Computers and Mathematics with Applications 64 (2012) 1242–1251

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

MLAIN: Multi-leveled air indexing scheme in non-flat wireless data broadcast for efficient window query processing

Seokjin Im, JinTak Choi *

College of Information and Technology, University of Incheon, Incheon, Republic of Korea

ARTICLE INFO

Keywords:
Non-flat wireless data broadcast
Distributed indexing scheme
Spatial data
Window query

ABSTRACT

In ubiquitous computing, it is critical to allow a great number of clients to access information simultaneously at any place and at any time. Wireless data broadcasting provides effective information services due to its own high scalability. In this paper, we propose a novel indexing scheme for spatial data items that adopts multi-leveled grid partition to support window queries in non-flat data broadcasting that considers clients' skewed data access patterns. In the proposed scheme, each cell of the partition is restricted to its number of data items. This constraint makes the proposed scheme different from other schemes that use a space partition. Cell indexes of cells that keep link information between cells are interleaved with data items on the channel. The scheme allows clients to access queried data items quickly by reducing the broadcast cycle and the spacing between indexes, and providing multiple paths to a cell on the channel. We show the efficiency of the proposed scheme with regard to the access time, tuning time, and energy consumption using intensive simulation studies.

1. Introduction

Ubiquitous computing has evolved tremendously with technological advances in wireless networks and context aware systems and is becoming more and more pervasive in daily life [1–4]. In the environment, numerous mobile clients try to access a variety of information at any place and at any time while moving freely. It is critical to provide clients with information services efficiently regardless of the number of clients [5–11].

A wireless data broadcasting system that enables a server to disseminate data items periodically over a wireless channel provides effective information services, in addition to offering high scalability because it can accommodate an arbitrary number of mobile clients simultaneously [5]. In such a system, each client accesses the channel and downloads the desired data items. Microsoft's MSN Direct Services, which provide information to clients, runs using wireless data broadcasting [12].

In particular, the system broadcasting spatial data items, which have geographic location information regarding the features of a particular area, such as data on hotels in a city, provides numerous clients with efficient location dependent information services. Using the system, each client independently processes its own spatial query, such as a window query to find data items within a given query window or a kNN query to find k nearest neighbors from a given query point, by monitoring the wireless channel. For example, clients can find hotels within a 100 m square centered at their current location in a metropolitan area such as New York City.

Due to the limited battery life of clients, they operate in two modes: active mode (the energy-consuming mode that listens to the channel to download data items) and doze mode (the energy-saving mode). In order to download data items satisfying

* Corresponding author.
E-mail address: choi@incheon.ac.kr (J. Choi).

0898-1221/$ – see front matter
doi:10.1016/j.camwa.2012.03.068

a given query from the wireless channel, clients must listen to the channel until the items appear on the channel. This causes clients to consume a great deal of energy during data retrieval because of the long stay in active mode that is required.

Air indexing schemes have been proposed for energy efficient query processing by letting clients selectively listen to only the data items queried. An air index for spatial data items stores the location information about the items and their broadcasting times. The index is then interleaved with the items on the channel. After accessing the index, clients filter out the queried items and selectively download them with their broadcasting times.

The performance of the system is characterized by access time and tuning time [5]. Access time refers to the time elapsed from the beginning of processing a given query to the receipt of the answer to the query. Tuning time is the amount time necessary in active mode for the access time.

According to the uniformity of the broadcast frequency of data items, the wireless data broadcasting systems can be categorized into flat data and non-flat data. In flat data broadcasting, the server disseminates all data items with the same frequency in a broadcast cycle [5]. On the other hand, in non-flat data broadcasting, the server disseminates popular data items, called hot data items, more frequently than regular ones in a broadcast cycle [13]. When clients' data access patterns are skewed to hot data items, non-flat data broadcasting reduces the average access time because clients can receive hot data items more quickly than regular ones on the channel.

Various data scheduling and air indexing schemes for spatial data have been developed for efficient spatial query processing. In flat data broadcasting, to support spatial queries, a Hilbert curve index (HCI) and distributed spatial index (DSI) were proposed on a Hilbert curve (HC) in [14,6], respectively. The two indexing schemes cause clients to spend an excessive amount of time accessing queried data items; these two indexing schemes extract data items by listening to many candidates that were decided with HC values for each item rather than their real coordinates. In [7], a cell-based distributed index (CEDI) using a regular cell partition was proposed for time efficient query processing. With CEDI, clients listen to only their desired data items by extracting them using their real coordinates. Access time is longer in CEDI when data items are skewed in several cells because the difference between the distributed indexes increases. The three works on spatial data can cause the deterioration of client performance in skewed data access patterns to hot items because of flat data broadcasting. In non-flat data broadcasting, a grid-based distributed index for non-flat broadcast (GDIN) organized over a regular space partition was proposed to reduce the access time when clients' data access pattern's skewed to hot data [8]. GDIN repeats hot cells, which are cells containing one or more hot data items as well as regular ones, several times in a broadcast cycle. GDIN rapidly increases the number of data items repeated in a cycle when numerous regular data items are skewed around hot data items in a hot cell because GDIN repeats all regular items in a hot cell at the same frequently as with the hot items. This lengthens GDIN's access time because the length of a broadcast cycle gets longer and the differences between indexes on the broadcast channel also increases.

In this paper, we focus on non-flat data broadcasting with spatial data items, proposing a multi-leveled air indexing scheme in non-flat data broadcasting (MLAIN) to efficiently process window queries with the popularity of spatial data items considered. MLAIN partitions the data space by recursively subdividing it into four quadrant cells of multiple levels until all cells satisfy the constraint that the number of data items in a cell is not over a specified number. It is this constraint that makes MLAIN different from other indexing schemes using a regular space partition. Each cell having one or more hot data items, called a hot cell, is broadcast more frequently than regular cells to help clients quickly access hot data items on the channel. A cell index for each cell satisfying the constraint is interleaved with data items in the cell on the wireless channel in the distributed manner. Using the indexes on the channel, clients can process their own queries and download desired items from the channel.

MLAIN contributes to improved access time in the following ways:

- Reducing the number of regular items repeated with hot data items through the use of the constraint. This shortens the length of a broadcast cycle. Therefore, clients can quickly access both hot data items and regular ones because the average access time is half the cycle under the assumption that the probability distribution of the access to the channel of clients is uniform within a cycle.
- There is shortened spacing between indexes because of the constraint, although lots of regular data items are skewed around hot items, unlike a regular space partition. This enables clients to access index information quickly and to start processing a given query immediately.
- There are multiple access paths to a cell on the channel. A MLAIN cell index keeps link information that clients can access in the upper level, same level, and lower level. The link information provides clients with multiple paths to a cell. The paths enable clients to improve the access time because they can take another path to access a cell although they miss a path to the cell, rather than waiting for the next cycle.

For the tuning time, MLAIN allows clients to not spend an excessive amount of time in active mode accessing data items because it lets them extract their desired data items using the real coordinates of spatial data items before downloading data items from the wireless channel. In addition, MLAIN employs a linear table structure to match clients' linear channel access patterns.

The rest of the paper is organized as follows: Section 2 reviews the related works in wireless data broadcasting. We present the proposed MLAIN in Section 3. Then, we describe simulation environments and evaluate the performance of the proposed scheme in Section 4. Finally, we conclude the paper in Section 5.