Measuring early usage of Dartmouth's wireless network

Pablo Stern
pablo.stern.01@alum.dartmouth.org
Senior Honors Thesis
Dartmouth College Computer Science

Technical Report TR2001-393
Advisor: Professor David Kotz

June 1, 2001

Abstract

In Spring 2001, Dartmouth College installed a campus-wide 802.11b wireless network. To understand how that network is used, we examined the usage characteristics of the network over a five-week period. We monitored access points to determine user behavior, and user and network traffic characteristics. Because our study coincided with the deployment of the access points, our analysis captures the growth of a wireless network. The results of this study help understand the behavior of mobile users and provide a reference to network engineers wishing to deploy and expand similar wireless networks.

1 Introduction

Dartmouth College, in conjunction with Cisco Systems, Dell Computer, Apple Computer and its alumni, has installed a system of Cisco Aironet 350 access points throughout its campus. These devices create a campus-wide wireless network, allowing the entire community to access both the campus network and the Internet through untethered means. The new system provides universal access to the network and allows freedom of mobility for users [1].

This paper examines a five-week trace of the network during Spring 2001. We installed a system to monitor the network and determine the characteristics of the wireless user. We analyze user behavior, such as periods of activity, movement, and types of protocols used. The study also examines access-point statistics, including traffic density and error rates.

During the five-week trace, Dartmouth's Computing Services department was actively installing access points. At the beginning of the monitoring period there were six access points, providing wireless access to a limited subset of the campus. By the study's end, 278 access points had been installed, covering 90% of the campus. The growth in wireless coverage during the study, coupled with growing knowledge among users about the presence of the network,

This project supported by the Cisco University Research Program.

provides data that highlights the deployment of the network rather than a steady-state use of the network.

A total of 249 user devices were recorded during the trace, generating 160GB of traffic.[1] Throughout the trace the number of users and amount of traffic grew notably. Generally, activity was greatest during the business day, and lowest on weekends and during the early morning hours. A protocol examination of a subset of campus users shows that web surfing was the predominant use for the network, however protocols like FTP, SMTP, and SSH also figured prominently.

The results of the study provide insight into how a large-scale wireless network grows and how it is used. Characteristics, such as mobility and user concentration, aid in optimizing the network and deploying similar systems.

In the next section, we examine how a wireless network is implemented, what research has been done into these networks, and the role of the Dartmouth community in the study. Then in Section 3, we expose our method for capturing data. In Section 4, we analyze the behavior of individual users on the network, and in Section 5 we examine the characteristics of access points. In Section 6, we look at the traffic through the entire wireless network. Finally, in Section 7, we focus on protocol statistics for a subset of the wireless community.

2 Background

In this section, we look at general concepts for wireless networking, followed by a review of existing literature about the use of wireless technologies, and an examination into the growth of Dartmouth's wireless network.

2.1 Implementation of a wireless network

In recent years we have witnessed the increasing availability of wireless networks, largely due to new standards for communication. The basic components for most wireless networks are an access point (or base station) and a network interface (also known as a client card). The access point provides an interface between the tethered network and the wireless world. The client card plugs into a computer and communicates with the access point. Certain wireless standards exist, such as Bluetooth and IEEE's 802.11b. The 802.11b standard is well established and is supported by many manufacturers, such as Cisco, Apple, Lucent, and 3Com. The Cisco Aironet access points installed at Dartmouth support the 802.11b standard and allow for 11Mbps transfer rates [6]. The Aironet access points have a range of 130-350 feet indoors, and 800-2,000 feet outdoors [2]. This range depends on modulation and interference.

An access point's range is relatively small. Thus, one of the goals of wireless networking is to allow the user to roam seamlessly between different access points. By placing access points within range of each other, a large wireless network is created. To the user, transitions between access points (known as *hand-offs* or *roaming*) are transparent.

The IEEE 802.11 protocol is designed to interface with the canonical IEEE 802 wired network. Therefore, a wireless user will have access to the all network resources on adjoining wireless or tethered networks (such as web-surfing, and peer-to-peer communications).

[1] For our analysis, references to bytes are in powers of 2 (i.e., 1KB: 1024 bytes, 1MB: 1024KB, etc.).

2.2 Previous research

In this section, we examine previous studies of wireless networks, which provide comparison data for our analysis. The wireless domain is a recent source of research.

In 1994, Carnegie Mellon began a wireless initiative with the support of the National Science Foundation. The University based its wireless network on Lucent's WaveLAN access points [4, 5]. Currently, CMU's network covers all academic and administrative buildings on campus, as well as, common spaces, offices, and some outdoor areas. In 1996, a study into the error characteristics of wireless networks was performed on the CMU network [3].

A 1997 study from Worcester Polytechnic Institute examines the performance issues of a wireless network on a campus [7]. This study discusses the effects of obstacles (such as walls and floors) and interference on wireless networking. The authors conclude that a properly planned wireless network can provide a practical extension to a wired network.

Two Stanford studies [8, 9] analyze wireless networks to determine usage characteristics. The first study used a seven-week trace of a metropolitan-area packet radio wireless network based on Metricom's Ricochet technology. This large-scale (14,053 repeaters) study examines mobility and network use, but does not address issues concerning a computer-based network, such as protocol information. We also expect that the lower bandwidth (28Kbps) and larger metropolitan area may lead to characteristics different than a high-speed campus network.

The second study is a twelve-week trace of a WaveLAN local-area wireless network in Stanford's Computer Science building. For their analysis, the authors routed all access points through a *sniffer* and used Simple Network Management Protocol (*SNMP*) information to determine wireless client information. The study finds sub-communities in the user community, each with characteristic usage traits. These traits include daily and weekly use patterns, and protocol biasing. This study also examines the overall network characteristics, noting usage peaks in the mid-afternoon and a greater amount of bytes arriving at the access points than leaving.

The study of Stanford's Computer Science building provides a model for our campus-wide analysis. The analysis examines user behavior, network traffic characteristics, and user traffic characteristics. Where relevant, in our study, we draw parallels to this Stanford study.

2.3 User community

To analyze Dartmouth's wireless network, we must have a sense for its geographic and demographic composition. This information provides preliminary insight into the user community.

Relative to most universities, Dartmouth's campus is small. The campus covers approximately 200 acres, and the bulk of academic, administrative and dining facilities line the perimeter of a 400 square-foot park, known as the Green. A large portion of the access points have been installed around this core (see Figure 1). The College has 5,500 students, with graduate students accounting for one-fifth of the total. The faculty includes 1,215 full-time professors. Dartmouth established itself as a primarily Apple-based campus, although in recent years there have been more Windows and Unix computers in use. Every student is required to own a computer, and most students have desktop Macintosh or Windows systems. Students are increasingly selecting portable computers such as

