Network Cache Model for Wireless Proxy Caching

James Z. Wang   Zhidian Du   Pradip K Srimani

Department of Computer Science
Clemson University, Box 340974
Clemson, SC 29634-0974, USA
+1-864-656-7678

{jzwang, zdu, srimani}@cs.clemson.edu

Abstract

We propose a wireless network cache model to facilitate the cooperative proxy caching in wireless base stations. A self-configured and adaptive virtual proxy graph (VPG) is used to achieve efficient data search, data cache and date replication. Based on demand, the aggregate effect of data caching, searching and replicating actions by individual proxies automatically distributes cached web documents closer to the interested mobile clients. Our performance study demonstrates that the proposed individual-based wireless proxy caching scheme outperforms existing caching schemes in terms of different performance metrics.

1. Introduction

The popularity of wireless networks grows with the advances of wireless technologies and internet applications. New generation wireless networks, such as G4 and WiMAX, are ready to bring sophisticated web applications into wireless. In typical wireless internet architecture as depicted in Figure 1, mobile hosts access the wireless network through base stations, which are inter-connected by access routers to form wireless LANs and, in turn, are connected to the internet through gateway routers.

Among the numerous studies carried out on the enhancement of the wireless internet performance, caching popular web documents at locations close to mobile clients is an effective solution to improving the quality of wireless web applications. Especially, building a cooperative proxy cache system among the distributed wireless base stations is desirable.

A cooperative proxy cache system can be either hierarchical or distributed. The hierarchical approach [1] is not suitable for wireless environments due to the non-hierarchical nature of the distributed wireless base stations. Conversely, the distributed cooperative proxy cache systems [2, 3] employ sophisticated caching and searching schemes, such as centralized or distributed directory lookup, making the wireless network more complex. It is necessary to design a new cache model to take advantage of the cooperation among distributed caching proxies in wireless base stations. In addition, new challenges have to be addressed for cooperative proxy caching in wireless base stations. These challenges include constant movement of mobile devices, stringent constraint of user request latency, and diversities of mobile clients.

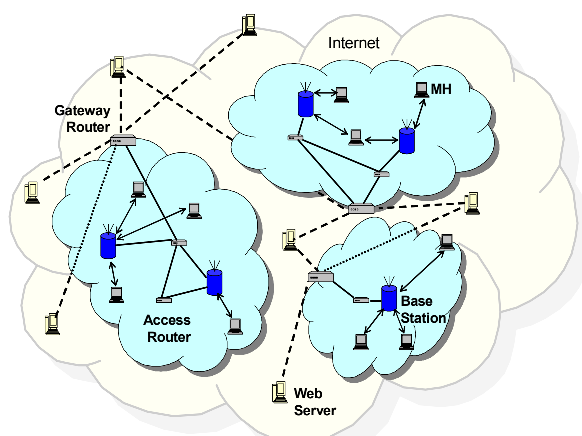

Figure 1. Wireless Internet Architecture

To address these challenges, we propose a new proxy caching scheme for wireless base stations, in which a network cache model is used to facilitate the caching cooperation, and a virtual proxy graph is used to assist the individual-based data cache, data search and data replication. The rest of the paper is organized as follows. In section 2, we propose the individual-based wireless proxy caching scheme. In section 3, we use simulation to demonstrate the advantages of our individual-based wireless proxy caching scheme. Finally we have our concluding remarks and discuss the future studies in section 4.

2. An individual-based wireless proxy caching scheme

To cope with the characteristics of the cached web data and wireless environments, a dedicated cache model is proposed in [4], where each mobile host is assigned a dedicated cache space in the proxy server to assist the cache relocation, which tries to predict the next base station that a mobile host will move to and copies the cached web documents in the dedicated cache space to it before handoff is actually taken place. However this dedicated cache model causes duplication of the same web document in the same proxy cache although it is simple for cache relocation. In addition, inaccuracy of predictions causes unnecessary data copies among the base station proxies and copying web documents based on the prediction of the mobile host movement may introduce data oscillation in wireless network, hindering the efficiency of the cooperative proxy cache system.

