We briefly sketch the HWB model[3] and its features in this section, since the work of this paper is based on the HWB model.

The HWB model comprises a broadcast server, lots of the base stations, and a large number of clients. The bandwidth for broadcast is classified as the main channel and the on demand sub channel. Additionally, each base station connects with the broadcast server through the Internet, and serves the local clients via the wireless channel by using the base station cache. The three data delivery ways of the HWB model have different features.

For the push-based broadcast of the main channel, the response is not affected by the load, namely, even the number of clients increases, the cost of broadcast has no change. However, clients may not be able to acquire the reply quickly even though the total number of queries is small, since they need to wait until the desirable broadcast arrives. The average response time depends on the number of the broadcast items.

The pull-based broadcast of the sub channel can meet the individual requirements of the global clients; at the same time any on demand response can be accessed by an arbitrary number of clients. Therefore, it is efficient to respond to the queries which require for the same data. However, the response depends on the number of queries which request for the different items, namely, the waiting time increases with the number of different requests.

In contrast, the pull-based wireless communication can meet the individual requirements of the local clients, by using the base station cache. However, unlike the on-demand sub channel, the on demand response through the wireless channel cannot be shared due to the point-to-point communication. Hence, the response depends on the number of requests, no matter they require the same item or not.

Among the three data delivery ways of the HWB model, how to minimize the impact of the defect of one mode by exploiting the advantages of the others? It is helpful to determine more effective system control, if taking these features into account.

1.3 Broadcast Scheduling and Caching

Traditionally, the server broadcasts different items at a same frequency, namely flat broadcast, regardless of their relative importance to the clients. By comparison, in multi-disks broadcast[1], important items can be broadcast more often than others; normally, hot items are determined to be put to the fast disks.

On the other hand, caching is used to improve performance by keeping frequently accessed or expensive data in memory. The basic problem for cache management is to determine which items should be cached to give the best responsiveness to the clients' requests. In pull-based systems, this is typically achieved by storing the items with the highest access probability, namely probability-based caching. Whereas in push-based multi-disks systems, a cost-based PIX caching is efficient to cooperate with multi-disks scheduling. In such a system, caching strategy cannot favor items that are frequently broadcast, since those items can be loaded within a short time. However, it remains an unclarified question which is more suitable to the HWB system, probability-based caching or PIX caching.

Generally cache management is performed by the client side, of which the role of cache is to reduce the response latency for the client itself. In contrast, the HWB system employs the base station cache, by which serves the local client community rather than an individual client. Therefore, caching of base station aims at improving the average response of the local client community by increasing the cache hits of base station. Moreover, cache management in the HWB environment is not a solely caching problem, which relates to the broadcast scheduling.

Accordingly, to construct a more effective joint control strategy of the HWB environment, it is necessary to consider the special role of base station cache and its cooperation with the push-based broadcast and the pull-based broadcast.

The HWB model provides two pull-based data delivery ways, and performs more effectively under the skewed access[3]. Considering probability-based caching is suitable for the pull-based system; meanwhile, multi-disks scheduling is a good program for the non-uniform access, we attempt to employ multi-disks scheduling and probability-based caching as the basic processing mode into our joint control of the HWB system.

1.4 Overview of the Paper

In this paper, we propose several joint control strategies which take into account the complementary features of HWB data dissemination and the cooperation of base station caching and broadcast scheduling. Additionally, we investigate the following interrelated issues by evaluating the system performance under different strategies:

- Whether multi-disks broadcast performs better than flat broadcast in the HWB environment?
- Which is the optimal way to process the data items with high access probability? Keeping them in the base station cache or broadcasting them with high broadcast frequency?
- Whether it is necessary to assign the whole data items for broadcasting? How about only broadcasting the date items without cache-resident?

All the work in this paper is assumed based on the HWB model. In addition, we make the following assumptions:

- There is no update either on data of the database and access pattern of clients.
- Clients' access pattern vary according to the base station areas they belong to.
- Broadcast server and base station have the knowledge of the clients' access pattern. Additionally, broadcast server is aware of the cache of each base station.
- Size of the base station cache is relatively smaller than that of the interest data set of the local clients.
- Client has no capability of caching, like PDA.

The rest of paper is organized as follows. In section 2, we present our integrated control and several implementary strategies of the HWB system. In section 3, we show the evaluation results of the simulation study and give a further discussion on the joint control of the HWB system. Section 4 introduces some related works. Finally, we summarize this paper in section 5.