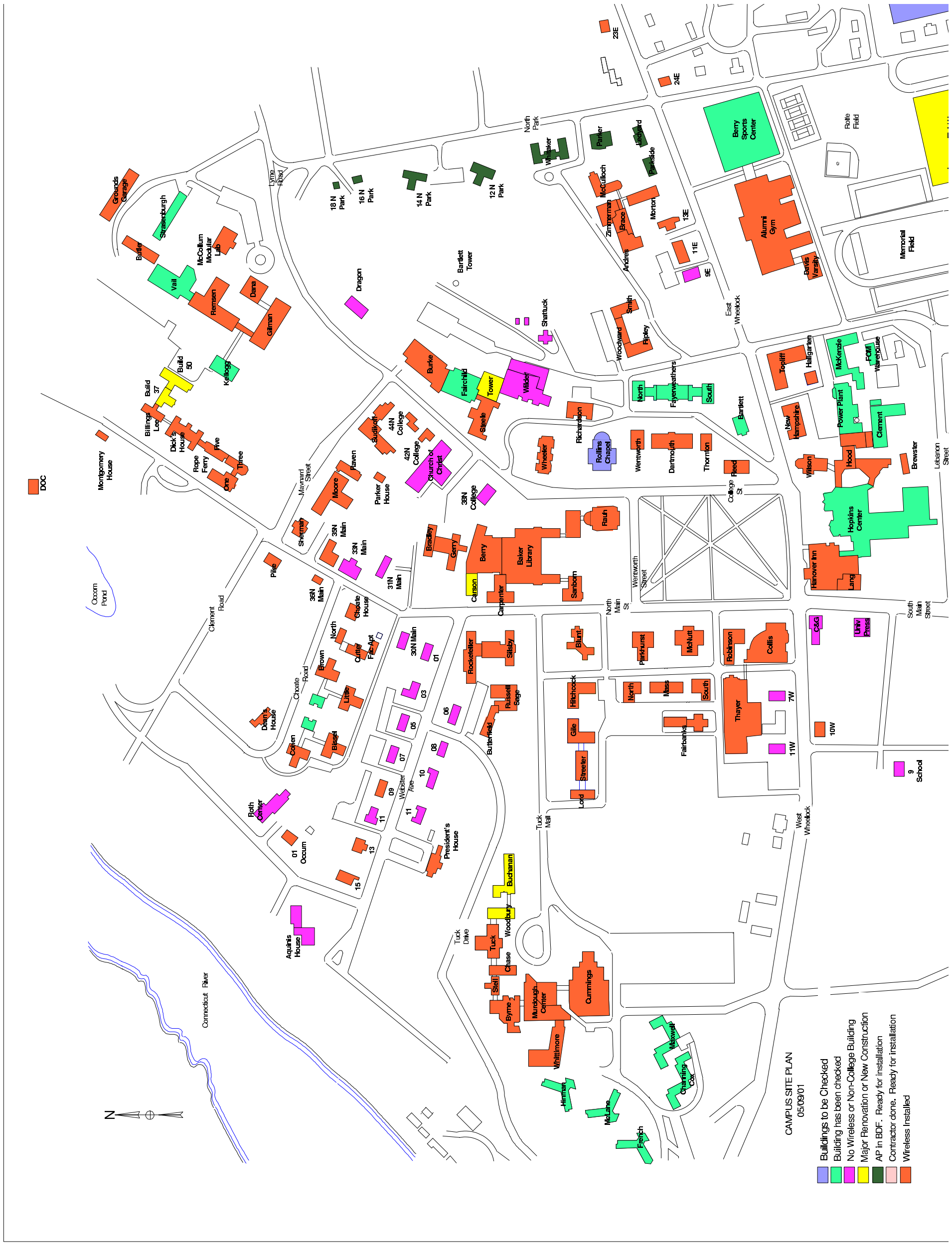

Figure 1: Dartmouth College campus map. Shading indicates the status of the wireless network on May 9, 2001, the mid-point of our study. Best seen in color.

the iBook, the PowerBook, and Windows-based laptops. Last year, the ratio of desktops to laptops purchased by incoming undergraduates was 2:1, whereas in other years the percentage of laptops had not exceeded 15% [10].

At Dartmouth, every building is connected to the wired network, and is structured with one subnet per building (mostly). Therefore, in each building, access points are on the same subnet.

In May 1999, Dartmouth installed 15 Lucent WaveLAN wireless access points in 10 buildings on campus. These devices allowed laptop users to use wireless network cards for access to the wireless network. During this initial deployment, the number of people with client cards is estimated to have been about 50 [10]. In May 2001, the Cisco access points have replaced the old access points and nearly covered the campus. As a result, a portion of the campus population already has the required equipment for wireless access. Furthermore, certain departments have seeded users with client cards. The Computing Services department distributed 30 Cisco Aironet client cards, and the Computer Science department distributed another 40. A further 12 Cisco cards were distributed in Cummings (the Engineering school building) for a course on Information Technology. Also, to date, about 60 people have purchased Cisco Aironet cards through the campus computer store. Also, the computer store has sold approximately 500 Apple Airport cards in the past year.

3 Data collection

Our five-week analysis involved two traces: an SNMP trace of all access points, and a tcpdump sniffer trace of selected access points.

For the SNMP trace we used a computer, with a database of active access points, to query each access point every five minutes. Using SNMP, we retrieved statistics from the access point's Management Information Base (MIB). We measured the amount of wireless traffic and statistics about all connected users (see Table 1). For each wireless user, we periodically logged the following statistics:

- timestamp
- MAC address
- IP address
- data sent/received
- data errors sent/received
- connection state

The MAC address provides a unique identifying tag for the user's client card, and the IP address is specific to the subnet. The connection state identifies whether the client is authenticated and/or associated. Authentication proves the identity of a device to another, while association is the logical connection between a mobile host and an access point [6]. In this study, we monitor sessions based on authentication status.

The second trace is a *tcpdump* trace of packet header information.[2] We set-up tcpdump to capture Internet Protocol (IP) traffic, such as TCP and UDP packets. We collected this trace using a sniffer device connected to the same hub as the access point that is being monitored (see Figure

[2] The packet header information specifies information (such as source, destination, port, and size) for a given packet. This data specifically excludes the contents of the data, such as usernames, passwords, filenames and files.

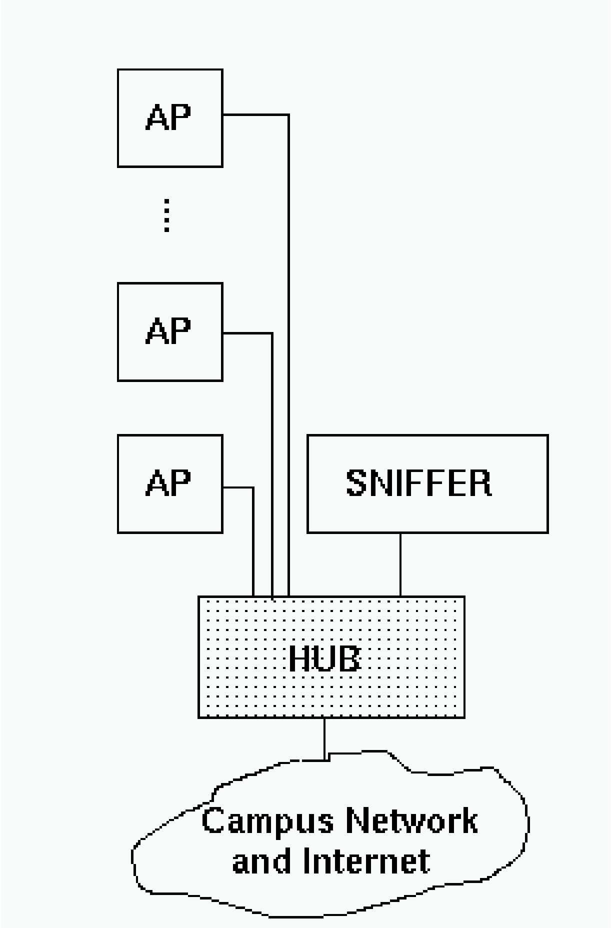

Figure 2: Diagram of how a sniffer monitors the traffic on a set of access points.

