International Journal of Computer Applications Technology and Research
Volume 4– Issue 5, 336 - 339, 2015, ISSN:- 2319–8656

Heuristic Algorithm for Efficient Data Retrieval Scheduling in the Multichannel Wireless Broadcast Environments

A. Porselvi
Dept. of CSE
Panimalar Institute of Technology
Chennai,India

S.Brindha Devi
Dept. of CSE
Panimalar Institute of Technology
Chennai,India

Abstract: Wireless data broadcast is an efficient way of disseminating data to users in the mobile computing environments. From the server's point of view, how to place the data items on channels is a crucial issue, with the objective of minimizing the average access time and tuning time. Similarly, how to schedule the data retrieval process for a given request at the client side such that all the requested items can be downloaded in a short time is also an important problem. In this paper, we investigate the multi-item data retrieval scheduling in the push-based multichannel broadcast environments. The most important issues in mobile computing are energy efficiency and query response efficiency. However, in data broadcast the objectives of reducing access latency and energy cost can be contradictive to each other. Consequently, we define a new problem named Minimum Cost Data Retrieval Problem (MCDR) and Large Number Data Retrieval (LNDR) Problem. We also develop a heuristic algorithm to download a large number of items efficiently. When there is no replicated item in a broadcast cycle, we show that an optimal retrieval schedule can be obtained in polynomial time.
Keywords – Multichannel, Wireless data broadcast, MCDR, LNDR

1. INTRODUCTION

BROADCAST is a means by which a single server can transmit data to an unlimited number of clients in a scalable way [3], [4]. Unlike unicast transmission, broadcast is scalable because a single transmission of an item satisfies all outstanding requests for it Generally, there are two types of broadcast systems: push-based and pull-based.

In a push-based system, the server will broadcast a set of data items to the clients periodically according to a fixed schedule; while in a pull-based system, the clients will first send requests to the server and the server will provide timely broadcast according to the requests received. Response time is the time interval between the moment a client tunes in a broadcast system with a request of one or more data items to the moment all requested data are downloaded. It is obvious that shorter response time is more desirable. On the other hand, in wireless communication environments, most clients are mobile devices operating on batteries. The smaller the amount of energy consumed during retrieving data is, the longer the battery life of a mobile device will be. Therefore, saving energy is another important issue for designing wireless data broadcast system. The fast development of wireless communication technologies such as OFDM (Orthogonal frequency division multiplexing) makes efficiently broadcasting data through multiple channels possible [25]. How to allocate the data onto multiple channels to minimize the expected response time has become a hot research topic and lots of scheduling algorithms are proposed [11], [19], [21]. When a query requests only one data item, to schedule the retrieving process is straightforward. However, it is common that a query requests multiple data items at a time [9], [15], [18] (e.g., a user may submit a query of the top 10 stocks). In such cases, different retrieving schedules may result in different response time. Moreover, in a multi-channel broadcast system, retrieving data will probably need switchings among the channels, which not only consumes additional energy, but also causes possible conflicts [17], [22], [26]. The LNDR problem takes the "deadline" into consideration and therefore also describes the time-critical scenario. For push-based broadcast, we derive a polynomial time $(1-\frac{1}{e}-\epsilon)$-approximation scheme for LNDR, and we also propose a heuristic algorithm for it based on maximum independent set. For the case that all channels are synchronized, we propose a polynomial time optimal algorithm for LNDR. When channels are unsynchronized, we prove LNDR is NP-hard. When all the requested data items have to be downloaded, we formulate another problem, namely minimum cost data retrieval (MCDR), with the objective of minimizing the response time and energy consumption .We investigate the approximability of MCDR in push-based broadcast. Due to the strong in-approximability, we develop a heuristic algorithm for MCDR.

