Chapter 6
Statistical Inference-Based Cache Management for Mobile Learning

Qing Li
Zhejiang Normal University, China

Jianmin Zhao
Zhejiang Normal University, China

Xinzhong Zhu
Zhejiang Normal University, China

ABSTRACT
Supporting efficient data access in the mobile learning environment is becoming a hot research problem in recent years, and the problem becomes tougher when the clients are using light-weight mobile devices such as cell phones whose limited storage space prevents the clients from holding a large cache. A practical solution is to store the cache data at some proxies nearby, so that mobile devices can access the data from these proxies instead of data servers in order to reduce the latency time. However, when mobile devices move freely, the cache data may not enhance the overall performance because it may become too far away for the clients to access. In this article, we propose a statistical caching mechanism which makes use of prior knowledge (statistical data) to predict the pattern of user movement and then replicates/migrates the cache objects among different proxies. We propose a statistical inference based heuristic search algorithm to accommodate dynamic mobile data access in the mobile learning environment. Experimental studies show that, with an acceptable complexity, our algorithm can obtain good performance on caching mobile data.

INTRODUCTION
Mobile data management has been an increasingly hot research problem over the past few years, as it is crucial to the successful deployment into a number of applications including mobile learning. An important topic (sub-problem) there is efficient data access in a mobile learning system, especially from the perspective of data caching and quality of services (QoS). Generally speaking,
a local memory space is necessary for a mobile device to cache, in advance, some data objects since the data server on the Internet may need to handle tens of thousands of requests simultaneously. But it becomes unpractical to cache a lot of data on the so-called light-weight mobile devices (particularly, cell phones) due to their very limited capability. Instead, data caching over the proxies becomes a viable approach in order to enable the mobile clients to access the learning objects (data) efficiently. The proxies may serve as a secondary level cache if a mobile device can maintain a small local cache, or serve as a direct cache if the mobile device cannot hold a local cache at all. However, how to dynamically maintain the cache data across the multiple proxies in order to enhance the overall performance of a mobile learning system is a challenging problem.

There are two main difficult problems for maintaining the cache data on the proxies. First of all, a proxy is limited by its hardware capability, which may prevent it from caching all the required data for the mobile clients. When a mobile device needs to access some data from a location whose proxy does not have the data cached, the proxy must decide if it should replicate/migrate the data from other proxies, depending on the actual situations and constraints. Because an earlier decision may impact the subsequent performance, it is very complex and costly to find a global optimal replication/migration algorithm. Secondly, global knowledge on mobile clients’ movement and data access patterns is hard to obtain accurately, if not impossible. Yet even a rough approximation of such global knowledge will surely be useful for making sensible decisions.

In this article, we propose to develop an efficient mobile learning system through devising a statistical caching mechanism, in which each proxy makes use of prior knowledge (statistical data) to decide whether any cache objects should be replicated, migrated or deleted. Our method is different from the traditional approaches, e.g., Markov Chain-based models. Such kinds of models typically have an assumption which may not simulate real-life user movement accurately, and some of them even incur a very high computational cost. In contrast, our mechanism converts the distributed caching problem into a heuristic search problem and uses statistical inference to supervise the decision.

The rest of the article is organized as follows. We give a brief review of related work in the next section, after which our system architecture including the performance metrics is introduced. Next, we discuss the statistical caching mechanism in detail, followed by presenting our experiment results. In the last section, we summarize our work and offer a few research directions.

RELATED WORK

In this section, we review some earlier research works related to our research. Such relevant works can be divided into three categories: mobility model (movement prediction), data migration and data caching.

Mobility Model

Several user mobility models can be found in previous research papers for modelling the user movement. The fluid flow model (Thomas, R., Gilbert, H., & Mazziotto, G., 1998) is one of the early models widely used, in which mobile users are assumed to be uniformly populated and the users carrying terminals are moving at an average velocity with uniformly distributed moving-direction over [0, 2π]. During the last fifteen years, other mobility models have been suggested, such as random way point model, city model, highway model etc. Most of these models assumed that user behaviour is absolutely random. Kobayashi et.al. introduced a new mobility model based on HSMM process and used it to facilitate resource allocation in wireless networks (Kobayashi, H., Yu, S. Z., & Mark, B. L., 2000). To study and