2).[3] Attempting to perform this trace on the entire campus network of access points would have been costly and inefficient. We chose to install a few of these devices around campus to gather a trace of a limited subset of access points. Ultimately, we were able to connect a sniffer to a switch serving all the access points in Dartmouth's Computer Science building (Sudikoff), to trace all the wireless traffic in this building (see Table 1). Our other sniffers provided less information. We had one installed in Collis, a popular student area. Unfortunately, this device was destroyed during a power surge. Another sniffer, installed in an academic building (Silsby), was connected to a seldomly used access point. Therefore, we rely on tcpdump data captured in Sudikoff.

The Computer Science building proves to be one of the more interesting wireless buildings at the College. Apart from the disposition among professors and students to use the computer network for research, Sudikoff was one of the first buildings on campus to have the wireless system installed. Furthermore, Cisco donated 40 client cards specifically for persons in the department. As such, Sudikoff has been one of the more active wireless locations on campus.

In Table 1, we see overall statistics for both traces. We note that the scope of the sniffer trace is much more limited than the campus-wide SNMP trace.

[3] This device passively records the packet information traveling through the network. To assure user privacy, our sniffing was limited to header information. See Appendix A for privacy statement.

Table 1: Statistics for our traces.

	Campus (SNMP)	Sudikoff (Sniffer)
access points	278	6
users	249	49
traffic	160GB	11.0GB
duration	5 weeks	4 weeks

Table 2: Representation of client card manufacturers.

Manufacturer	Company ID	%
Agere Systems	00:02:2D	5.2
Apple Computer	00:30:65	35.7
Aironet Wireless Comm	00:40:96	45.7
Lucent Technologies	00:60:1D	10.8
Other		2.0

4 User behavior

Each IEEE 802.11b client card has a unique Medium Access Control (MAC) address that distinguishes that card. This address tends to individualize wireless devices and approximately individualizes users. While the majority of users have one wireless client card, some have multiple cards. Also, the MAC address provides client card manufacturer information.

In this section, we examine several aspects of user behavior, asking:

- How many people use the wireless network?
- What are their session characteristics?
- What are the traits of the mobile user?

4.1 Number of users

Although certain users may have multiple wireless client cards, it is difficult to distinguish these users. Thus, for the purpose of this study each client card identifies one user.

In this study, we define an active user as one who is authenticated, but not necessarily associated. By default, access points will deauthenticate a wireless client after 30 minutes of inactivity.

During the study, we encountered 249 unique MAC addresses. Aironet Wireless Communications, the manufacturers of the access points, figured prominently among client card users (see Table 2).[4] By the end of the study, 142 Cisco Aironet client cards had been sold or distributed on campus. Due to the large population of Apple users, Apple cards also figured prominently. Between these two manufacturers the campus store has been selling one to three client cards per day. The Lucent cards are a mainstay of the earlier Lucent WaveLAN network.

4.2 Session characteristics

Due to the continued deployment of access points during the study we notice a growth in the number of active users throughout the study (see Figure 3). This trend indicates a growing knowledge among the user community of the network's presence. In this early period, knowledge about the wireless network was limited to technical staff and researchers. There was no concerted effort to advertise the network or encourage usage. The majority of the users learned about wireless accessibility by trial and error, or word of mouth.

[4] Cisco Systems acquired Aironet in November 1999.

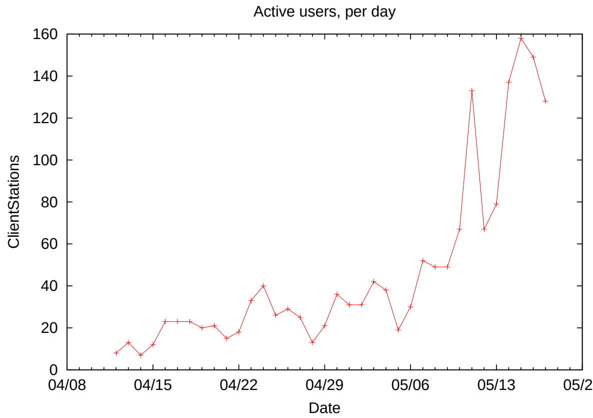

Figure 3: Number of active devices on each day of the study. The labeled dates are Sundays. The large spike on 05/11 was caused by the addition of 212 new access points to the study. Taken from campus SNMP trace.

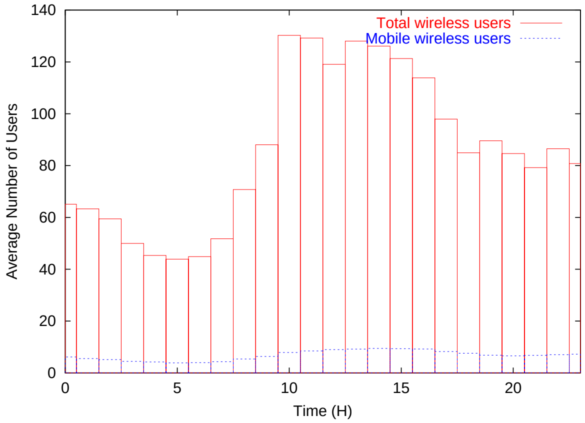

Figure 4: Hourly use of the network averaged over the life of the trace. User mobility is defined as the presence of one user on two or more access points in a given hour. Taken from campus SNMP trace.

Another factor potentially causing the increased use of the wireless network has been the weather. As springtime weather has improved, students have been frequenting outdoor areas with their laptops.

Weekends demonstrate a sharp decline in wireless use. The greatest drop is generally noted on Saturdays, with a 30% decline from the previous day. We mainly attribute this drop to the departure of administrators and faculty during the weekend. While some researchers and faculty members may use the network on weekends, we suspect that the majority remain off-campus during these days. Also, we note that Saturday appears to be the quietest day. This trend may indicate a day off after a hard school week. The slight rise on Sunday could be homework and studying to be completed for the return to classes the following day.

During the day, user presence peaks in the mid-afternoon (see Figure 4). During the early hours of the morning, when most people are sleeping, network use is lowest. Network use peaks during the academic school day (between 10am and 4pm), and is generally high during the business day (between 9am and 5pm). There is a slight drop in the evening, reflecting the departure of non-resident users. This trend in user presence is analogous to the findings in the Stanford study [9].

We define user mobility as the presence of one user on two or more access points within a given hour. In Figure 4, about one-tenth of total active users are considered mobile by our metric. In Dartmouth's attempt to cover the entire campus with wireless access, there exist many locations where the range of access points overlap. As mentioned in the *Background* section, the IEEE 802.11b standard allows for a seamless transition from one access point to another. Therefore, a motionless user could shift associations from one access point to another without knowing. Likewise, a minimal amount of motion (such as moving a laptop across a desk) could cause this shift. Thus, it is common for users to have visited more than one access point without physically moving. These users are classified in the same mobile category as users who walk around while remaining connected. Regardless, the fact that a resounding majority of users are generally immobile is suggestive. Either these users are not in range of a tethered connection, or they prefer the simplicity of a wireless connection. All of Dartmouth's buildings are equipped with numerous network jacks; the only campus locations where a cable connection is infeasible is outdoors and inside big halls, such as cafeterias and classrooms.

