Analysis of Movement and Mobility of Wireless Network Users

Jeremy Shaffer
Electrical and Computer Engineering
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213 USA
jshaffer@ece.cmu.edu

Daniel P. Siewiorek
Electrical and Computer Engineering
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213 USA
dps@cs.cmu.edu

Asim Smailagic
Institute for Complex Engineered Systems-Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213 USA
asim@cs.cmu.edu

Abstract

This paper describes a system developed for determining locations of devices on 802.11 wireless networks. Data representing 18,000 computers registered on Carnegie Mellon's Wireless Andrew is presented from traces taken in 2003 and 2005. Developers make many assumptions when creating applications for wearable and pervasive computers. The data collected by Locator@CMU provides a clearer understanding of large-scale wireless networks and their usage for implementing services and programs. Among the findings, we examine the mobility of a user as defined by percentage of time spent at their home site and favorite sites. Our results show that only a small number of wireless users exhibit high mobility and our data suggests that typical mobile users utilize the network only in a handful of sites. These basic patterns have remained steady over the past two years.

1. Introduction

The enormous growth of wireless local-area networks is well documented. In the rush to construct applications for them little work has been done to understand how these networks are utilized. Many assumptions have been made on how users might move around a network and models have been constructed based on them. These assumptions often include: relatively uniform distribution of users, constant connectivity/usage, and high movement of users.

This paper presents results from an analysis of user movement on Wireless Andrew at Carnegie Mellon University [1,2], one of the first large scale 802.11b/g networks constructed. It covers all buildings on the campus and virtually all open areas. A number of studies have analyzed wireless local and metropolitan area networks. One of the first was the work of Tang and Baker [5,6] who analyzed a building wide wireless local-area network and also a metropolitan local area network. Recently, researchers at Dartmouth University [4] have also analyzed their campus wide wireless local-area network. All three studies answered a wide array of questions. Many of these questions related to network traffic. Unlike past papers that only briefly consider mobility issues, we solely focus on issues related to mobility in this paper. We do not track any packet headers but use the wireless network infrastructure to locate the unique MAC Addresses of each wireless card.

Section 2 presents an overview of the Wireless Andrew network and Location Service implemented at Carnegie Mellon. Section 3 defines terminology, provides overall network statistics, and identifies two distinct user communities. Section 4 analyzes in detail the mobility and movement of wireless users. This allows us to draw a number of conclusions related to mobility in Section 5 and discuss future work in Section 6.

2. Wireless Andrew and Locator@CMU Location Service

Wireless Andrew, started as a research project in 1994 and has been greatly expanded since its conception. Currently, Wireless Andrew uses 640 access points to cover all 77 academic and administrative buildings of main campus, dormitories and key off-campus buildings. The wireless network has evolved from an initial 900MHz network to its current 802.11b/g protocol in the 2.4GHz range [7]. Even before the Wireless Andrew network was created its potential use as a research platform was noted [2].

Location information on wireless users, a key element of context aware computing, has been difficult to obtain. Interesting research has been done in the area of developing location systems that can provide precise coordinates and even orientation of the user. However, these systems [3,5,6] are limited in scope to single rooms, floors, or at most a building. It is

Proceedings of the 2005 Ninth IEEE International Symposium on Wearable Computers (ISWC'05)