2 Joint Control Policy

2.1 Access Probability

Our joint control of the HWB system depends on the clients' access probability, which is distinguished as the Local Access Probability (LAP) and the Global Access Probability (GAP). LAP plays an important role in the base station caching. While GAP has much more relation with the broadcast scheduling.

The requests for item i from base station b generates the local access probability $P_b(i)$, where $1 \leq i \leq N$ and $\sum_{i=1}^{N} P_b(i) = 1$. Here N denotes the total number of the data items in the database. Additionally, the requests for item i from each base station are aggregated to the global access probability *P(i)*, which is calculated by

$$P(i) = \frac{1}{M} \sum_{i=1}^{M} P_b(i),$$

where M denotes the total number of base stations, and $\sum_{i=1}^{N} P(i) = 1$ is also satisfied.

The clients in different regions have an heterogeneous access, while the clients in a same geographical area often have similar interests in the local information, for instance, a location dependent query. Therefore, some items may be hot in one area, but cold at other areas. Taking these into account, we assume that all the items of the database are divided into several data groups in accordance with the base stations. Requests in each base station have a high tendency (k) to issue the skewed queries, and a low tendency to send uniform queries. The skewed queries follow Gaussian distribution $Gau(i, \mu, \sigma)$ with the center of hot spot μ and deviation σ. The value of μ is set as the center of each data group, while the value of σ can be varied to reflect the different skewness of the clients' queries. Let $F_b(i)$ denote the query mode of the base station b for item i, which is given by

$$F_b(i) = k \times Gau(i, \mu, \sigma) + (1 - k)/N.$$

Hence, the local access probability is estimated by

$$P_b(i) = F_b(i) / \sum_{j=1}^{N} F_b(j).$$

2.2 Joint Control in HWB

Figure 1 indicates the control model of the HWB system. Generally speaking, data items in the database can be assigned to the schedule of the main channel to be broadcast, or be placed into the on-demand sub channel to respond to the queries of the global clients, or be kept in the base station cache and on-demand transmitted by the wireless channel to respond to the queries of the local clients. It is essential to choose the optimal processing way, but it may be unrealistic to assume data organization of the system complies with the individual access pattern of all clients. Therefore, we focus on processing the items with high access probability (i.e., hot items), for which to choose the optimal way.

The broadcast server serves clients in a global area, whereas the base station provides services only for the local clients. Accordingly, we tend to keep the items with the highest local access probability in the base station cache. It is good at increasing the cache hits

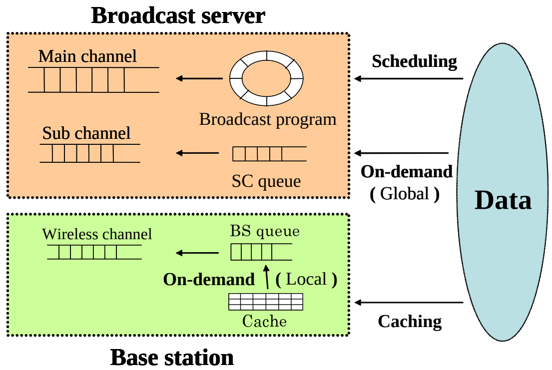

Figure 1: Control Model

if keeping the local hottest items in the base station cache, since the size of base station cache usually is much smaller than that of the database, also smaller than that of the interest set of the local clients.

On the other hand, the server needs to be aware of the base station caching in order to manage the broadcast program. In view of the multi-disks broadcast, if the hottest items of the local area is retained in the cache of the base station, these items should be pushed to the slower disk, otherwise the server wastes a significant portion of the fastest disk for items already in the cache of the base station.

For the meanwhile, chopping off part of the broadcast schedule, namely shortening the broadcast cycle, has the effect of increasing the available bandwidth. Therefore, we consider the processing: all the items that have resided in the base station cache will not be broadcast any more. It is because in the HWB environment clients can pull the items that are disappeared in the main channel through the on-demand sub channel or the on-demand wireless channel.

Consequently, we design the following procedures for our joint control by using the LAP and the GAP.