2. RELATED WORKS

Scheduling is an important issue in the area of wireless data broadcast. Acharya et al. first proposed the scheduling problem for data broadcast [1], and Prabhakara et al. suggested the multi-channel model for data broadcast to improve the data delivery performance [14]. Since then, many works have been done for scheduling data on multiple channels to reduce the expected access time [20,22,2]. Besides, some researches began to study how to allocate dependent data on broadcast channels (see, e.g., [10,19,21,5,6]). With respect to index, many methods have been proposed to improve the search efficiency in data broadcast systems (see, e.g., [8,16,18,19,21]).

Jung et al. proposed a tree-structured index algorithm that allocates indices and data on different channels [11]. Lo and Chen designed a parameterized schema for allocating indices and data optimally on multiple channels such that the average expected access latency is minimized [12]. In terms of data retrieval scheduling, Hurson et al. proposed two heuristic algorithms for downloading multiple data items from multiple channels [7]. As both push-based and pull-based approaches have their own strengths and drawbacks [15,16], hybrid scheduling is regarded as a prospective approach to better scheduling.

N. Saxena et al. [17] proposed a probabilistic hybrid scheduling, which probabilistically selects push operation or pull operation based on the present system statistics. Their results show that hybrid scheduling generally outperforms other purely push-based or pull-based algorithms in terms of access time. However, the above are all non-real-time scheduling. Huang and Chen proposed a scheme based on a generic algorithm to handle a similar problem [5].

3. PROPOSED WORK

In graph theory, an independent set or stable set for a graph G is a subset of vertices that are pairwise non-adjacent. A maximum independent set is an independent set with the maximum cardinality. As we mentioned in Section 2, a valid retrieval schedule for an LNDR instance is a set of triples without conflicts. Thus, finding a valid schedule with the largest number of requested data items is equivalent to finding a maximum independent set, considering conflicts as edges and triples as vertices. Although finding a maximum independent set is NP-hard, we still can devise heuristics that provide solutions not necessarily provable, but usually efficient for practice. We next present a sequential greedy heuristic that guarantees a maximal valid retrieval schedule (i.e., a valid set of triples that is not a subset of others).

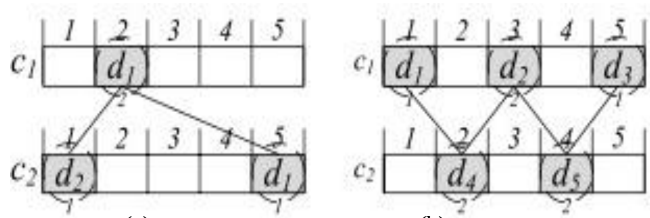
Heuristic Algorithm:

1. **Input**: an LNDR instance which is represented by a set of triples.
2. Construct a graph G of triples and add edges between conflicted triples;
3. Let P<-Ø (P denotes the set of triples selected);
4. **While** G is not empty **do**

select a triple in G with the minimum degree;
put it in P and delete its neighbors;

5.**end while**
6.output P;

Generally, when a subset of elements need to be selected, a greedy based algorithm will construct a solution by adding elements sequentially. Decisions on which element is to be added is based on certain rule. In SGH each time we add a triple with the minimum degree. It can be shown that choosing a vertex and removing its neighbors repeatedly will achieve a maximal independent set. Thus, the solution resulted by SGH is maximal. Moreover, based on our observation, SGH is very efficient in practice, e.g., in Fig. 2a, data item d1 appears twice and SGH will select the one at time 5, because of its relatively low degree. As a result, data item d2 can also be downloaded (the number below a data item indicates its vertex degree). In Fig. 2b, SGH will select data items in channel c1. As a result, three data items can be downloaded. If selecting data items in channel c2, at most two data items can be downloaded. We will demonstrate the efficiency of SGH through simulation in Section 6. Since we convert LNDR into MIS only based on the conflicts, it is clear that SGH can be applied for non-uniform size data items and non-uniform bandwidth channels.