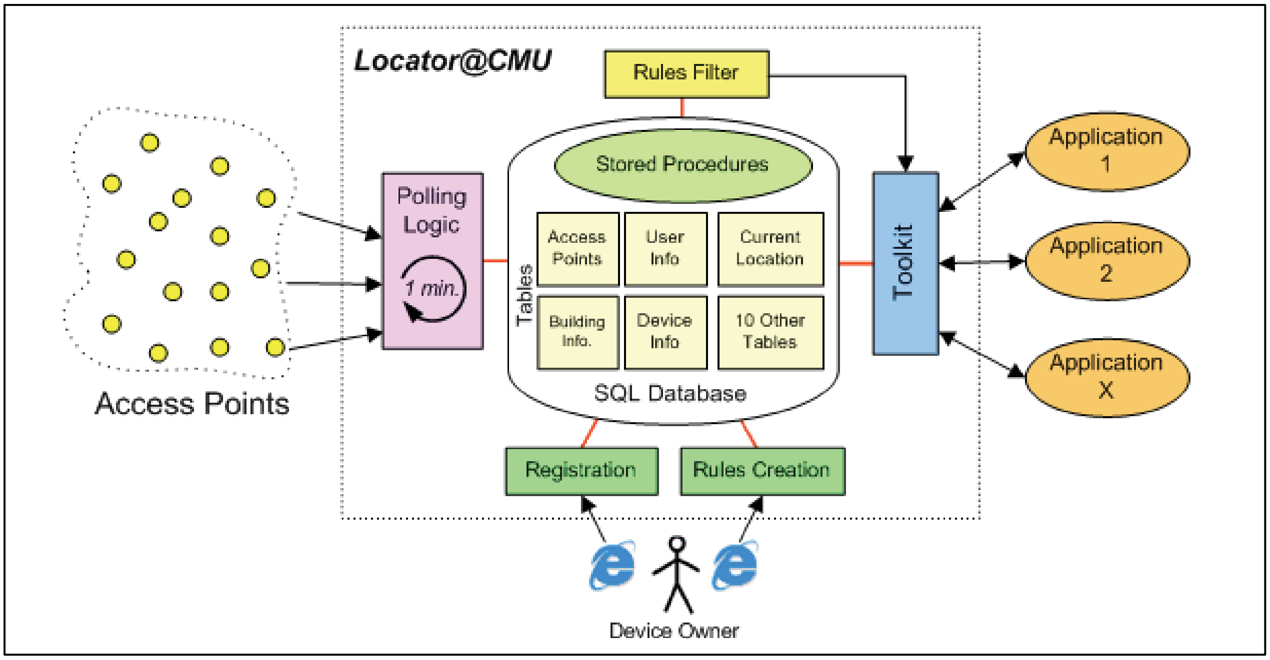

Figure 1: ***Locator@CMU System Diagram***, *all campus access points are polled every minute, information is stored in the Locator@CMU system (dotted line). Wearable and context-aware applications utilize the associated toolkit to access data.*

impractical to require new hardware on a large scale throughout the campus environment.

Locator@CMU is a centralized wireless location service and runs on a server computer with a database to store and retrieve all device locations. Figure 1 depicts a high-level system diagram. Wireless access points keep track of information on all connected devices. This information is stored in a format defined by the IEEE 802.11 MIB. When a device connects to an access point, it is entered into a table in the MIB. Locator@CMU uses SNMP calls to retrieve from each access point the list of all connected wireless devices defined by their MAC Addresses. Locator@CMU queries all 640 access points every minute.

All of the information retrieved from the access points is stored in database tables optimized to enhance system performance. Tables exist for instant retrieval of immediate user locations as well as historical movement patterns. The system allows users to create accounts and associate their wireless MAC addresses with their personal identity. Locator@CMU can provide location self-awareness to programs running on a user's wearable or mobile device. Additionally, the system allows the exchange of information between users and the programs they are running. With the correct permissions users have the ability to locate their friends and colleagues. This functionality supports a wide scope of location-based programs.

The issue of ensuring user privacy and system security is extremely critical. Without trust in the system many users would never voluntarily register their wireless MAC Addresses and associate their user identity with it. To develop this trust Locator@CMU has a web based interface to allows users to specify privacy rules on usage of the location information. These rules can be created based on a number of factors including time, user identity, location of requestee, and relationship to requestor (friend, classmate, professor). With these factors rules can be constructed for location divulgence such as: "only show to specific persons X, Y, and Z", "block person A when I am in building B", and "only show between 9AM-9PM to professors." To address security concerns the system has been designed to meet industry standards of encrypting all network traffic with SSL, limiting system administrative access, and monitoring of outside connections for potential attacks. The combination of the privacy rules and tight security produces trust in both the users that register with the system and the applications that rely on it for information.

Wearable and mobile applications that need wireless location information are supported by a toolkit/API for Locator@CMU. The toolkit, developed in Java, provides functions and data types to allow for quick integration of location data into programs. The Java implementation directly supports the device and operating system independent principles of the overall approach. Additionally, programs that desire to directly interface with the database can do so through ODBC connections and the SQL querying language.