Throughout the study one-third of all users visited only one access point, while the majority of users had visited less than five (see Figure 5). A couple of outliers (shown) can be attributed to technical staff and researchers, who were probably testing access points to ensure proper functionality. A large percentage of users (41%) remained solely within the same subnet (i.e., building). One-third of users did a minimal amount of roaming beyond their primary location, but remained largely within the same subnet. The remaining 26% of users could not be associated with a particular location.

Most users (90%) who roamed between two and five access points remained within the same subnet at least half the time. These figures suggest that the users were roaming within a building, or were motionless but nonetheless switched access point. With a majority of these users remaining in the same subnet (usually equivalent to the same building), we can conclude that many probably had an office or work area in this location. It is likely that these users have tethered network access nearby, but chose the wireless connection because of its simplicity.

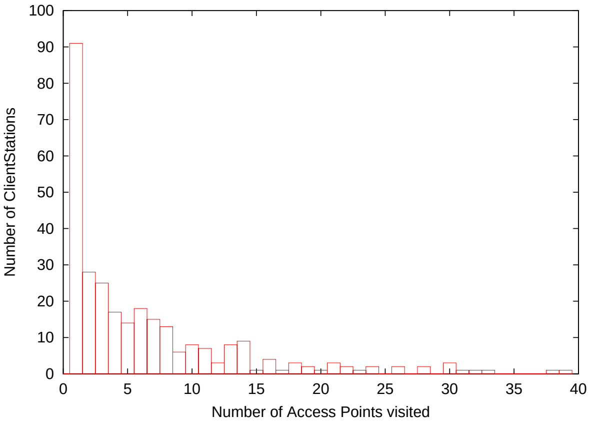

Figure 5: Number of access points visited by wireless users during five-week trace. Taken from campus SNMP trace.

We define a session as the duration of a user's presence on one access point. We polled each access point once every 5 minutes; if a user was no longer present during a poll their session was over. The granularity of our measurement is 5 minutes, therefore our session length calculations can be up to 10 minutes shorter than actual sessions.

By our metric, roaming is considered starting a new session, therefore some of these shorter sessions may have been continued on another access point.

As we can see in Figure 6, the highest peak in session duration is between 20 and 30 minutes. The access points have a 30 minute timeout for users that do not disassociate. Hence, if a user was to log onto the network and then break the connection without disassociating, the access point would not terminate their session for 30 minutes. Therefore a number of sessions between 20 and 30 minutes may represent short sessions (less than 10 minutes) that fail to disassociate. Since real sessions may last up to 10 minutes longer than our calculated duration, a short un-disassociated session (i.e., one that times out after a total of 30 to 40 minutes) would fall into the 20 to 30 minute range in the plot.

The majority of sessions last between 20 and 60 minutes. As session length increases, there is a gradual decline in the number of sessions. A consideration for the high number of short sessions is roaming. It is possible for a stationary user to roam between two or more access points without knowing it. Also, in a school environment where classes last between 60 and 120 minutes and where users may bring their laptops to libraries to work for short periods of time, this session-length distribution does not seem unlikely.

As noted in the caption, there do exist a few sessions that have lasted longer than a day. These users are probably administrators, faculty or graduate students that leave their machines on, at their desks, in perpetuity.

5 Access-Point Characteristics

At the beginning of the study there were six active access points. There were 278 by the end of the five-week study. In this section, we examine access-point characteristics by asking:

- Which access points were most active (i.e., where are people using the wireless network)?
- Are there differences between the amount of traffic and the number of active users?

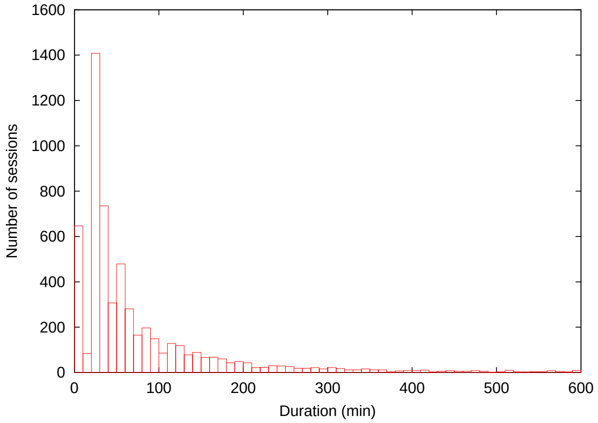

Figure 6: Distribution of session lengths (divided into 10 minute increments). Session length is characterized by the length of presence on one access point in an authenticated state. There exist a minimal number of sessions extending beyond 600 minutes. These sessions account for 0.06% of total sessions, with a maximum session length of 20 days. Taken from SNMP trace; due to the coarseness of the trace we cannot detect session of less than 5 minutes.

- How does the type of building affect activity?

Throughout the study, Dartmouth's Computing Services was actively installing access points. This growth causes an imbalance in the total amount of traffic seen by different access points. Therefore, for comparison purposes, we present individual access-point statistics averaged over the number of days that an access point existed.

In Figure 7, we note the growth in availability and use of access points. While not all access points are used in one day (the average number of access points in daily use was 51%), the number of access points in use gradually increases. On the second-to-last day of the study, 187 out of 278 access points were in use (67.3%). We can account for this rise largely by the deployment of new access points over the course of the study; as the reach of the wireless network extended so did its use. There are two notable rises in the amount of access points installed, on 04/23 and 05/10. We added access points to our study as Dartmouth's Computing Services informed us of their presence. Thus, there may have been a delay between the installation date and when we were notified of an access point's presence. Also, we see that on several occasions the total number of access points dropped, due to some testing and repair of misconfigured access points.

We can draw a parallel between Figures 7 and 3, both in the growing use of the network and in the dips noted on weekends.

We now consider the busiest access points, in terms of bytes moved (we rank the access points based on traffic per day because access points were activated on different dates during the study). The busiest access points (see Table 3 and Figure 8) can be traced to Berry/Baker library. The Berry/Baker library facility is primarily used for research, and visited by many students. It is also home to Computing Services, (including, Technical Services, Information Systems, and User Communications) which may be a reason for the large amount of traffic. For example, the berry1-ap, berry4-ap, baker12-ap, and baker11-ap access points are located close to a computing department and thus get used by technical staff. As mentioned, this department was seeded with 30 wireless client cards. The berry13-ap, berry3-ap, and berry2-ap access points are located in the library's cafe (Novack), which is a popular location for students to do work.

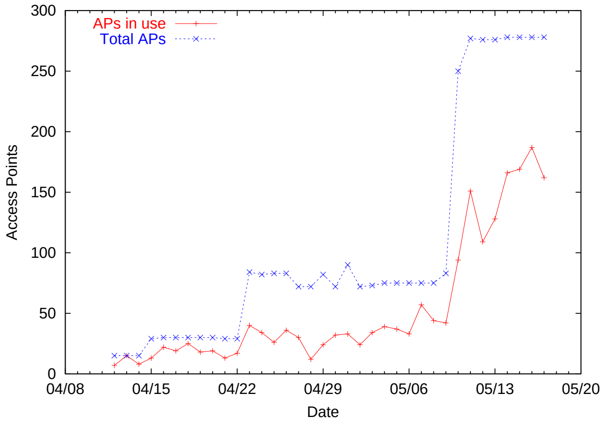

Figure 7: A daily trend of the number of access points with non-zero wireless traffic. The labeled dates are Sundays. Taken from campus SNMP trace.