(a) (b)

Fig 2: Two examples

MCDR Greedy Heuristic:

1. Input: a broadcast schedule with requested data item $d_1, d_2, \ldots, d_k$ and two parameters *p* and *q* (*p<q)*.
2. Let $P \leftarrow \emptyset$;
3. construct a set T_{d_i} for each data item d_i;
4. while $|P| < k$ do

5.let $\tau \leftarrow \max_{1 \le i \le k}\left(t_{Tr_f}\left(T_{d_i}\right)\right)$;

6.if there exist a channel c and a time interval [x,y] such that $|c[x,y]| > p$, $y-x \le q$ and $y \le \tau$ then

7.Put that triples in c[x,y] into *P* and delete the conflicted triples;

8.else

let *Tr* be the triple with the maximum e(*Tr*);

9.put *Tr* into P, and delete the conflicted triples;
10.end if
11. end while
12. output P;

1) Let [x,y] be a time interval and c be a channel, define c[x,y] to be the set of data items in the time interval [x, y] of channel c.
2) For each triple Tr= (d_{Tr}; c_{Tr}; t_{Tr}), define e(*Tr*) to be the earliest time that data item d_{Tr} is downloadable if we do not download T_r at time t_{Tr}.
3) For each requested data item d, define T d to be the set of triples of d.
4) Let *T* be a set of triples, define $Tr_f(T)$ and $Tr_e(T)$, respectively, to be the first and last triples in *T* according to the broadcasting time.
In MGH (Algorithm 5), P holds the triples selected and t is the earliest possible time that all the requested data items can be downloaded. Each time MGH searches for a channel broadcasting a significant number of data items during a short time interval before t. If there exists such a channel, it downloads those data items; otherwise, it selects a triple Tr greedily with the maximum e(*Tr*). The two parameters p and q would be chosen according to α, λ_{Active}, λ_{Doze} and λ_{Switch}. When α=0 and λ_{Doze}=0, we can ignore the response time and set q to be greater than the cycle length, which converts the MCDR problem into a set cover problem, and thus brings an O(log k)-factor approximation solution. When α=1, we can decrease q and increase p to minimize the response time, regardless of the energy consumption.

4. CONCLUSION

In this paper, the data retrieval scheduling for multi-item requests over multiple channels is studied. Two optimization problems, LNDR and MCDR, are defined and some approximation and heuristic algorithms are proposed. The algorithms are analyzed both theoretically and practically. Their efficiencies are also demonstrated through simulation. For LNDR in push-based broadcast, MM can download the maximum number of data items when the channels are synchronized. When the channels are unsynchronized, SGH always achieves a better solution with respect to GL, NO, MM and RS, and it scales well. AS is slightly better than SGH but it cannot be applied to download a large number of data items. For LNDR in pull-based broadcast, GL is better than NO, and other algorithms cannot be applied. For MCDR, MGH always outperforms MH, GL, NO and RS.

RS is also an efficient scheduling when a large percentage of data items have to be downloaded. To the best of our knowledge, we do not find any algorithms in the literature which are designed for pull-based data scheduling at the server side over multiple unsynchronized channels. As a direction for further research, one can study the data scheduling problem for unsynchronized channels from the server's point of view.