The Locator@CMU system can locate all active users on a wireless network. Its access point based approach requires no software to be installed on the client device. The system also requires no training data and little configuration. It provides accurate location information across the entire CMU campus to within the radius of an access point. The two main drawbacks are granularity of its location information (75'-100') and the difficulty in following rapidly moving users due to the need to wait on polling information obtained from the access points. For the majority of applications this moderate level of granularity and updating speed are more than sufficient. Locator@CMU allows these applications to obtain accurate location information without the enormous overhead and costs associated with other methods. Over ten applications have been implemented using the Locator@CMU system including: a location aware-instant messaging client, activity recommending agent, wireless network visualization tool, and integration with a wearable context-aware phone system.

3. Overall Network Analysis

Students, faculty, and staff at Carnegie Mellon University use the Wireless Andrew network extensively. Approximately 50% of all students use the wireless network regularly in 2005. This percentage has doubled from 2003. By monitoring when wireless network cards connected and disconnected we are able to determine a wealth of information on individual and aggregate mobility and movement patterns. We utilize data collected from network traces on Locator@CMU during 2003 and 2005. We have selected comparable weeks of trace data in the spring semesters of the two years as a basis.

3.1 Terminology

Before describing user mobility, a number of terms are defined:

Wireless Card - A physical network device that allows network connectivity. The card is identified by its unique MAC address.

Access Point (AP) – Device connected to the wired campus network that allows client *wireless cards* to connect to the network. There are over 600 of these on the Wireless Andrew network.

Mobility - The number of access points the *wireless card* has been connected to within a period of time. For most metrics in this paper a 24 hour day is used for the period of time.

Dwell Time – The length of time at a given *access point*.

Home Site – *Access point* with the greatest *dwell time*.

Favorite Sites – Top five *Access points* where greatest *dwell time* is spent ranked in order from greatest *dwell time* to least.

User – For simplicity we associate a user with a given *wireless card*. From observation there is normally a one to one mapping between the *wireless card* and the human using it. However, it is possible for a *wireless card* to be shared or even more likely for a person to own multiple *wireless cards.*

User Session – The time spend by a user at a given *access point*.

3.2 Global Statistics

The Wireless Andrew network has over 18,000 registered wireless cards as of April 2005 this is slightly more than double the 8,000 registered cards in April of 2003. Peak usage during the observation period in 2005 was measured to be 4,965 unique devices for an average of 7.7 devices per access point during maximum utilization. The 2003 peak usage was 2,603 unique devices for an average of 4.0 devices per access point.

3.3 Categorization of Access Points

With the entire campus covered by the Wireless Andrew network there is a possibility for wide usage variation between access points. Exploring the usage of specific access points is useful to application developers such as e-coupons for vendors and collaborative team builders. Access points are located in different types of spaces varying from residential buildings, offices, classrooms, and open public space. Figure 2 shows the total number of user sessions for each access point over a four month monitoring period. From this graph we identify three different categories: hot, normal, and underutilized.

Proceedings of the 2005 Ninth IEEE International Symposium on Wearable Computers (ISWC'05)
0-7695-2419-2/05 $20.00 © 2005 IEEE

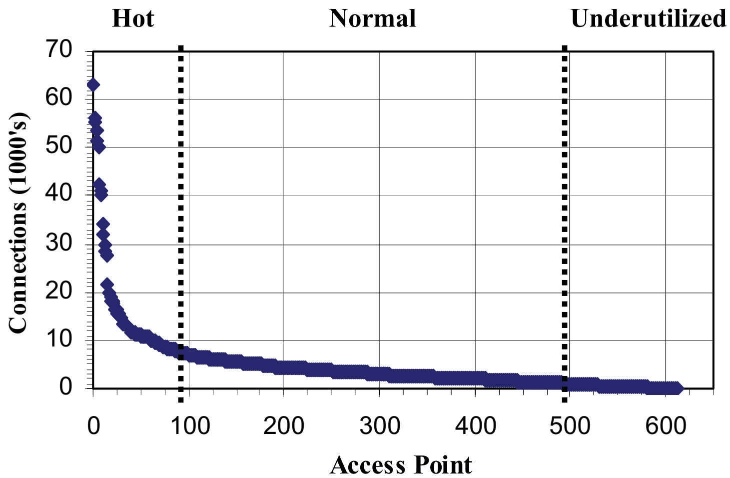

Figure 2 - *Connections Per Access Point*

Table 1 - ***Hot Spots by Building*** ** - The nine buildings that contained fewer than 4 hot spots were condensed together to save table space*