In Table 3, all the access points with names sudi?-ap are located in Sudikoff. Most of the activity on access point silsby2-ap originates from one user that remained connected throughout the study. This user is generally idle, except for two transfers of 2GB each on 4/13 and 4/18. The reason for the high amount of traffic in Sudikoff

Table 3: Access points with the greatest amount of traffic. Taken from campus SNMP trace.

Rank	Access Point	Where	(MB) per day	Total Users
1	berry1-ap	Library (Computing)	7,264	21
2	berry13-ap	Library (Cafe)	1,597	20
3	berry3-ap	Library (Cafe)	1,536	11
4	berry4-ap	Library (Computing)	1,526	14
5	berry10-ap	Library	1,255	18
6	berry12-ap	Library	966	9
7	baker12-ap	Library (Computing)	893	8
8	berry8-ap	Library	549	11
9	berry5-ap	Library	332	15
10	baker11-ap	Library (Computing)	293	10
11	berry9-ap	Library	225	12
12	berry7-ap	Library	205	6
13	berry2-ap	Library (Cafe)	204	12
14	sudi7-ap	Computer Sci.	166	42
15	cummings6-ap	Engineering	148	33
16	silsby2-ap	Economics	144	12
17	lodge2-ap	Residence	131	6
18	russell-sage3-ap	Residence	115	9
19	murdough2-ap	Eng. Library	113	34
20	sudi11-ap	Computer Sci.	110	34
21	new-hamp3-ap	Residence	99	4
22	cummings21-ap	Engineering	98	20

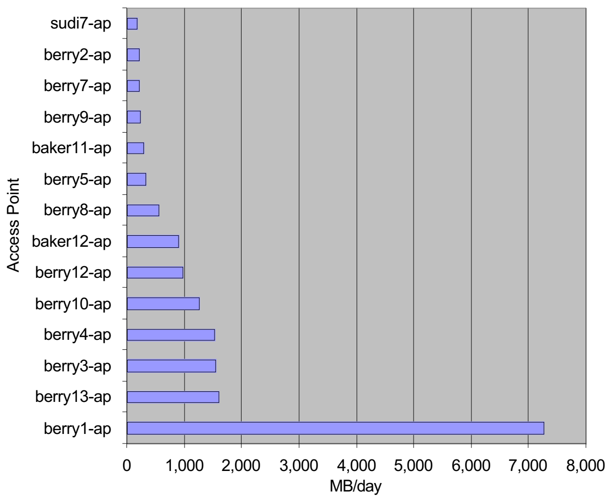

Figure 8: A chart representing access points with greatest amount of traffic.

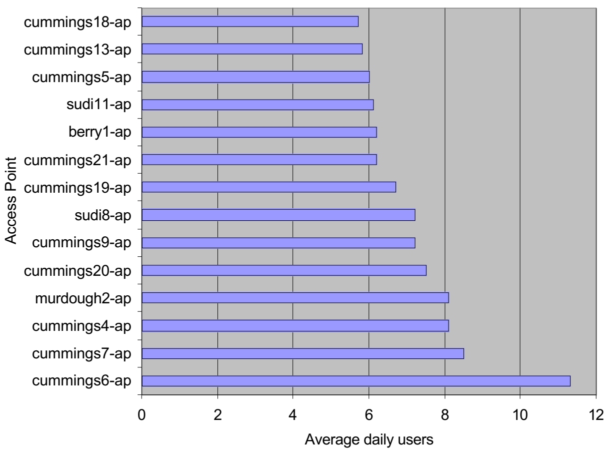

Figure 9: A chart of the busiest access points based on average number of daily users.

and Cummings is primarily the high concentration of network-card owners. As noted in the *Background* section, Sudikoff faculty and graduate students were given 40 Cisco client cards prior to the study.

Although traffic volume indicates which access points have been getting the largest load, to examine user distribution we focus on per capita statistics (see Table 4 and Figure 9). There are about two dozen access points that see more than four users per day. A majority of these access points are in Cummings, which is the Engineering school building. We attribute the high concentration of users in Cummings to several factors. Because Cummings is an academic building with a focus on technology, its community must rely on network connectivity for research. Furthermore, in Spring 2001 there was a course taught at Cummings on Modern Information Technology with a focus on wireless technology. In this course, all the students have been given wireless-enabled laptops and PDAs. Other locales with high per capita statistics are Sudikoff (the Computer Science building) and Berry/Baker (the main campus library).

The seven access points with the greatest amount of traffic per user are located in Berry/Baker. These access points have seen more than 70MB per user per day during the study. A large portion of this traffic may be attributed to Dartmouth Computing Services staff members who have been researching and installing the wireless network.

We examine traffic per building in Tables 5, 6, and Figures 10, 11. We note that a majority of traffic occurs in the libraries (primarily Berry/Baker), while traffic in other building types, such as academic, is spread out among a greater number of buildings. Berry/Baker is a large building, therefore it contains an above-average number of access points. Most buildings have between 3-9 access points; the Berry/Baker combination houses 28 access points. While libraries are the busiest buildings, computing services are second, followed by academic buildings,

Table 4: A distribution of the busiest access points based on the average number of daily users. Taken from campus SNMP trace.

Access Point	Where	Average daily users	Average traffic per user per day (KB)
cummings6-ap	Engineering	11.3	4,606
cummings7-ap	Engineering	8.5	802
cummings4-ap	Engineering	8.1	1,202
murdough2-ap	Eng. Library	8.1	3,416
cummings20-ap	Engineering	7.5	1,088
cummings9-ap	Engineering	7.2	1,961
sudi8-ap	Computer Sci.	7.2	2,887
cummings19-ap	Engineering	6.7	1,228
cummings21-ap	Engineering	6.2	5,027
berry1-ap	Library	6.2	354,208
sudi11-ap	Computer Sci.	6.1	3,301
cummings5-ap	Engineering	6.0	2,468
cummings13-ap	Engineering	5.8	351
cummings18-ap	Engineering	5.7	215
cummings10-ap	Engineering	5.7	1,408
berry10-ap	Library	5.4	71,385
dartmouth2-ap	Languages	5.2	369
hitchcock4-ap	Residence	5.2	8,134
berry13-ap	Library	5.0	81,772
sudi7-ap	Computer Sci.	4.9	4,038
cummings12-ap	Engineering	4.5	1,537
berry4-ap	Library	4.4	111,618
baker6-ap	Library	4.3	373
rauner2-ap	Library	4.2	1,277
collis2-ap	Social Space	4.1	721
cummings17-ap	Engineering	4.1	531

Table 5: A distribution of average daily traffic per building for the busiest buildings. Types: academic, computing, library, administrative, residential, social space, and other. Taken from campus SNMP trace.

Building	Avg. daily traffic (MB)	Type
berry/baker	14,704	Library
berry/baker	13,641	Computing
sudi	3,361	Academic
silsby	2,805	Academic
raven	924	Administrative
cummings	743	Academic
collis	666	Social space
new-hamp	521	Residential
steele	486	Academic
smith	450	Residential
russell-sage	439	Residential
moore	361	Academic
lodge	319	Residential

and residential buildings. The split between library and Computing Services traffic is intangible because certain access points in Berry/Baker saw both types of traffic. Academic buildings house offices for professors and students, plus wireless-enabled classrooms. Because many students (especially, Sophomores, Juniors and Seniors) still have desktops, and many residence halls have yet to be covered with access points the amount of traffic in residential halls has not reached its peak. As incoming students purchase wireless-enabled laptops, this figure will probably rise.