1. Retain the hottest items with the largest estimated LAP values in the base station cache
2. Modify the LAP values for all the items resided in the base station cache
3. Aggregate the GAP by using the modified LAP of each base station
4. Construct a broadcast program with the newly aggregated GAP

To reflect the base station cache, once an item is cached, its LAP and GAP value will be modified. Additionally, in terms of multi-disks scheduling, it needs to rank all the items with the modified GAP to generate a broadcast disk program. By greatly decreasing the corresponding LAP and GAP values, the cached items are impossible to be pushed to the fastest disk. Moreover, if the modified GAP values of some items become zero, these items will not be broadcast any more.

2.3 Control Strategies

To investigate how our control idea could be implemented efficiently, and what influence on the system performance under different processing, we design four joint control strategies as follows.

GAP-MD(no_Ca) Strategy

In this strategy, the global hottest items with the largest GAP values are kept in the cache of each base station; hence all base stations have an identical cache. Moreover, the LAP values for all the cached items are set as zero, i.e.,

$$P_b(i) = 0 \mid i \in Cache,\ b = 1 \sim M,$$

where $Cache = Cache_1 = \cdots = Cache_b = Cache_M$. Here $Cache_b$ denotes the set of items kept in the cache of the base station b ($1 \leq b \leq M$), while $Cache$ denotes the total set of items kept in each base station cache. Therefore, the GAP values for those cached items also become zero. That is to say, all the items identically cached in each base station will not be broadcast any more. The other items without cache-resident adopt multi-disks scheduling.

LAP-MD(no_Ca) Strategy

In this strategy, each base station keeps the local hottest items with the largest LAP values in the cache. Differing from the strategy GAP-MD(no_Ca), cache of each base station is different, since each base station has distinct LAP. Moreover, the LAP values for all the items resided in any base station cache are set as zero, i.e.,

$$P_b(i) = 0 \mid i \in Cache,\ b = 1 \sim M,$$

where $Cache = Cache_1 \cup \cdots \cup Cache_b \cup Cache_M$. Therefore, the GAP values for those items resided in any base station also become zero. In other words, the items once cached in any base station will not be broadcast any more. The items without any cache-resident adopt multi-disks scheduling.

LAP-FT((no_Ca)) Strategy

The cache management of this strategy is the same as the LAP-MD(no_Ca), which also satisfies

$$P_b(i) = 0 \mid i \in Cache,\ b = 1 \sim M,$$

where $Cache = Cache_1 \cup \cdots \cup Cache_b \cup Cache_M$. Thus the items once cached in any base station will not be broadcast any more. The difference is the broadcast schedule; in this strategy, the items without any cache-resident adopts the flat broadcasting rather than multi-disks scheduling.

LAP-MD(all) Strategy

This strategy also let each base station keep the local hottest items with the largest LAP values in the cache, however, differing from the above strategies, which only modifies the LAP values for the items resided in its own base station cache as zero, i.e.,

$$P_b(i) = 0 \mid i \in Cache_b,\ b = 1 \sim M.$$

Hence, the GAP values of the cached items are just greatly decreased, but not become zero. Consequently, all the items in the database will be broadcast with the multi-disks scheduling, based on the ranking of the modified GAP values.

In addition, we introduce two traditional strategies into our evaluation to compare with our proposed joint control strategies.

Table 1: Parameter Settings

Parameters	Values
Database Size [Data Items]	10,000
Data Item Size [KB]	100
Number of Base Stations	10
Cache Size of BS [Data Items]	200
Main Channel Bandwidth [Mbps]	100
Sub Channel Bandwidth [Mbps]	10
Wireless Bandwidth [Mbps]	5
Time Slot [D/Bm]	30,000
Query Interval [ms]	50∼1,500
Data Group Size [Data Items]	1,000
Query Tendency [%]	80
Deviation for Gaussain	150
Number of Disks	3
Broadcast Frequency of $Disk_{1,2,3}$	4, 2, 1
Size of $Disk_{1,2,3}$ [Data Items]	500, 1,000, uncertain

PIX-MD(all) Strategy

This strategy, employing the traditional BD approach, broadcasts all the items in the database with multi-disks scheduling, while the cache management of base station adopts PIX caching. Therefore, the global hottest items are broadcast with the fastest broadcast frequency. Thus, this strategy completely differs from our proposed strategies.

LAP-FT(all) Strategy

The cache management of this strategy is the same as the strategy LAP-FT(no_Ca), namely keeping the local hottest items in the base station cache, moreover both of them adopt flat broadcast. However, this strategy broadcasts all the items in the database, instead of only broadcasting the items without cache-resident in LAP-FT(no_Ca). This strategy was adopted in our previous study of the HWB system[3], approximately.