Building	Hot Spots	Top Ten Spots	Total Building APs	% of Building APs Hot
GSIA (Business)	17	9	18	94.4%
Hamburg Hall (Public Policy)	7	1	22	31.8%
Wean Hall (Comp. Sci.)	7	0	33	21.2%
Newell-Simon (Comp. Sci.)	6	0	15	40.0%
Baker/Porter	4	0	33	12.1%
Morewood E-Tower (Dorm)	4	0	11	36.4%
Other*	14	0	149	9.4%

Hot access points are defined as having over 10,000 user sessions over the monitoring period. There are 59 hot access points corresponding to roughly 10% of the total network. The hot access points are concentrated in four main buildings. The Graduate School of Industrial Automation (GSIA) is home to nine of the top ten access points. Additionally, of the building's 18 total access points 17 were in the top 10% of usage. This makes nearly the entire building a single hot zone of user activity. Upon investigation it was discovered that all GSIA students are required to have a mobile computer and much of the curriculum is actively designed around the mobile computer activities.

The top six buildings based on number of hot spots is shown in Table 1. Of the 77 buildings on Wireless Andrew six contained 76.3% (45 of 59) of all the hot spots. Hot access points are thus clustered in a handful of areas on campus. This shows that assumption of relatively balanced user distribution is not the case at CMU.

4. User Mobility and Movement

It is often assumed that all wireless users are mobile and actively move around the network. However, this need not be the case. For example, we have observed that there are even some desktop computers on Wireless Andrew. To study this assumption and define mobility we look at the number of unique access points that a user connects to over a given time period. In most cases we define this time period to be a single day. In classifying user mobility we examine dwell time of users at locations, the time spent at the user's home node, the user's favorite node, and also the overall mobility statistics defined by the average and median of the traces.

4.1 Dwell Time

The amount of time that a user spends at each location is an important characteristic of mobility. Table 2 shows the percent of total connections that fall within the given time categories for all users both in 2003 and 2005. Nearly 90% of connections were less than one hour in length in 2005 a significant jump from the 68% in 2003. Also interesting was that 5.7% of all users remained connected for between 18-24 hours a day in 2003 and in 2005 the number had fallen to below 1%. These users were most likely desktops with wireless cards. As Wireless Andrew has become more mature this practice has been discouraged leading to the drop.

We further studied the large number of connections less than an hour by dividing them in ten-minute increments as shown in Table 3. The largest grouping of connections for both 2003 and 2005 was between 11 to 20 minutes. A majority of all connections were

Table 2 - ***Dwell Time by Hour Based Categories***

Connection Length	2003 Percent of Total Connections	2005 Percent of Total Connections
0-1 hour	67.9%	89.3%
1-2 hours	12.1%	6.5%
2-3 hours	6.2%	1.8%
3-6 hours	5.5%	1.6%
6-12 hours	2.2%	0.6%
12-18 hrs.	0.7%	0.2%
18-24 hrs.	5.7%	0.1%

less than 30 minutes for both years. Many applications such as checking and responding to email do not require a long amount of time and correspond to this usage pattern. The median connection time in 2003 was 26.0 minutes in 2005 it had fallen 36.5% to 16.5 minutes which was nearly identical to the 16.6 minutes observed on the Dartmouth network [4]. The short network connections suggest that many applications should be designed for quick interactions.

4.2 Home Site

The home site is the network access point that the user is connected to for the greatest amount of time. Users who spend extensive time at a home site range from completely stationary workstations to mobile computers that are utilized primarily in a single office or dormitory room. To classify the home site we calculated the time spent by each MAC address at each access point. The access point that had the greatest time spent at it was classified as the home site.

