Measurements and modeling of information and client associations in wireless infrastructures

Maria Papadopouli
Department of Computer Science, The University of North Carolina at Chapel Hill

July 23, 2003

1 Project Summary

Wireless devices are becoming smaller, more user-friendly, and more pervasive. They are not only carried by people but are also integrated into physical objects. The effect of Moore's Law is transforming a niche technology into a ubiquitous one, expanding the innovations in an increasingly networked world. The complexity and demand for information and power is growing fast and the power remains a very important challenge for mobile computing. As more people will have access in this networked world, there will be more possibilities for single point failures and service degradation. In addition, mobile devices will experience disconnections from the Internet due to their mobility. The overhead in accessing the Internet via wireless infrastructures can be prohibitive for several real-time multimedia applications on mobile devices. Research on new efficient communication and access paradigms for mobile users becomes important.

My long-term objective is to *develop the principles for designing robust and intelligent wireless infrastructures that improve the performance of the network, enhance the information access, and take into consideration the energy constraints of the devices.*

The proposed research focus on the *development of models of the information, user access, and association patterns using actual traces from the wireless infrastructure*.

2 Modeling information accessed via wireless infrastructures

Characteristics of the information accessed in a wireless environment can be catalysts in the design of caching and prefetching paradigms. We plan to analyze actual traces from different wireless infrastructures and develop models that characterize the information properties, user access, and association patterns. We will use the models to evaluate the feasibility of caching mechanisms and in the location and association prediction mechanisms at the access points (APs).

There are many challenges on how to provide efficient caching and prefetching mechanisms and content distribution networks that enhance the information access of mobile users. To our knowledge, this is the first empirical study that analyzes spatial locality properties of the wireless web access. We believe that spatial locality can have a dominant impact on the mobile information access and this study sets the directions for exploring further such issues.

The high temporal locality of data can enable prefetching systems and replacement algorithms that maximize the data availability over short time periods and minimize the extended storage of data when it has not been recently requested. The spatial locality can have dominant impact on different caching mechanisms. Our research goals are to investigate properties of the information accessed in wireless environments and evaluate the feasibility of caching. In particular, we will

- Measure:
  - the **locality** properties that information accessed via wireless infrastructures exhibits,
  - the **cacheability** and **expiration** of objects,
  - the **locality for collaborative, location-dependent, and peer-to-peer applications**,
  - the **traffic correlation with user location, device type, movement patterns, and activity**.
- Analyze and compare
  - **wireless traffic from different settings (administration, different levels of education, corporation)**,
  - **wireless vs. wired traffic**.

The peer-to-peer caching systems that initially motivated this study, such as 7DS [7], require the objects to be cacheable. Stale objects should not be distributed, but many popular objects on the web are not cacheable by the HTTP standard [4]. It appears that content providers inhibit cacheability to force reloads of their pages for reasons *other* than document freshness (such as distributing new advertisements). This use of the cacheability mechanisms works well enough in fully connected environments, but is a limiting factor for weakly connected systems as we describe here. We intend to measure cacheability of objects in wireless traffic and address this issue; ideally, an object should be cached only for its *true* useful lifetime, while content providers receive the feedback they need.

In this research, we will focus on the wireless web traffic, because the web provides a ready testbed to study the prevalence of locality among wireless devices. We will extend our measurements to learn how different traffic types correlate with buildings, user devices, association pattern, activities according to the university regulations. We will perform this studies in UNC and possibly corporate environments. Actually, it would be *extremely* useful to repeat these studies in IBM and compare their wireless traffic characteristics. We also plan to draw a comparison between wired vs. wireless traffic. We will apply statistical techniques, including feature extraction, multidimensional clustering, and visualization to detect user traffic patterns.

Several papers report that proxies that prefetch and cache data for mobile users or run mobile code and customize the data access based on the user profile can be very beneficial. For the support of such environments, it is crucial to predict the next association of mobile clients. Also, the probability that a client will revisit an AP can provide some guidelines for the expiration of objects related to that client in the AP cache. For that, we will model user association and mobility patterns. In the next Section we discuss our plan on modeling of user associations with APs.

3 Modeling of user associations with APs

By predicting the user associations with APs, an AP can apply a more effective admission control and resource utilization. In the area of PCS and wireless ATM networks there is related research on location management or tracking to locate a given mobile device at any given time. Two strategies are used: location updating and location prediction. Location updating is a passive strategy in which the system periodically records the current location of the mobile in some data bae that it maintains. Tracking efficiency depends on the frequency of such updates. In location prediction, the system dynamically and proactively estimates the mobile's location based on a model. While most recent studies have focused on the update approach, relatively little has been done on the prediction [1, 6, 3]. There is also an effort to simulate mobility modelling and trajectory tracking of mobile users with signal strength measurements using extended Kalman filter-based estimation algorithms [8].

