Wireless Networks 9, 239–247, 2003

A Cost-Efficient Scheduling Algorithm of On-Demand Broadcasts

WEIWEI SUN *, WEIBIN SHI and BOLE SHI
Department of Computing and Information Technology, Fudan University, PR China

YIJUN YU
Department of Electrical Engineering, Gent University, Belgium

Abstract. In mobile wireless systems data on air can be accessed by a large number of mobile users. Many of these applications including wireless internets and traffic information systems are pull-based, that is, they respond to on-demand user requests. In this paper, we study the scheduling problems of on-demand broadcast environments. Traditionally, the response time of the requests has been used as a performance measure. In this paper we consider the performance as the average cost of request composed of three kinds of costs – access time cost, tuning time cost, and cost of handling failure request. Our main contribution is a self-adaptive scheduling algorithm named LDFC, which computes the delay cost of data item as the priority of broadcast. It costs less compared with some previous algorithms in this context, and shows good adaptability as well even in pure push-based broadcasts.

Keywords: scheduling algorithm, cost modeling, data broadcast, wireless network, mobile computing

1. Introduction

In a client/server architecture with fixed networks, clients would send requests when they want to retrieve data from the server. Then the server will respond to the requests and send data to clients. Compared with fixed networks, wireless networks have low bandwidth and low communication quality [7,10]. To support numerous mobile users to access data in server concurrently, a new method of data-transmission is put forward, that is, the server broadcasts data on air and clients acquire data that way, which is called *data broadcasting*.

Data broadcast technology has many applications in the field of public information dissemination, such as stock market quotation or traffic and landmark information. One important issue in broadcast technology is to determine an optimal broadcast sequence according to the access probability distribution of mobile users, i.e. the data broadcast scheduling. To evaluate the effectiveness of one broadcast scheduling strategy, we need to consider two basic aspects:

(1) Access Time (shortened as *AT*). It indicates the time elapsed between the query submission and receipt of the response. *AT* determines the response time of query made by mobile users. We need to concentrate on the arrangement of frequency and location of data items in one broadcast period, so as to make the average *AT* smaller, according to various access probabilities of data items. The study on this issue includes [1–3,5,9,14,17], etc.

(2) Tuning Time (shortened as *TT*). It indicates the total time that mobile users spend actively listening to the channel in a complete access period. *TT* determines the power consumption of mobile users because they could slip into doze (stand by) mode when they are not actively listening on the channel. As most of mobile users depend on limited battery supply, the reduction of *TT* would also be an important issue in data broadcast technology. A widespread method is to insert index segments into broadcast period in order to reduce *TT*. The study on this issue includes [11,13,16], etc.

In on-demand broadcasts, we cannot obtain the access profiles of mobile users, that is to say, their access pattern would have some unpredictable changes. Thus, we need a kind of new scheduling algorithm, to determine the contents and organization of data broadcast on the basis of circumstances of recent access and scheduling.

The study of on-demand broadcast scheduling problem includes [4,8], etc. In this broadcast environment, mobile users communicate with the server via wireless channels. These channels include an uplink channel and a downlink channel. Mobile users use this uplink channel to send data access request, and the contents of broadcast will reach the mobile users through downlink channel. First, mobile users make the access request; second, the server considers all pending requests to decide the contents of next broadcast. One core issue is to determine the priorities of data items to be broadcasted, that is, which data items should be broadcasted in the next period. A FCFS (First-Come-First-Served) scheduling algorithm is put forward in [5], which sequences data items according to their requested time. Because of its time sequencing principle, any access request would get responded after waiting for a finite period. There does not exist any case of endless waiting. But it has the deficiency of low average performance, because it considers only the requested time, and does not take into account the difference of access frequency of various data items. MRF (Most-Request-First) scheduling algorithm will broadcast those data items with the

* Corresponding author.
E-mail: wwsun@fudan.edu.cn