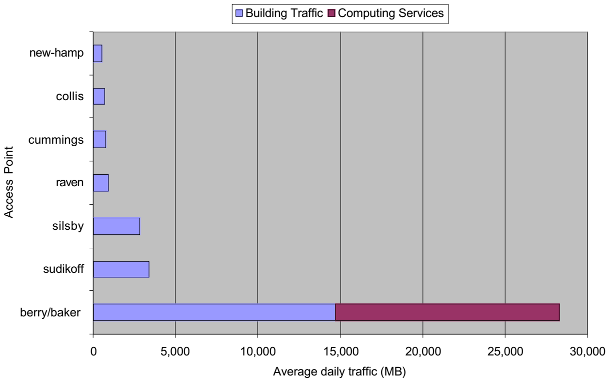

Figure 10: A chart of the busiest buildings.

Table 6: A distribution of average daily traffic per building type for all buildings. Taken from campus SNMP trace.

Type	Traffic (MB)
Library	15,345
Computing	13,641
Academic	9,536
Residential	3,165
Administrative	1,539
Social space	952
Other	99

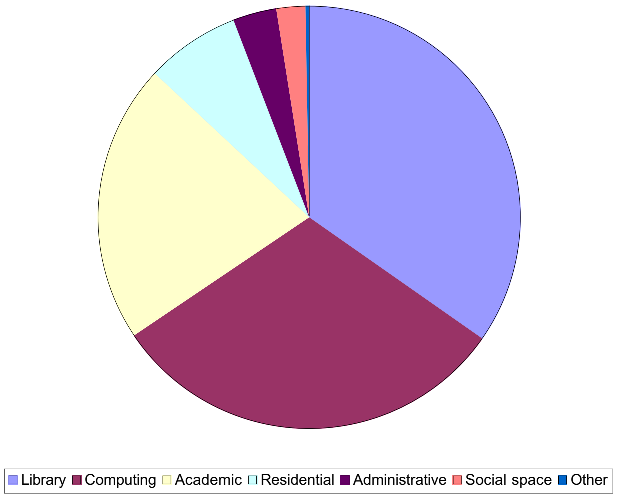

Figure 11: A pie chart of average daily traffic per building type.

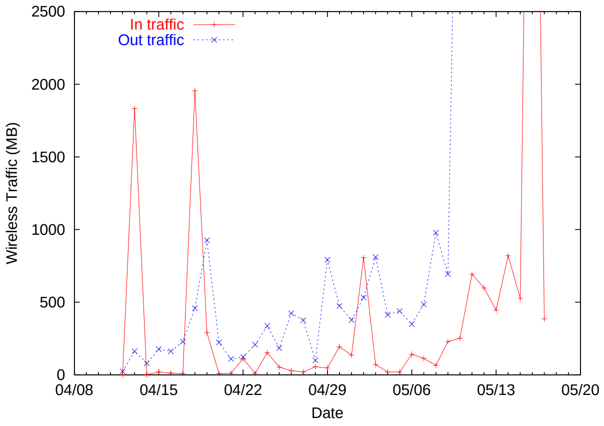

Figure 12: Daily Traffic on the Wireless Network (close-up view of Figure 13). Labeled dates are Sundays. Taken from campus SNMP trace.

6 Network Traffic Characteristics

In this section, we try to answer the following questions:

- How does growth of the network affect traffic characteristics?
- Is there symmetry between inbound and outbound traffic?
- What are the error rates on the network?

For the purpose of this study, traffic is measured at the access point; thus *inbound traffic* is defined as data received by the access point over the wireless network and, conversely, *outbound traffic* is data sent by the access point to wireless users.

The total traffic on the wireless network demonstrates a growth similar to the user base (see Figures 12 and 13). This trend, however, is disrupted by several anomalies. In Figure 12, we notice two dates (04/13 and 04/18) that demonstrate considerable peaks in the amount of inbound traffic. On both occasions, we isolate the same user transferring 2GB of data in about 60 minutes. This is the same traffic that caused a high load on silsby2-ap. We conjecture that this user was backing up their laptop's disk onto a network server. In both figures we note a sharp peak on 05/11. On this date, we added 212 new access points to our study. The most notable addition was Berry/Baker library, which has recently been seeing about 24.7GB of traffic per day.

In Figure 14, we note an exponential rise in the amount of daily traffic as more access points were installed during the study. This fact is partly due to the growing user presence during the study. However, the late addition of high traffic areas, such as the Berry/Baker library access points, magnifies this rise.

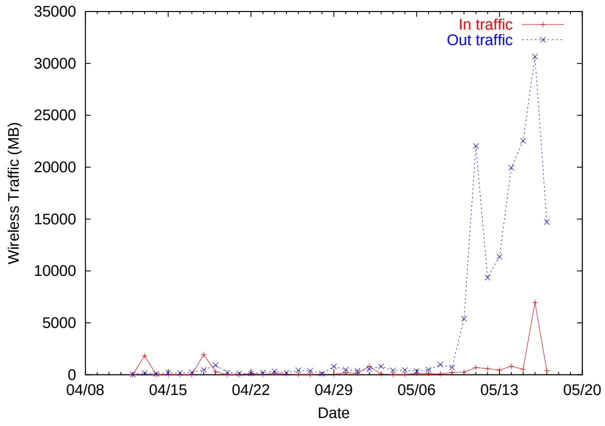

Figure 13: Daily Traffic on the Wireless Network, showing the full range. Labeled dates are Sundays. Taken from campus SNMP trace.

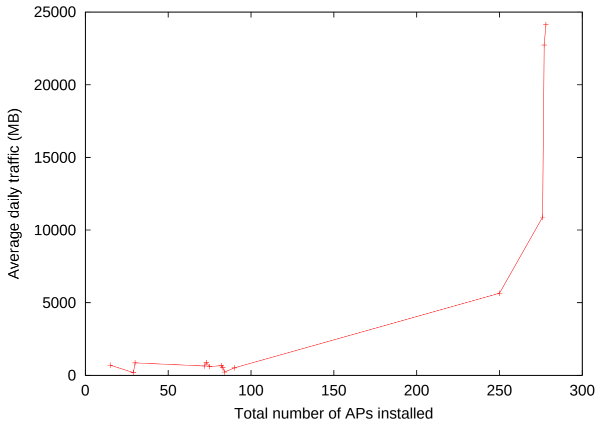

Figure 14: Daily traffic on the Wireless Network based on the number of access points installed on the network. Taken from campus SNMP trace.

6.1 Network symmetry

The total amount of inbound data (16.7GB) is much less than the outbound data (143.3GB). The ratio of inbound to outbound traffic is almost 1:10. This asymmetry contrasts Diane Tang's study where the ratio of inbound to outbound traffic was 3:1 (34GB vs. 12GB) [9]. Although Tang hypothesized that outbound traffic (mainly due to web surfing) would dominate in her study, the nature of her user population (computer scientists performing code development, for example, using shared file servers) may explain the difference. We suspect that the asymmetry in our study was due to a more diverse user community that focused on web surfing, which heavily favors outbound traffic.