3 Simulation Experiments

In order to evaluate the performance of the above mentioned strategies, we conduct the simulation study. The performance metrics are the average waiting time and the success rate of queries. Table 1 presents the default system parameter settings used in the experiments. The number of items in the database is 10,000; all data items have an equal size 100KB. The number of base station is 10; each base station has a cache with 200 items. The bandwidth for the main channel, sub channel, and wireless channel are 100Mbps, 10Mbps, and 5Mbps, respectively. In the case of the multi-disks scheduling, the number of multi-disks is 3, the broadcast frequency of the three disks are 4, 2, and 1. To gain an objective evaluation result, the size of the fastest disk is identically set as 500 items, while the medium disk is 1,000 items. On the other hand, the size of the slowest disk and broadcast cycle depend on the strategy. Comparing the broadcast cycle of all strategies, LAP-FT(no_Ca) is the shortest, while LAP-MD(all) and PIX-MD(all) are the longest. In addition, LAP-MD(no_Ca) is the shortest among all the strategies adopting multi-disks scheduling.

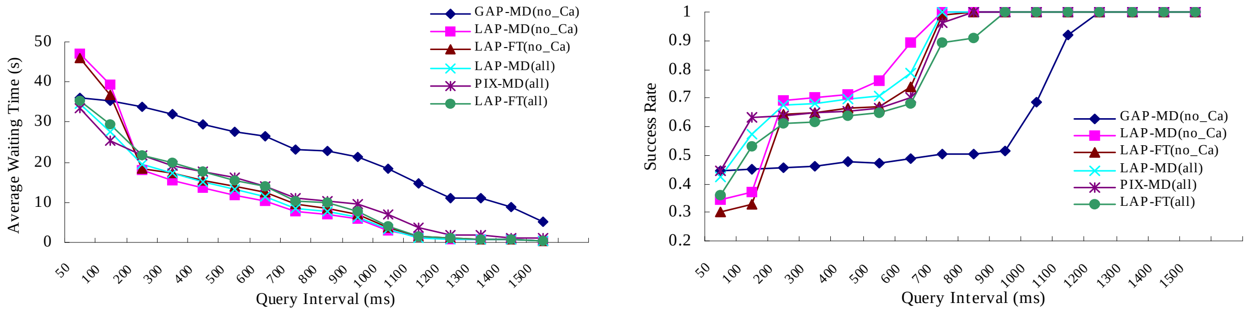

Figure 2: Impact of Query Interval

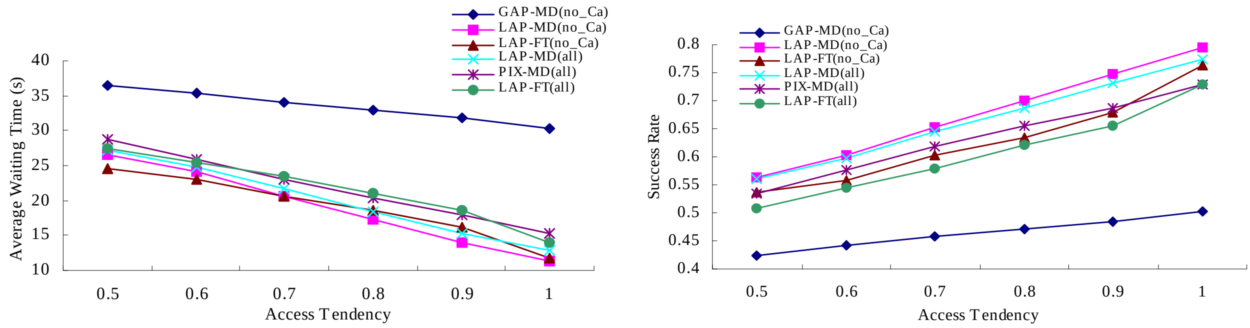

Figure 3: Impact of Access Tendency

In our simulation, the base station takes responsibility for the query processing[3]. When the base station receives a query, it needs to compare the three waiting times for the main channel, the sub channel, and wireless channel and take a corresponding action according to the selection with the shortest waiting time.