Figure 3 shows that a majority of users spend over half of their time at their home node. Specifically, in 2003 62% of all wireless users spent 50% or more of their time at their home location. In 2005 this figure rose to 73% of all wireless users. Slightly over 8% of users in 2003 and 19% of users in 2005 never connected to more than a single access point. These users exhibited no mobility and are likely laptop computers left in an office and never moved. The home node and its performance are particularly critical to the large number of users who do not exhibit extensive movement patterns. Using the home site metric alone would suggest that mobility has decreased significantly from 2003 to 2005.

4.3 Favorite Sites

In addition to a home site many users have favorite sites that exhibit high frequency of usage. Many users

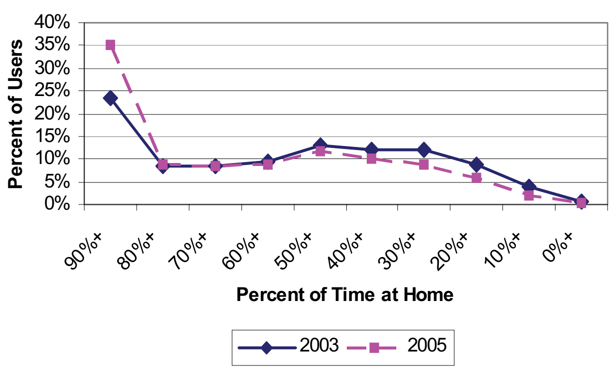

Figure 3 - *Percent of Time at Home by user groups*

Table 3 - ***Dwell Time for under one hour connections***

Connection Lengths Under 1 Hour	2003 Percent of Total Connections	2005 Percent of Total Connections
1-10 mins.	14.2%	24.3%
11-20 mins.	26.4%	34.5%
21-30 mins.	12.3%	15.5%
31-40 mins.	7.1%	7.2%
41-50 mins.	4.6%	4.2%
51-60 mins.	3.2%	3.3%

follow set routines that take them to a handful of places throughout the day. For example, a student's class schedule can provide a predictable pattern of movement. Many research faculty also have a set schedule in which time is split between different labs and an office.

To classify the favorite sites we calculate the time spent by each MAC address at each access point. We then take the top five access points by time; this includes the home site discussed earlier. In 2003 just 3% of users spent less than 50% of that time at their top 5 sites. However, this number had jumped to 12% in 2005 showing that a new more mobile sub-group of users was developing. Still, both in 2003 (73%) and 2005 (69%) over two-thirds of wireless devices spent more than 90% of their connected time at their top 5 favorite locations. The majority of users are only concerned about network services and applications available in a handful of locations

The overwhelming feature to note is how little time all users spend away from their favorite sites. For the overall user community slightly over one-third of users spent no time outside of their top five favorite sites.

4.4 Mobility

The metric of mobility is defined as the number of access points the user connects to per day. Stationary users will exhibit mobility values of 1 while high movement users will have large values. To calculate the mobility values of all wireless users on campus we determine the number of unique access points that each MAC address accessed for each day over the entire evaluation period. For a user's total mobility we take the sum of all their mobility values for days they are on the network and divided by the number of days. We felt this was more reliable then using the total number of days since many users are away from the network on the weekends, breaks, or trips.

The mobility results in Figure 4 show that most users are not very mobile confirming the same trends through our study of home and favorite sites. In 2003 only 10.5% of all users had average mobility values of

Proceedings of the 2005 Ninth IEEE International Symposium on Wearable Computers (ISWC'05)
0-7695-2419-2/05 $20.00 © 2005 IEEE

Table 4 – Average, median, and standard deviation of mobility for users

	Average Mobility	Median Mobility	Standard Deviation
2003	2.85	2.18	2.75
2005	2.81	2.01	3.01

5 or above this number had increased slightly to 11.7% in 2005. The median mobility actually fell slightly from 2.18 to 2.01 from 2005 to 2003 due to the increase in users in 2005 whose average mobility was less than 2. Users who would be considered highly mobile and benefit the greatest from context-aware application make-up only a small percentage of Wireless Andrew users. However, there does exist a small subset of users that is actively embracing the network. The percentage of network users that averaged connecting to 10 or more access points per day in 2003 was 1.2% and this ratio more than doubled to 2.7% in 2005.

Table 4 shows the comparison of 2003 and 2005 mobility statistics for users. These numbers appear to show that mobility decreased slightly between 2003 and 2005. It is important to remember that the entire user population also doubled between these years and different trends were at work often having conflicting effects on mobility as discussed in the Conclusions section.