6.2 Error rates

We examined error rate as reported in the access-point MIB (and recorded in our SNMP trace). Zeltserman [11] defines these data as the "number of inbound/outbound packets that were discarded due to errors." He states that a high percentage of inbound errors can indicate a problem receiving data (or a line problem on a wired network), while a high percentage of outbound errors indicates a potential hardware problem. The error rates across the network were three orders of magnitude greater on outbound traffic than they were on inbound traffic. Averaged over the period of the study, the error rates were 0.09% (outbound) and 0.00009% (inbound). It is not clear from our data whether these rates are indicative of a problem, although we do note an exceptional imbalance in the amount of outbound errors. If a user is in the range of various access points, the client card may disassociate from one access point and associate with another, causing data from the first to be discarded during the transition. The error rate asymmetry may also be caused by differences between access point and client card antennas.

7 Computer Science Building

Our tcpdump trace provides detailed information about the packets transfered across the network. We mention above that this trace was limited to the Computer Science building, Sudikoff, so this analysis is much more isolated in scope. This trace began a week after the SNMP trace, so in this section we have adjusted the Sudikoff SNMP statistics to coincide with the duration of the tcpdump trace.

By analyzing packet headers we hope to answer the following questions:

- What protocols are being used most often?
- What applications and hardware are being used most frequently?
- Is there a notable asymmetry in traffic?

7.1 General statistics

We can see from Figures 15 and 16 that the data from Sudikoff is more suggestive of a consistent weekly use of the wireless network, largely because the network had been in place for a year and the user population was well established. All six Cisco access points in the building were installed two months before the study began, and 40 wireless client cards were distributed shortly thereafter. The only anomaly is on outbound traffic between 05/10 and 05/11. This surge is attributed to one user performing a large transfer. The consistency in Sudikoff is because the wireless network was in place long before the study began.

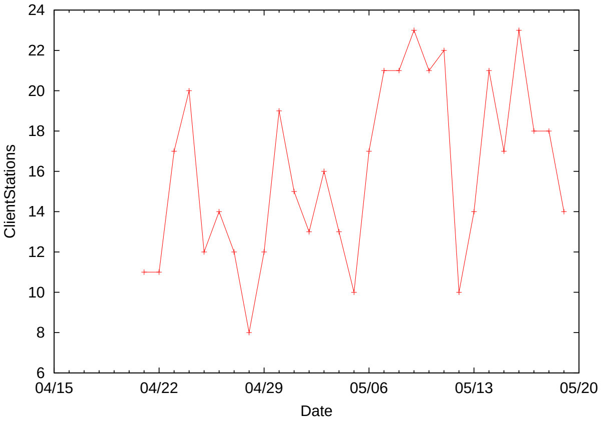

Figure 15: Number of daily wireless users in Sudikoff during the tcpdump trace. Labeled dates are Sundays. Taken from Sudikoff SNMP trace.

During the Sudikoff sniffer trace, we saw 49 MAC addresses. The percentage of mobile users (visiting two or more access points within an hour) is about 20% greater than the campus average (mobile users in Sudikoff account for about 30% of Sudikoff users). This mobility may be due to the relative proximity access points within the building. Many user are constantly in the range of two or more access points, so hand-offs are frequent. Like most of Dartmouth's wireless community, in Sudikoff, outbound traffic heavily exceeds inbound traffic. In this building the ratio of inbound to outbound traffic is less pronounced, 1:7. Again, in the Sudikoff trace, we found a sharp decline in the user presence and traffic during weekends.

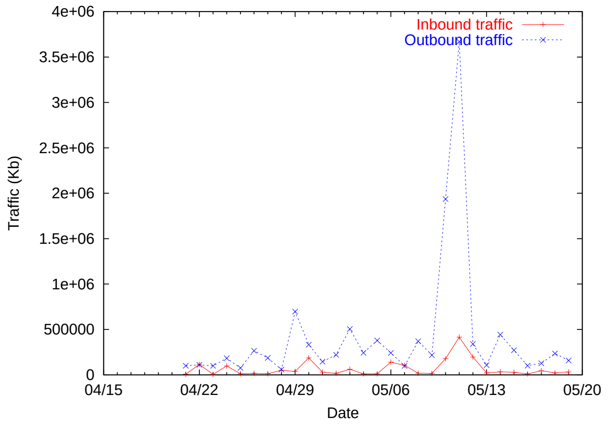

Figure 16: Daily traffic on Sudikoff's wireless network during the tcpdump trace. Labeled dates are Sundays. Taken from Sudikoff SNMP trace.

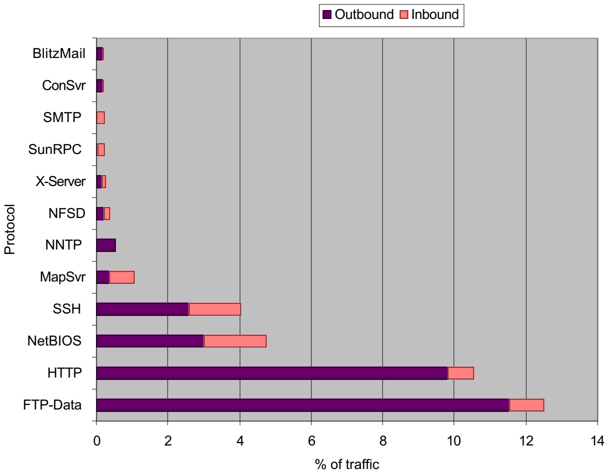

Figure 17: A chart of common protocols by IP bytes transferred.

7.2 Protocol statistics

We measure protocol statistics by examining the port number in the TCP or UDP header (both based on IP). Port numbers often have standard associations (eg., HTTP is port 80). We can see from Table 7 and Figure 17 that most of the bytes moving on the network were from FTP data transfers and HTTP. For both these protocols the majority of traffic was outbound. This statistic suggests that users are web surfing and downloading files. We also see a prevalence in session applications such as SSH, although SSH is often used for secure tunnels, hiding other traffic such as e-mail and file transfer. Also on the list is SMTP, an e-mail protocol. E-mail protocols are generally less data intensive, so account for a smaller portion of the total traffic. The BlitzMail protocol represents Dartmouth's proprietary e-mail system. Many computer science majors and faculty members use their Computer Science e-mail accounts, rather than their Dartmouth BlitzMail accounts. Thus, BlitzMail traffic in Sudikoff probably represents a smaller percentage than it would elsewhere on campus. With these protocols, the large discrepancy in the direction of traffic indicates that the wireless users are e-mail clients.

The NNTP protocol stands for Network News Transfer Protocol, and is used for Usenet newsgroups. We note with NNTP that a majority of the traffic is outbound; the server is sending news to the client. NetBIOS is primarily used by clients to access LAN resources, such as Microsoft Windows' File and Printer sharing and to identify users. The Address Resolution Protocol (ARP) is used to convert 32-bit IP addresses into 48-bit Ethernet addresses. We also see certain protocols that indicate UNIX-based traffic (i.e., NFSD, X-Server, SunRPC). NFSD is the Network File System Daemon. This UNIX-based program handles requests for file system operations. The X-server protocol allows client to use a graphical interface on a foreign server.

Table 7: Distribution of common protocols by IP bytes transferred in Dartmouth's Computer Science Building. We also show the division between inbound and outbound bytes. Taken from Sudikoff tcpdump trace.

