Data Dissemination in Mobile Databases

Agustinus Borgy Waluyo
Monash University, Australia

Bala Srinivasan
Monash University, Australia

David Taniar
Monash University, Australia

INTRODUCTION

The development of wireless technology has led to *mobile computing*, a new era in data communication and processing (Barbara, 1999; Myers & Beigl, 2003). With this technology, people can now access information anytime and anywhere using a portable, wireless computer powered by battery (e.g., PDAs). These portable computers communicate with a central stationary server via a wireless channel. Mobile computing provides *database applications* with useful aspects of wireless technology known as mobile databases.

The main properties of mobile computing include mobility, severe power and storage restriction, frequency of disconnection that is much greater than a traditional network, bandwidth capacity, and asymmetric communications costs. Radio wireless transmission usually requires a greater amount of power as compared with the reception operation (Xu, Zheng, Zhu, & Lee, 2002). Moreover, the life expectancy of a battery (e.g., nickel-cadmium, lithium ion) was estimated to increase time of effective use by only another 15% (Paulson, 2003). Thus, efficient use of energy is definitely one of the main issues.

Data dissemination (can also be called *data broadcasting*) is one way to overcome these limitations. With this mechanism, a mobile client is able to retrieve information without wasting power to transmit a request to the server. Other characteristics of data dissemination include: scalability as it supports a large number of queries; query performance which is not affected by the number of users in a cell as well as the request rate; and effective to a high-degree of overlap in the user's request. In this article, the terms data dissemination and data broadcasting are used interchangeably.

The ultimate challenge in data dissemination is to minimize *the response time* and *tuning time* of retrieving database items. Response time is the total of elapsed time required for the data of interest to arrive in the channel and the download time, while tuning time is the amount of time that a client is required to listen to the channel, which is used to indicate its energy consumption. In some cases, the response time is equal to the tuning time.

This article describes a state-of-the art development in data dissemination strategies in mobile databases. Several strategies for improving the query performance by disseminating data to a population of mobile users will be explained.

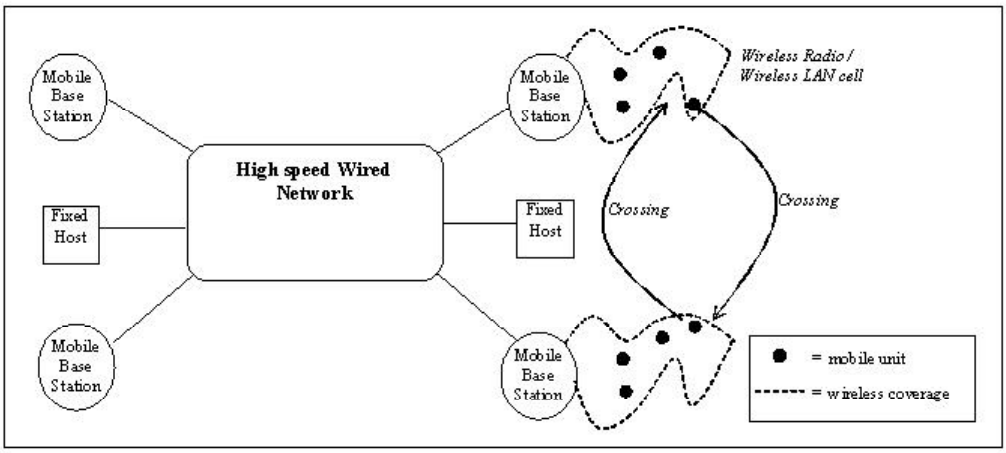
BACKGROUND

In general, each mobile user communicates with a mobile base station (MBS) to carry out any activities such as a transaction and information retrieval. MBS has a wireless interface to establish communication with the mobile client, and it serves a large number of mobile users in a specific region called a "cell". The number of mobile clients in a cell can be infinite. In mobile environment architecture, each MBS is connected to a fixed network as illustrated in Figure 1.

Mobile clients can move between cells while being active and this intercell movement is known as the handoff process (Trivedi, Dharmaraja, & Ma, 2002). Each client in a cell can connect to the fixed network via wireless radio, wireless local area network (LAN), wireless cellular, or satellite. Each of the wireless networks provides a different bandwidth capacity. However, this wireless bandwidth is too small compared with the fixed network such as asynchronous transfer mode (ATM) that can provide a speed of up to 155Mbps (Elmasri & Navathe, 2003).

Data dissemination refers to the periodic broadcasting of database items to mobile clients through one or more wireless channels (or also called broadcast channels), and the clients filter their desired data on the fly. Access to data is sequential. The behavior of the broadcast channel is unidirectional which means the server disseminates a set of data periodically to a multiple number of users. This mechanism is also known as the *push-mechanism* (Malladi & Davis, 2002; Yajima, Hara, Tsukamoto, & Nishio, 2001). It must be noted that data dissemination is different from the data replication mechanism. Conventional data replication distributes a set of database items to one or more identified clients according to a pre-determined requirement. However, data dissemination broadcasts the database items periodically to an unbounded

Figure 1. Mobile environment architecture



number of mobile clients, and the clients filter the data on air based on individual interest.

Figure 2 shows the mechanism of data dissemination. In this article, the term data item corresponds to database record or tuples, and data segment contains a set of data items. A complete broadcast file is referred to as a broadcast cycle. The terms mobile client, mobile computer, mobile unit, mobile user and client are used interchangeably.

DATA DISSEMINATION

Data dissemination schemes are classified into two categories: one is to minimize query response time, and the other minimizes tuning time.

Minimizing Query Response Time

There are several data dissemination schemes, which include:

(i) Selection of Data Items to be Broadcast,
(ii) Non-Uniform Frequency Distribution of Broadcast Data Items,
(iii) Distribution of Data Items over Multiple Channels, and
(iv) Organization of Data Items.

These schemes aim to minimize the query response time by either reducing the waiting time for the desired data to arrive, or, both waiting and download time.

Figure 2. Data dissemination mechanism