There are only a few user association and movement characterization studies in the wireless infrastructure. Our proposed research will extend the studies by Kotz and Essien [5], Balachandran *et al.* [2], and Tang and Baker [9] by focusing more closely on the association and mobility patterns of individual clients rather than on the entire population of mobile clients and in a finer time granularity. Balachandran *et al.* [2] performed similar measurements in a three-day conference setting, also focusing on the offered network load, and global AP utilization. They characterized wireless users and their workloads and addressed the network capacity planning problem. The study closest to ours is the one done by Kotz and Essien [5] that characterized Dartmouth's wireless network, by examining global traffic and AP utilization. Unlike the Dartmouth wireless infrastructure [5], the UNC wireless clients maintain one single IP address throughout their roaming in the wireless infrastructure and keep the same ID (based on its MAC address) throughout the entire trace. This allows us to correlate the wireless users with their web access and association patterns and carry out user-behavior analysis more accurately.

User association models and accurate predictions of the next association can be very beneficial in several domains such as network simulations, supporting intelligent mobile peers, and designing robust wireless infrastructures. We will explore the following issues:

- develop and evaluate next association prediction algorithms using history information of the client,
- expand the next association prediction by incorporating additional information such as weekday, academic calendar, duration at an AP, different time-scale association patterns for each user, and campus networking and physical topology,
- classify the clients into different categories and develop association models,
- investigate the correlation of information access (such as content type, application) with association models or device type,
- validate the models using not only the UNC wireless traces, but also traces from other universities, institutes, and corporations.

As mentioned earlier, our main goal is to support wireless clients with their quality of service requirements and energy and privacy constraints while efficiently utilizing the wireless bandwidth at each cell. Using predictions of the total and individual client traffic and movement pattern and coordination among neighboring APs, the system can guide clients and more efficiently serve them.

4 Milestones and time plan

4.1 Term 1

- Prepare and set up the testbed
- Implement the scripts that monitor, record, and analyze the traffic
- Collect traces of two weeks, run and analyze the traces
- Start monitoring and collecting traces over an extended period of time of several weeks Main focus is the spatial locality of the http-get requests
- Collect syslogs from APs that show the associations of mobile hosts to APs.

Deliverables: Traces of the http-get requests and the wireless hosts associations to APs

4.2 Term 2

1. Analyze the collected trace, detect trends, and classify the users and traffic with respect to time, AP, location, category of information
2. Model the associations of mobile users to APs using different order markovian chains (we may also try other models)
3. Observe inefficiencies of the 802.11 wireless infrastructure, potential problems, areas of improve-ment
4. Extend the measurements to record other useful information such as packet retransmissions, traffic of other protocols, load at each AP, web page expiration

Deliverables: Measurement analysis of the wireless http traffic with main focus the spatial locality of the information and user access pattern (item 1). Model of the wireless hosts associations to APs (item 2)

4.3 Term 3

1. Continue the classification analysis (of the previous semester)
2. Investigate mechanisms to improve the performance of the wireless 802.11 infrastructure using the observations from the measurement study and the models of users associations to APs and wireless information access
3. Comparison with the measurements in the wired infrastructure

4. Propose architecture changes of the infrastructure, 802.11 protocol changes, deployment of caches, and prefetching algorithms to improve the wireless data access (e.g., improve the response time).

Deliverables: Complete analysis study (item 1 and 3). Proposal of changes for improving the wireless access (item 4).
Publications. We plan to regularly publish our results in conferences, journals, and on the web.

References

[1] Ian F. Akyildiz and Joseph S. M. Ho. On location management for personal communications networks. *IEEE Communications Magazine*, 34(9):138–145, September 1996.

[2] Anand Balachandran, Geoffrey Voelker, Paramvir Bahl, and Venkat Rangan. Characterizing user behavior and network performance in a public wireless lan. In *Proceedings of the ACM Sigmetrics Conference on Measurement and Modeling of Computer Systems*, 2002.

[3] Amiya Bhattacharya and Sajal K. Das. LeZi-update: an information-theoretic approach to track mobile users in PCS networks. In *Proceedings of the Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pages 1–12, Seattle, Washington, USA, August 1999.

[4] Bradley M. Duska, David Marwood, and Michael J. Feeley. The measured access characteristics of world-wide-web client proxy caches. In *USENIX Symposium on Internet Technologies and Systems, 1997*, Monterey, CA, December 1997.

[5] David Kotz and Kobby Essien. Analysis of a campus-wide wireless network. Technical Report TR2002-432, Dept. of Computer Science, Dartmouth College, September 2002.

[6] Tong Liu, Paramvir Bahl, and Imrich Chlamtac. Mobility modeling, location tracking, and trajectory prediction in wireless atm networks. In *IEEE Journal on Selected Areas in Communications*, August 1998.

[7] Maria Papadopouli and Henning Schulzrinne. Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices. In *ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc)*, Long Beach, California, October 2001.

[8] P. Pathirana, A. Savkin, and S. Jha. Mobility modelling and trajectory prediction for cellular networks with mobile base stations. Annapolis, Maryland, June 2003.

[9] Diane Tang and Mary Baker. Analysis of a local-area wireless network. In *ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom)*, pages 1–10, Boston, Massachusetts, USA, August 2000.