To address the problems, we propose an individual-based wireless proxy caching scheme, which can take advantage of the cooperation of base station proxies to increase the cache hit ratio and reduce the web request latency. In this caching scheme, a mobile cache line is used to facilitate the data cache and data replication. Besides the cached web document itself, the head portion of the mobile cache line consists of consists of *ID*, *Tag*, *State bytes, Link fields, Client list and Origin*. The *ID* field contains a *UUID* which the URI of the cached web document will be hashed to. The *Tag* is the name of the cached web document. *State bytes* are used to store information for cache coherency protocol as well as some other statistic data. Link fields (<*NID, Dist*>) are used to provide the links to the neighbor proxies where web replicas can be found. *NID* is an integer used to lookup the neighbor table for a neighbor's IP address. *Dist* is round trip distance to reach the cached web document through the neighbor specified by *NID*. *Client List* contains IDs of the clients that are currently accessing the cached web document. *Origin* indicates whether the current web replica is fetched from the original web server or is migrated from another proxy in the cooperative proxy cache system.

Since our mobile cache lines use links to connect the cached web replicas, it is necessary to avoid one caching proxy having to connect to too many other proxies. To satisfy this constraint, we use a virtual proxy graph (VPG), an overlay proxy network independent of the underlying network structures, to facilitate the data linkage and data exchange among the caching proxies. In our wireless proxy cache system, there is no centralized controller for building the virtual proxy graph. Instead, the VPG is automatically configured by individual base station proxies using some simple rules. The individual caching proxies can also adjust the VPG based on dynamic data flow among them.

An essential part of our cooperative proxy caching scheme for wireless base stations is to design a data search and data cache strategies so that individual proxies can manage the cached data on the proposed VPG using the designed network cache model based on their local knowledge of the global cache state. We expect that the data search, data cache and data replication actions at individual proxies will create a group behavior so that the aggregate effect can manage the global caching state for the wireless proxy cache system. When a new web request reaches a certain proxy server from its mobile client, the request is satisfied if the entire mobile cache line is cached in this proxy. Otherwise, if the head of the mobile cache line is cached in this proxy server, the cached web document can be easily found by following the links established by mobile cache lines. Finally, if nothing is cached in this proxy server, a query message will be created in this proxy server and disseminated to the nearby caching proxies within the VPG to search for the cached web document. If the web document is not found in the wireless proxy cache system, a request is sent directly to the original web server.

Many unique features in the proposed cooperative proxy caching scheme make it an excellent caching scheme for distributed wireless base stations:

Data Cohesion: Because individual proxies on the search path create and update the heads of mobile cache lines to establish links to the cached web replicas, nearby cached web replicas will be linked together to form a cohesive group. Data cohesion allows the nearby proxies to cooperate in making cache replacement decisions.

Search Alignment: By caching and updating the mobile cache line heads in individual proxies during search, later web queries can quickly locate the cached web replicas by following the traces of previous web requests. This search alignment feature can not only reduce the user request latencies but also reduce the network overhead generated by search messages.

Replication-by-demand: When a query locates a cached web document in the wireless proxy cache system, the cached web document will be transferred to the requesting proxy server directly from the proxy server where it is located. Usually the web document will be replicated at the requesting proxy server because of the increasing demand for this web document at the requesting proxy server. The replication-by-demand not only moves the cached web documents to the interested clients but also achieves load balance by distributing highly demanded web documents to distributed mobile base stations.

Furthermore, our individual-based wireless proxy caching scheme is very efficient in handling the mobile host movement. When the mobile host moves from one base station to its neighbor base station, it is not necessary to move its cached web documents to the new base station prior to or at the handoff. Our unique data search scheme can quickly find the cached web documents from the proxy to which the mobile host previously connected if the cached web documents have not been removed by cache replacement. The request latencies of querying web documents from neighbor proxies are much smaller than that of fetching them from the original web servers. Thus, the user request latency is reduced.