6. REFERENCES

[1] J.E. Hopcroft and R.M. Karp, "An n5=2 Algorithm for Maximum Matchings in Bipartite Graphs," SIAM J. Computing, vol. 2, no. 4, pp. 225-231, 1973.
[2] H.D. Dykeman, M. Ammar, and J.W. Wong, "Scheduling Algorithms for Videotex Systems under Broadcast Delivery," Proc. IEEE Int'l Conf. Comm., pp. 1847-1851, 1986.
[3] S. Acharya, R. Alonso, M. Franklin, and S. Zdonik, "Broadcast
Disks: Data Management for Asymmetric Communication Environments," Proc. ACM SIGMOD Int'l Conf. Management of Data, pp. 199-210, 1995.
[4] N. Vaidya and S. Hameed, "Log Time Algorithms for Scheduling Single and Multiple Channel Data Broadcast," Proc. Ann. Int'l Conf. Mobile Computing and Networking, pp. 90-99, 1997.
[5] U. Feige, "A Threshold of lnn for Approximating Set Cover," J. ACM, vol. 45, no. 4, pp. 314-318, 1998.
[6] T. Imielinski, S. Viswanathan, and B.R. Badrinath, "Data on Air: Organization and Access," IEEE Trans. Knowledge and Data Eng., vol. 9, no. 3, pp. 353-372, May/June 1997.
[7] D. Aksoy and M. Franklin, "Scheduling for Large-Scale On-Demand Data Broadcasting," Proc. IEEE Int'l Conf. Computer Comm., pp. 651-659, 1998.
[8] D. Aksoy and M. Franklin, "R _W: A Scheduling Approach for Large-Scale On-Demand Data Broadcasting," IEEE/ACM Trans. Networking, vol. 7, no. 6, pp. 846-860, Dec. 1999.
[9] C. Kenyon and N. Schabanel, "The Data Broadcast Problem with Non-Uniform Transmission Time," Proc. ACM-SIAM Symp. Discrete Algorithms, pp. 547-556, 1999.
[10] C.D. Manning and H. Schutze, Foundations of Statistical Natural Language Processing. MIT Press, 1999.
[11] K. Prabhakara, K.A. Hua, and J. Oh, "Multi-Level Multi-Channel Air Cache Designs for Broadcasting in a Mobile Environment," Proc. IEEE Int'l Conf. Data Eng., pp. 167-176, 2000.
[12] W. Mao, "Competitive Analysis of On-line Algorithms for On- Demand Data Broadcast Scheduling," Proc. Int'l Symp. Parallel Architectures, Algorithms and Networks, pp. 292-296, 2000.
[13] Y.D. Chung and M.H. Kim, "Effective Data Placement for Wireless Broadcast," Distributed and Parallel Databases, vol. 9, no. 2, pp. 133-150, 2001.
[14] G. Lee, M.S. Yeh, S.C. Lo, and A. Chen, "A Strategy for Efficient Access of Multiple Data Items in Mobile Environments," Proc. IEEE Int'l Conf. Mobile Data Management, pp. 71-78, 2002.
[15] W.G.Yee , S.B. Navathe, E. Omiecinski, and C. Jermaine, "Efficient Data Allocation over Multiple Channels at Broadcast Servers," IEEE Trans. Computers, vol. 51, no. 10, pp. 1231-1236, Oct. 2002.
[16] W.G. Yee and S.B. Navathe, "Efficient Data Access to Multi- Channel Broadcast Programs," Proc. ACM Int'l Conf. Information and Knowledge Management, pp. 153-160, 2003.
[17] J.L. Huang, M.S. Chen, and W.C. Peng, "Broadcasting Dependent
Data for Ordered Queries without Replication in a Multi-Channel Mobile Environment," Proc. IEEE Int'l Conf. Data Eng., pp. 692- 694, 2003.
[18] M.V. Lawrence, L.S. Brakmo, and W.R. Hamburgen, "Energy Management on Handheld Devices," ACM Queue, vol. 1, pp. 44- 52, 2003.