4.4 GSIA User Sub-Group Mobility

In section 3.3 we showed that the entire GSIA building is a single hot zone on campus representing 9 of the 10 most actively used access points on campus. Approximately 1/7th of the total Wireless Andrew users belong the GSIA sub-group. These students have institutionally embraced network and actively use it in their classes and group work. We believe that these users demonstrate the potential future utilization of all others.

We have calculated many of the mobility criteria for just this sub-group of users. The average mobility of GSIA users (3.81) is 42% higher than that of all non-GSIA users (2.69). The dwell time statistics for GSIA and non-GSIA users are nearly the same, within 2% of each other for all categories. This shows that GSIA students utilize the network more, but the time patterns of usage are still in proportion to non-GSIA users. When studying the home site and favorite site metrics we also observe that GSIA users are much more mobile. GSIA users spend an average of 34% less time at their home site then do non-GSIA users.

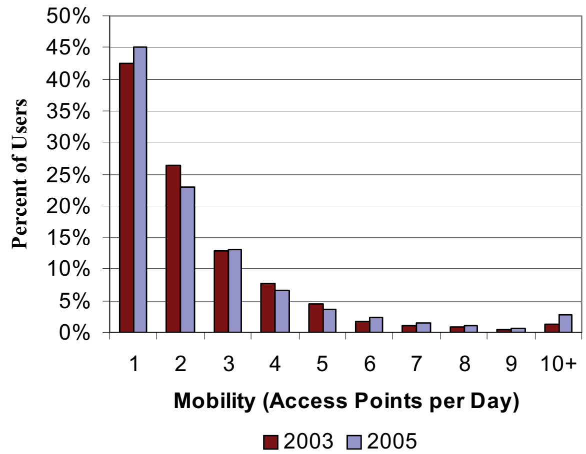

Figure 4 - *Mobility Percentages.* *Graph shows % of users whose average mobility falls within the given ranges.*

5. Conclusions

This paper provides a better understanding of user mobility on a wireless local-area network. Current users are not very mobile, on average using less than 3 access points per day both in 2003 and 2005. From the data however we observe two trends. First, a large increase in total users has occurred. Between 2003 and 2005 it is now common for laptops and other portable devices to have built-in wireless connections. No longer is it just users who must make the effort to purchase the wireless card who use Wireless Andrew. This increase in the overall number of users has led to many more people who use the network as a convenience in 1 or 2 locations. However, we also detect another trend. A small highly mobile sub-group of users is developing. These users connect to the network in more than 10 different spots per day on average. Our analysis shows that the total number of these highly mobile users has more than quadrupled in two years while the total number of users only doubled. As more users acquire wireless devices such as PDAs, wearable computers, and tablet computers we would expect this group to continue to grow.

The analysis of usage by home site and favorite sites shows how strongly a user's perception of the network can be effected by the services and applications available in a specific location or small set of locations. Overall users tend to spend most of their time in a handful of places. Devices are seldom left on when traveling between locations and even when they are the total time spent at these intermediate points is negligible. Context-aware applications that are implemented sometimes support only a single building or area of campus. This approach may not work if there are not many home sites in that area or most of the access points that serve the area are in the underutilized category.

The current usage patterns of wireless networks do not lend themselves to programs that depend on the device to be continuously on. Application and device developers need to continue to provide tangible incentives and benefits to encourage users to more actively use their devices. Current programs such as email, web browsing, and instant messaging all are normally used when a user has arrived at a stationary position and will be there for more than 5 minutes. Services such as E-Coupons beamed to users walking in proximity of a store would not work under the current network usage patterns since few users keep their devices connected.

The Wireless Andrew network has increased in usage significantly in the past two years. Despite this large influx of new users the basic mobility and movement patterns have remained relatively the same due to two conflicting trends of more highly-mobile users while also more devices that have wireless standard for convenience. We anticipate network users are evolving into two distinct classes: first of highly mobile and active users and second those that appear to use the network at only a handful of places. If the rate of growth of the highly mobile users continues its quadrupling every two years it will provide a strong future user base for consuming context-aware applications.