3. Performance study

To quantitatively evaluate the performance of our individual-based cooperative caching (ICC) scheme for wireless base stations in terms of cache hit ratio, user request latency, and communication overhead, we use simulation to compare its performance with the following 3 different proxy caching schemes:

Cache Relocation (CR): The proxy cache relocation scheme proposed in [4] is implemented with the assumption of 100% accuracy in prediction of the mobile host movement, although it is unlikely for any prediction to be 100% accurate.

Non-cooperative Cache (NC): In this scheme, the caching proxies make cache decisions independently without cooperation.

Multicast-based Cooperative Cache (MCC): In this scheme, caching proxies search for cached web documents through multicast.

We assume that there are 100 base stations forming a 10x10 mesh-like grid in the wireless network. To evaluate the impact of the mobile host movement, we use a variant random waypoint movement pattern [5] to simulate the movement of the mobile hosts. The time for a mobile host to stay with one base station follows the Poisson distribution with an average duration of 180 seconds. We assume the user request latency between a pair of adjacent base stations are 100ms ignoring the other overheads. The average user request latency for a cache miss is 2000ms, ignoring all other overheads. These assumptions are based on the data collected in [6].

We also assume there are 10,000 distinct web documents on the internet and the average web document size to be 60K. 35% of the web documents have size less than 10K. 60% of them have size in the range from 10KB to 100KB. The sizes for the rest web documents are in the range of 100KB to 1MB. The assumption on web document sizes is based on the latest web access statistics of several different web servers [7]. It is assumed in [8] that, on average, each user clicks a web link once every 12.5 seconds. Thus, in our simulation, the inter-arrival time of web requests generated by a mobile host follows the Poisson distribution with an average inter-arrival time of 12.5 seconds. The access frequencies to the web documents follow a Zipf-like distribution. The access frequency for each web document i is determined as follows:

$$f_i = \frac{1}{i^z \cdot \sum_{j=1}^{m} 1/j^z}$$

where m is the number of the web documents in the system, and $0 \le z \le 1$ is the Zipf factor. We use $z = 0.75$ in our simulation. We vary the cache size in each base station proxy server, observing the performance of the wireless proxy cache system under various caching schemes. We start to collect data after the system is in a stable state. Cache hit ratio, average user request latency, and the average data exchange per web request between a pair of proxy servers are used as our performance metrics.