3.1 Impact of Query Interval

In the first experiment, we evaluate the system performance under different query intervals. Figure 2 shows that as the query interval increases, i.e., workload decreases, the performances of the average waiting time and the success rate upgrade for all the strategies. Notice that our proposed strategy LAP-MD(no_Ca), LAP-FT(no_Ca) and LAP-MD(all) occupy the top three among all the strategies at the main phase of the evaluation, and LAP-MD(no_Ca) always performs the best.

The reasons for these behaviors are as follows. These three strategies all keep the local hottest items in the base station cache, which can provide low-latency response to the local clients who usually have a similar access pattern. However, these strategies differ in the broadcast scheduling. LAP-MD(no_Ca) and LAP-MD(all) both adopting multi-disks scheduling though, LAP-MD(all) broadcasts the whole items, while LAP-MD(no_Ca) only broadcasts the items without any cache-resident. By shortening the broadcast cycle and by utilizing the two on-demand channel to pull the items disappeared in the main channel, LAP-MD(no_Ca) can make more efficient use of the three bandwidths and thus has the best performance. By comparison, LAP-FT(no_Ca) adopts flat broadcast. In the case of the skewed access, and the cache of base station can only keep part of hottest items, it is more efficient to push the remained hot items to the fastest disk rather than flat broadcast them.

As for GAP-MD(no_Ca), it always has the poorest performance, due to the lowest cache hits. Only when the access load is extremely heavy, it has a better performance. It is because in this case the load of the wireless channel for the strategies with a higher cache hits becomes much heavier. However, in GAP-MD(no_Ca), most queries are served by the frequent broadcast, since it broadcasts most of hottest items with the fastest disk.

On the other hand, the traditional strategies all the while performs much more poorly. The reason is that PIX-MD(all) has the longest broadcast cycle and lower cache hits, which give rise to the longer average response and lower utilization of the bandwidth. LAP-FT(all) has a higher cache hits though, it broadcasts the whole items in the database, which also cannot make efficient use of the bandwidth. Only when the access load of the system gets much heavy, PIX-MD(all) has the best performance. The best account is that in this case a large amount of queries can be better responded by the frequent broadcast of the hottest items.

3.2 Impact of Access Pattern

Next, we evaluate the impact of access pattern through two interrelated factors: the access tendency and the query deviation of the skewed access. Figures 3 and 4 show LAP-MD(no_Ca) outperforms the other