6. Future Work

This work continues the research that has been conducted on the Wireless Andrew network at CMU. Future work will continue the broad thrust on exploring 802.11 wireless location services and their use by wearable and context aware applications. It is our intention to use feedback from various outside research groups to independently confirm our findings and verify the completeness of a new CMU location service toolkit being developed. Implementations of additional wearable applications that use the location service are also planned.

In addition to the access point polling approach, we have also developed another location service, called Triangulation, Matching and Interpolation (CMU-TMI) [8]. This service will be integrated with the Locator@CMU toolkit to allow an additional location data source from the device. When the high overhead is overcome to make the service available on a device, this information is more accurate and timely than access point location data. To obtain the location position in CMU-TMI we perform two distinct transformations on the raw data of signal strengths. We first calculate the client's position on a continuous coordinate grid, assuming signal strengths map directly to distance. We then map the resulting coordinates onto real space coordinates using a set of trained values. Combined with the current access point approach CMU-TMI will allow for continuous monitoring with accuracy of 5 meters over 90% of the time. For rapidly prototyped applications that require high levels of accuracy and continuous location updates, the CMU-TMI algorithm combined with the toolkit offers the best options.

Both the Locator@CMU access point approach and CMU-TMI have unique strengths and weaknesses. Locator@CMU offers the capability of implementing a robust, device independent service that can cover the entire network area no matter how large. CMU-TMI requires software to be installed on each client. This often presents difficulties since only the most popular operating systems and hardware combinations can be supported. Additionally, CMU-TMI requires training measurements to be taken in the desired location area. This upfront work can be prohibitive to effective implementations across large areas. CMU-TMI offers location accuracy on the order of feet for wearable and mobile devices. Locator@CMU is limited to the accuracy of the access point range typically 75' to 100' on the CMU campus. CMU-TMI offers instant updates on location positioning while Locator@CMU is limited by the polling rate of the access points.

In general Locator@CMU can be used to provide an accurate building/floor level location service for every wireless device. CMU-TMI can be implemented to provide for greater detail to support certain devices and locations that need greater accuracy. Working in tandem the two services can provide the optimal coverage for wearable and mobile applications.

Privacy and security is a major issue that has arisen with the implementation of the location service and applications. We are continuing to explore this complex area with the Locator@CMU system. Various privacy rules and features created will be implemented and tested using context-aware applications to evaluate their effectiveness in a variety of situations.

Lastly, we also continuing to conduct experiments and evaluate network traces. There are many wireless device types that utilize wireless networks. The results presented in this paper present the aggregate information on these device types. By allowing users to specify their device type (mobile computer, PDA, wearable computer, etc.) on network registration we will be able to further analyze network usage and mobility patterns.

7. Acknowledgements

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. NBCHD030010 and the National Science Foundation under Grant Nos. 0205266 and 0203448.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Advanced Research Projects Agency (DARPA), or the Department of Interior-National Business Center (DOI-NBC).

8. References

[1] Benngington B. and C. Bartel. “Wireless Andrew: Experience in building a high speed, campus-wide wireless data network.” *In Mobicom 1997*, pages 55-65. ACM press, September 1997.

[2] Hills, A. and D. Johnson, “A wireless data network infrastructure at Carnegie Mellon University.” *IEEE Personal Communications* 3(1): 56-63, 1996.

[3] Hightower, J., Boriello, G. “Location Systems for Ubiquitous Computing.” *IEEE Computer* 33(8), August, 2001.

[4] Kotz, D., and K. Essien. “Analysis of a campus-wide wireless network.” *Proceedings of lthe Eighth Annual International Conference on Mobile Computing and Networking,* pages 107-118, Atlanta, Georgia, 2002

[5] Tang, D., and M. Baker. “Analysis of a local-area wireless network.” *Proceedings of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pages 1-10, Boston, Massachusetts, August 2000.

[6] Tang, D., and M. Baker. “Analysis of a metropolitan-area wireless network.” *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pages 13-23, Seattle, Washington, August 1999.

[7] WaveLAN. http://www.wavelan.com. (*now Agere*)

[8] Smailagic, A. and Kogan, D., “Location Sensing and Privacy in a Context-Aware Computing Environment”, IEEE Wireless Communication, Vol. 9, No. 5, 2002, pp. 10-17.