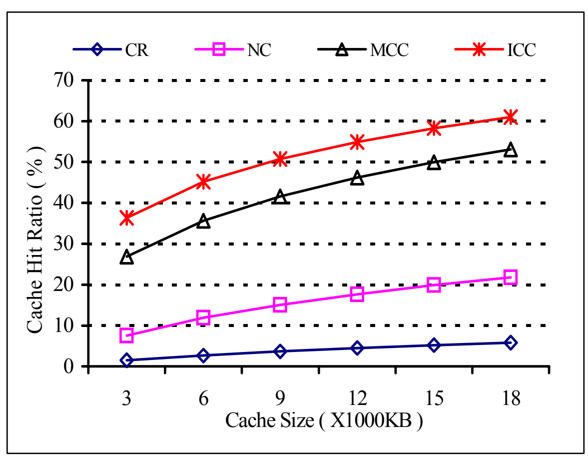

Figure 2. Cache hit ratio vs. cache sizes

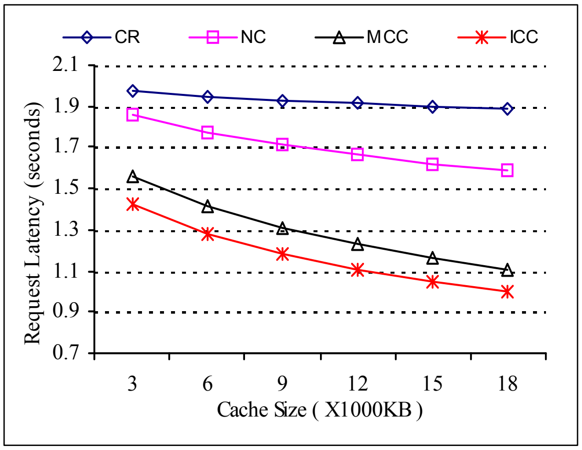

Figure 3. Average request latency under various cache sizes

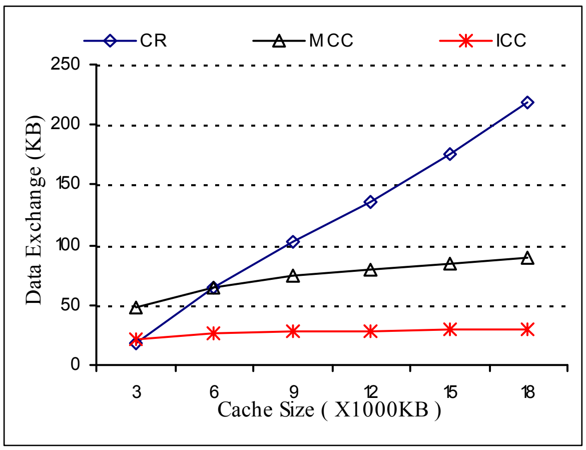

Figure 4. Data exchange vs. cache sizes

The simulation results depicted in Figure 2, Figure 3 and Figure 4 show that our individual-based wireless proxy caching scheme significantly outperforms the non-cooperative single proxy caching approach and the cache relocation approach in terms of cache hit ratio and user request latency. Furthermore, compared to the multicast-based caching scheme, our caching scheme requires much less network bandwidth within the wireless network while achieving better performance in terms of cache hit ratio and average user request latency. The superiority of our proposed caching scheme is due to its novel network cache model and its excellent data cache, data search and data replication strategies. We must note here that NC scheme does not require proxy cooperation. Thus, no data exchange is necessary in this scheme.

4. Conclusion and future studies

In this paper, we propose a novel individual-based cooperative proxy caching scheme for distributed wireless base stations. A network cache model is designed to facilitate the cooperative proxy caching in wireless base stations. Based on demand, the aggregate effect of data caching, searching and replicating actions by individual proxies automatically distributes cached web documents closer to the interested clients. Our performance studies indicate that our proposed individual-based wireless proxy caching scheme is an excellent scheme for web caching in distributed wireless base stations.

Currently we are investigating other factors, such as document size distributions, date replication rate, and mobile host handoff models, on the performance of the cooperative proxy cache system. We are also investigating how to better predict the mobile host movement and integrate with our proposed individual-based wireless proxy caching scheme.

5. Reference

[1] Duane Wessels, K Claffy. ICP and the Squid Web Cache. *IEEE Journal on Selected Areas in Communication*, 16(3):345-357, 1998.

[2] L. Fan, P. Cao, J. Almeida and A. Broder. Summary Cache: A Scalable Wide-Area Web Cache Sharing Protocol. *IEEE/ACM Transactions on Networking*, 8(3):281-293, 2000.

[3] Alex Rousskov and Duane Wessels, Cache Digests. *Computer Networks and ISDN Systems,* 30(22-23):2155-2168, June 1998.

[4] Stathes Hadjiefthymiades and Lazaros Merakos, Using Proxy Cache Relocation to Accelerate Web Browsing in Wireless/Mobile Communications, pages 26-35, *WWW10*, May 1-5, 2001, Hong Kong.

[5] W. Su, S.-J. Lee, and M. Gerla, Mobility Prediction in Wireless Networks, *Proceedings of the IEEE Military Communications Conference (MILCOM)*, Los Angeles, CA, October 2000.

[6] The Measurement Factory, *http://www.measurement-factory.com/*

[7] Web server statistics, *http://www.lescroupiersrunningclub.org.uk/ace/logfile.php*

[8] Microsoft Internet Security & Acceleration server: ISA Server Scales Out to Meet Enterprise-Class Caching Demands, *http://www.microsoft.com/isaserver/evaluation/competitive/scaleout.asp.*