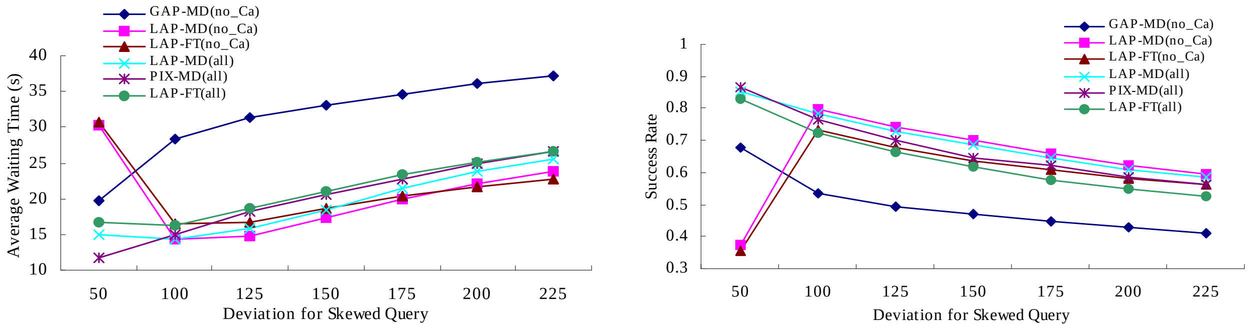

Figure 4: Impact of Query Deviation

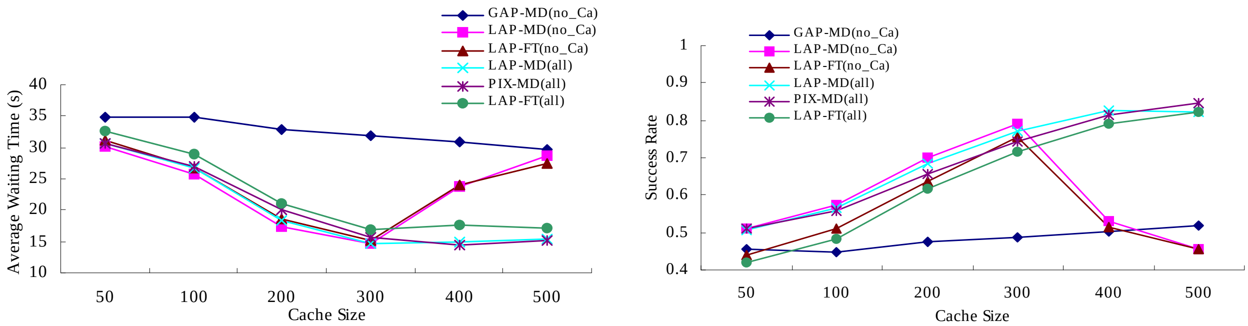

Figure 5: Impact of Base Station Cache Size

strategies under the ordinarily skewed access, i.e., access tendency is above 0.7 and query deviation is from 100 to 175. As the reason explained above, LAP-MD(no_Ca) has a higher cache hits and more efficient bandwidth usage, moreover adopting the multi-disks scheduling for the skewed access.

When the query deviation of the skewed access become extremely skewed, for example, only about 50 items, PIX-MD(all) behaves best. In this case, a large amount of queries concentrate on a quite small part of hottest items, while for the probability-based strategies, these hottest items are almost all kept in the base station cache. However, the responses are transmitted by the point-to-point wireless communication; thus, the load of wireless channel becomes greatly heavy. On the contrary, PIX-MD(all) broadcasts the required hottest items more often, and the frequent broadcast can be shared by an arbitrary number of clients.

When clients' access becomes uniform, i.e., access tendency is below 0.7, query deviation is above 175 items, LAP-FT(no_Ca) performs the best. The reason is that this strategy adopts the flat broadcast with the shortest broadcast cycle, which is more suitable to the case when queries of clients are uniform or relevant to a large number of global data. In addtition, when the access tendency of the skewed queries approaches 1, the performance of LAP-FT(no_Ca) is also increased greatly. It is because in this case the skewed queries concentrate on the local data, while most of the requested items are cached in the local base station, less response from the broadcast, and thus performance difference between the multi-disks broadcast and the flat broadcast becomes small.

3.3 Impact of BS Cache Size

We then examine the impact of cache size of the base station. As Figure 5 shows, when the size of base station cache is below 300 items, namely relatively smaller than that of the access range of the local clients, the performance of all strategies upgrades markedly as the cache size increases, and LAP-MD(no_Ca) has the best performance. It is because in this case cache hits increases with the cache size. In addition, LAP-MD(no_Ca) makes the best use of the three bandwidths of the HWB model, by keeping a part of local hottest items in the base station cache, pushing the remained hot items to the fastest disk, and removing the cached items from the broadcast.

However, when the size of base station cache approaches 300 items, the performance differences of almost all strategies become very small. Moreover, when the cache size is above 300 items, the performance of LAP-MD(no_Ca) and LAP-FT(no_Ca) drop sharply, whereas there is a little change for the other strategies. The reason is that these two strategies keep the local hottest items in the base station cache, and these items once kept in the cache will not be broadcast any more. When the cache size of the base station is so huge to keep most of local hottest items, the load of wireless channel becomes very heavy; in the meanwhile, the bandwidth utilization of the main channel gets too low. On the other hand, PIX-MD(all) performs best under a large cache size, since the hottest items are pushed to the fastest disk, and the increasing cache size has no influence on the broadcast schedule.