[19] J.L. Huang and M.S. Chen, "Broadcast Program Generation for Unordered Queries with Data Replication," Proc. ACM Symp. Applied Computing, pp. 866-870, 2003.
[20] A.A. Ageev and M.I. Sviridenko, "Pipage Rounding: A New Method of Constructing Algorithms with Proven Performance Guarantee," J. Combinatorial Optimization, vol. 8, no. 3, pp. 307- 328, 2004.
[21] K. Foltz, L. Xu, and J. Bruck, "Scheduling for Efficient Data Broadcast over Two Channels," Proc. IEEE Int'l Symp. Information Theory,
pp. 113-116, 2004.
[22] J. Juran, A.R. Hurson, N. Vijaykrishnan, and S. Kim, "Data Organization and Retrieval on Parallel Air Channels: Performance and Energy Issues," Wireless Networks, vol. 10, no. 2, pp. 183-195, 2004.
[23] J.L. Huang and M.S. Chen, "Dependent Data Broadcasting for Unordered Queries in a Multiple Channel Mobile Environment," IEEE Trans. Knowledge and Data Eng., vol. 16, no. 9, pp. 1143-1156, Sept. 2004.
[24] E. Ardizzoni, A.A. Bertossi, S. Ramaprasad, R. Rizzi, and M.V.S. Shashanka, "Optimal Skewed Data Allocation on Multiple Channels with Flat Broadcast per Channel," IEEE Trans. Computers, vol. 54, no. 5, pp. 558-572, 2005.
[25] S. Jung, B. Lee, and S. Pramanik, "A Tree-Structured Index Allocation Method with Replication over Multiple Broadcast Channels in Wireless Environment," IEEE Trans. Knowledge and Data Eng., vol. 17, no. 3, pp. 311-325, Mar. 2005.
[26] B. Zheng, X. Wu, X. Jin, and D.L. Lee, "Tosa: A Near-Optimal Scheduling Algorithm for Multi-Channel Data Broadcast," Proc. IEEE Int'l Conf. Mobile Data Management, pp. 29-37, 2005.
[27] A.R. Hurson, A.M. Munoz-Avila, N. Orchowski, B. Shirazi, and Y. Jiao, "Power Aware Data Retrieval Protocols for Indexed Broadcast Parallel Channels," Pervasive and Mobile Computing, vol. 2, no. 1, pp. 85-107, 2006.
[28] Y. Yao, X. Tang, E.P. Lim, and A. Sun, "An Energy-Efficient and Access Latency Optimized Indexing Scheme for Wireless Data Broadcast," IEEE Trans. Knowledge and Data Eng., vol. 18, no. 8, pp. 1111-1124, Aug. 2006.
[29] J. Xu, W.C. Lee, X. Tang, Q. Gao, and S. Li, "An Error-Resilient and Tunable Distributed Indexing Scheme for Wireless Data Broadcast," IEEE Trans. Knowledge and Data Eng., vol. 18, no. 3, pp. 392-404, Mar. 2006.
[30] T. Jiang, W. Xiang, H.H. Chen, and Q. Ni, "Multicast Broadcast Services Support in OFDMA-Based WiMAX Systems," IEEE Comm. Magazine, vol. 45, no. 8, pp. 78-86, Aug. 2007.
[31] J. Chen, G. Huang, and V.C.S. Lee, "Scheduling Algorithm for Multi-Item Requests with Time Constraints in Mobile Computing Environments," Proc. Int'l Conf. Parallel and Distributed Systems, pp. 1-7, 2007.
[32] K. Liu and V.C.S. Lee, "On-demand Broadcast for Multi-Item Requests in a Multiple Channel Mobile Environment," Information Sciences, vol. 180, no. 22, pp. 4336-4352, 2010.
[33] Y. Shi, X. Gao, J. Zhong, and W. Wu, "Efficient Parallel Data Retrieval Protocols with MIMO Antennae for Data Broadcast in 4G Wireless Communications," Proc. Int'l Conf. Database and Expert Systems Applications, pp. 80-95, 2010.
[34] X. Gao, Z. Lu, W. Wu, and B. Fu, "Algebraic Algorithm for Scheduling Data Retrieval in Multi-channel Wireless Data Broadcast Environments," Proc. Int'l Conf. Combinatorial Optimization and Applications, pp. 74-81, 2011.
[35] J. Lv, V.C.S. Lee, M. Li, and E. Chen, "Profit-Based Scheduling and Channel Allocation for Multi-Item Requests in Real-Time On- Demand Data Broadcast Systems," Data & Knowledge Eng., vol. 73, pp. 23-42, 2012.
[36] Z. Lu, W. Wu, and B. Fu, "Optimal Data Retrieval Scheduling in the Multi-Channel Wireless Broadcast Environments," IEEE Trans. Computers, vol. 62, no. 12, pp. 2427-2439, Dec. 2013.