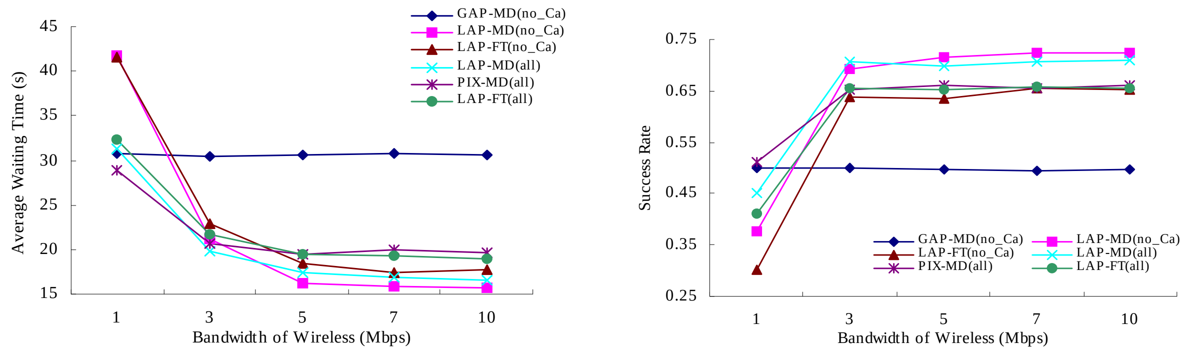

Figure 6: Impact of Bandwidth of Wireless Channel

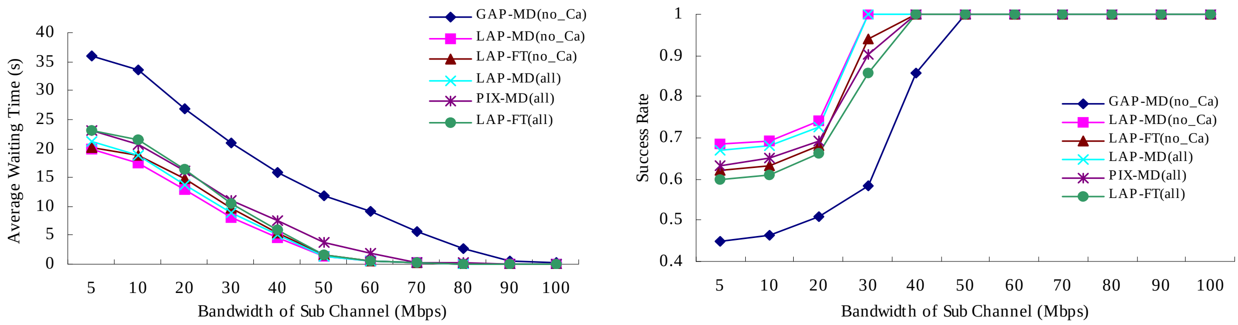

Figure 7: Impact of Bandwidth of Sub Channel

3.4 Impact of Pull Bandwidth

Finally, we examine the impact of pull bandwidth. As Figure 6 shows, when the bandwidth of wireless channel is very small, i.e., below 3Mbps, PIX-MD(all) performs the best. The reason is that the narrow wireless channel limits the strategies with a higher cache hits to behave better. Once the bandwidth becomes larger, these strategies perform better than the strategies with a lower cache hits. Our proposed strategy LAP-MD(no_Ca) performs the best when the bandwidth is above 5Mbps. Additionally, Figure 7 indicates that the performance of all strategies upgrade as the bandwidth of sub channel increases. However, when the bandwidth gets very large, the differences among almost all strategies become very small, since in this case the bandwidth of sub channel is so huge that all the strategies mainly use the sub channel to respond to the queries.

3.5 Discussion

From the above experimental results, we come to the conclusion of the evaluation as follows.

The most suitable strategy of the HWB system depends on the different situations. In the ordinary condition, i.e., when the system load is ordinarily heavy, clients' access is ordinarily skewed, and the cache size of the base station is relatively smaller than the number of local hot items, LAP-MD(no_Ca) is the most suitable strategy. In the case when the clients' access gets uniform, while system load is ordinarily heavy, LAP-FT(no_Ca) is the best choice. Additionally, in the case when the system load is extremely heavy, clients' access is extremely skewed, while the cache size of the base station is very large, or the bandwidth of the wireless channel is quite small, PIX-MD(all) is the best selection.

Therefore, it is clear that in a more normal situation our proposed LAP-MD(no_Ca) and LAP-FT(no_Ca) outperform the other strategies, by integrating broadcast scheduling with the base station caching, and by making the best use of the three bandwidths of the HWB model. Specifically, these two strategies adopt the following processing key points for joint control.

One point is to keep the items with the highest local access probability in the base station cache, which is more effective than broadcasting them with high broadcast frequency, also better than caching the global hottest items. This processing is suitable for the heterogeneous access of the clients in different areas, and can provide low-latency service to the clients in the same region with a similar access pattern. Results reveal that those strategies keeping the local hottest items in the base station cache normally have better performance.

Another point is to remove the cache-resident items from the broadcast program. This processing can greatly shorten the broadcast cycle, increase the available bandwidth of the main channel, and also can make effective use of the sub channel and wireless channel to pull the items disappeared in the main channel. The experiments confirm that the strategies adopting this processing perform better than those corresponding strategies which broadcast the whole items.

On the other hand, if there is no limit to use these two processing, when the increasing cache hits is beyond the bandwidth of wireless channel, and when most of the hottest items are removed from the main channel, system performance will become worse, due to the heavy load of the wireless channel and the low bandwidth utilization of the main channel.

4 Related Works

Recent advances in computer and wireless communication technologies have increased interest on combination of data broadcasting and data caching. Many researches investigated the collaborative control methods by integrating broadcast scheduling and cache management[1, 4, 8].

Broadcast Disks (BD) was well known joint control method by constructing multi-disks scheduling at the server side, and integrating with the cost-based PIX caching at the client side[1]. While, our proposed strategies adopt multi-disks scheduling, we integrate it with the cache management of the base station.

Moreover, Su et al. proposed an integrated method to simultaneously produce the broadcast schedule and the scheme of the prefetch management of the client [8]. Although the heterogeneous access of the clients was taken into account to construct the broadcast scheduling and client caching, it did not consider utilizing data caching to provide low-latency service to the clients with a similar access pattern.

Additionally, Ercetion et al. addressed the joint cache management and scheduling problem in two stage satellite-terrestrial wireless broadcast system [4]. In such a system, more efficient joint control methods were developed by taking account of lower average latency for the local clients, of which the motivation is similar with ours. However, there is no uplink between the clients and local stations or the local stations and the main server. They only suggest the further work needs to be extended to consider the hybrid system.

In these studies, it has been confirmed that the joint control methods by integrating broadcast scheduling and data caching can provide more efficient data delivery. However, their works are based on a push-based broadcasting, and mostly conduct cache management at the client side. To our best knowledge, none of studies discuss an integrated control strategy based on the hybrid broadcasting environments, which integrate base station caching and broadcast scheduling.

On the other hand, our previous study proposed the HWB model and confirmed the HWB approach can enhance system performance[3]. However it solely determined to adopt flat broadcasting and LRU caching, which did not investigate whether the cooperation of broadcast scheduling and base station caching in the HWB environment can further improve the system performance.

5 Conclusion

In this paper, we investigated a joint control of the HWB system by integrating broadcast scheduling and cache management of the base station, and designed several joint strategies to evaluate the efficient integrated control. The simulation study reveals that the system performance can be further improved by the suitable joint control. It is also confirmed that our proposed control strategies normally outperform the traditional strategies, by taking advantage of the complementary features of the HWB data dissemination and by taking advantage of the special role of the base station cache and its cooperation with the broadcasting.

The current work is conducted on the stable workload of the system. However, system workload is dynamically changed in a real environment. Hence, dynamic control of the HWB system is an interesting problem for further study. The work in this paper is helpful to design an appropriate control mechanism in a dynamic system; we will investigate a suitable way to take advantage of the proposed joint control.

Acknowledgments

This research was supported in part by Grant-in-Aid for Scientific Research (B) (15300033) and (16016260) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and as part of the 21st Century Center of Excellence Program of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] S. Acharya, R. Alonso, M. Franklin, S. Zdonik, Broadcast disks: Data management for asymmetric communication environments, Proc. ACM SIGMOD'95, pp.199-210, May 1995.

[2] S. Acharya, M. Franklin , S. Zdonik, Balancing push and pull for data broadcast, Proc. ACM SIGMOD'97, pp.183-194, May 1997.

[3] J. Cai, T. Terada, T. Hara, and S. Nishio, A query processing method for hybrid wireless-broadcast networks, Proc. MDDS'05, pp.1078-1082, Aug. 2005.

[4] O. Ercetin, L. Tassiulas, Push-based information delivery in two stage satellite-terrestrial wireless systems, IEEE Tran. on Computers, Vol.50, No.5, pp.506-518, 2001.

[5] Q.L. Hu , D.L. Lee , W.-C. Lee, Performance evaluation of a wireless hierarchical data dissemination system, Proc. MobiCom'99, pp.163-173, Aug. 1999.

[6] C.W. Lin, H.Hu, D.L. Lee, Adaptive realtime bandwidth allocation for wireless data delivery, Wireless Networks 10, pp.103-120, 2004.

[7] K. Stathatos, N. Roussopoulos, J.S. Baras, Adaptive Data Broadcast in Hybrid Networks, Proc. VLDB'97, pp.326-335, Aug. 1997.

[8] C.J. Su, L. Tassiulas, Joint broadcast scheduling and user's cache management for efficient information delivery, Proc. MobiCom'98, pp.33-42, Oct. 1998.