Protocol	%	Out (%)	In(%)
FTP-Data	12.3	92.1	7.9
HTTP	10.53	92.8	7.2
NetBIOS	4.68	63.3	36.7
SSH	3.99	63.9	36.1
MapSVR	1.05	32.0	68.0
NNTP	0.51	98.2	1.8
NFSD	0.37	51.1	48.9
X-Server	0.26	59.0	41.0
ARP	0.26	96.9	3.1
SunRPC	0.24	6.5	93.5
SMTP	0.22	2.5	97.5
ConSVR	0.19	95.4	4.6
BlitzMail	0.19	92.6	7.4

An examination of protocols based on bytes transfered skews the results to data intensive protocols like FTP. We found that TCP traffic accounts for 97% of all IP traffic; for TCP protocols, we can also count the number of connections. This metric provides a more accurate depiction of the connections created by the user, without biasing data intensive connections. In Table 8, we see that the Sun remote-procedure-call (SunRPC) protocol represents an overwhelming majority of connections. Many UNIX-based services use Sun Microsystems' RPCs, such as NFS (Network File System) and portmap, which we cannot distinguish in our trace. These services often run in the background and perform operations without a user's explicit knowledge. In a building with a large number of UNIX-based systems, such as Sudikoff, it is not surprising to see such a high concentration of RPCs. This prevalence of Sun RPCs reflects the general type of services being used (and the type of operating system), but they do not provide specific information about these services.

All SunRPC connections are inbound, indicating that the transfer is initiated by the wireless machine. Other protocols with a majority of inbound connections include web surfing, DBS, FTP, SUP, SOCKS, and NetBIOS. The Dartmouth Bulletin System (DBS) serves to inform users about events occurring on campus. The Software Update Protocol (SUP) is used by UNIX-based client machines to request file updates from a server. SOCKS is a network proxy protocol. A large number of inbound FTP connections suggests that users are requesting file downloads.

When examining the number of connections we note an emphasis on the amount of web surfing that occurs on the wireless network (i.e.,

Table 8: Distribution of common protocols by connections in Dartmouth's Computer Science Building, measured as a percentage of total TCP connections. Each connection is an information transfer, signaled by a SYN==1 and ACK==0 flag in the TCP packet header. We also show the division between inbound connections(initiated by mobile node) and outbound connections (initiated by non-mobile node). Taken from Sudikoff tcpdump trace.

Protocol	%	Out (%)	In (%)
SunRPC	0.71	7.9	92.1
HTTP	0.29	5.3	94.7
FTP	0.15	23.0	77.0
DBS	0.15	3.0	97.0
NetBIOS	0.14	8.8	91.2
SOCKS	0.13	1.2	98.8
SUP	0.12	0	100
SNMP-trap	0.12	5.2	94.8
SNMP	0.12	5.1	94.9
SSH	0.11	4.2	95.8

HTTP). This dominance is partly due to the nature of HTTP, which tends to have many brief connections. Services such as FTP and SSH also figure prominently.

8 Conclusion

Currently, Dartmouth's wireless network is growing. We document the increases in both the user base and the amount of traffic. Current traffic statistics indicate that the load on individual access points has not reached a critical stage, even in the areas of highest traffic.

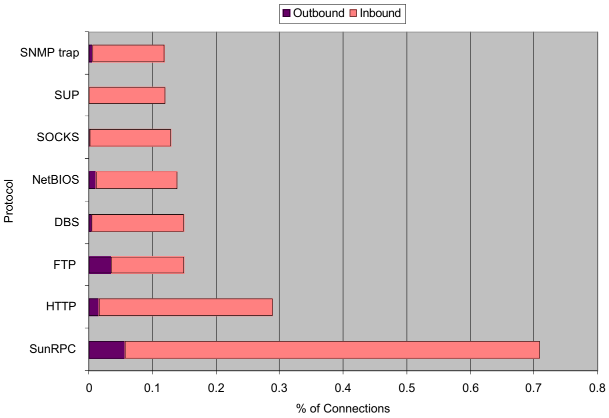

Figure 18: A chart of common protocols by connections.

The location of an access point is the greatest factor in the amount of traffic it sees. Access points located in wireless-user dense areas, such as Berry/Baker library, Cummings and Sudikoff, have seen the most traffic. Also, access points are at different stages of growth. Certain locations (such as Sudikoff) have reached more consistent use, while most places are seeing growing numbers of users.

Libraries and Computing Services had the greatest amount of traffic, followed by academic buildings, then residential halls and administrative buildings.

The amount of outbound traffic (sent by access points) greatly surpasses inbound traffic (sent to access points) measured in bytes. This disparity is due to the high amount of web surfing and file downloads, which mainly generate outbound traffic. For inbound traffic, a couple of sessions account for a majority of the traffic. These sessions were probably hard-disk backups.

A protocol analysis indicated web surfing as the main wireless activity. Also, many common protocols, including e-mail, file-transfers, and remote sessions, accounted for much of the remainder of traffic.

In the fall, as more incoming students will purchase wireless-enabled computers, the amount of users and traffic will increase dramatically. We speculate that more traffic will be seen in dorm rooms and student spaces. The network should still grow over the school year as upper-class students continue to switch to wireless technology; this growth will be much more subtle that what we witnessed in our study. Next academic year, the network usage characteristics will probably near a steady-state. Over the next few years, with each incoming freshman class, the number of wireless users should rise.

9 Future Work

This study provides some insight into the deployment of a wireless network. The majority of access points were not installed until the last few weeks of the study. Also, the user base for this study has been fairly limited. Next fall, incoming students will have the option to purchase wireless-enabled laptops. In the future, we hope to collect a new trace, where the network is in steady use, and the number of users is greater.

We also wanted to install network sniffers at different locations on campus (such as, cafeterias, dorm rooms, and classrooms) to get a diverse perspective of the types of wireless users. Unfortunately, many of these locations were not active until late in our study.

We are interested in a comparison study between the wireless and wired network. An analysis of the wired network would provide control data for the wireless study.

10 Acknowledgments

I would like to thank Professor David Kotz, my advisor, for providing innumerable resources and help. I worked closely with Kobby Essien, who helped set-up the traces and was a constant source for reference material. I also thank Steve Campbell, Punch Taylor, and Jim Baker, my contacts at Dartmouth's Computing Service, which administers the wireless system. Arne Grimstrup, with the D'Agents group, has also provided much time and guidance. This project is supported by the Cisco University Research Program.

References

[1] *Wireless Computer Networking coming to Dartmouth.* Dartmouth College News Service. `http://www.dartmouth.edu/~news/releases/feb01/wireless.html`, (02/28/01)

[2] *Data Sheet: Cisco Aironet 350 Series Access Points.* `http://www.cisco.com/warp/public/cc/pd/witc/ao350ap/prodlit/a350a\_ds.htm`, (05/24/01)

[3] Eckhardt, D. and Steenkiste, P. *Measurements and Analysis of the Error Characteristics of an In-Building Wireless Network.* SIGCOMM '96. August 1996.

[4] Hills, A. and Johnson, D. *A Wireless Data Network Infrastructure at Carnegie Mellon University.* IEEE Personal Communications. February 1996. p. 56-63.

[5] Hills, A. *Wireless Andrew: Bringing mobile computing to a university community of 10 000.* IEEE Spectrum. June 1999. p.49-53.

[6] O'Hara, B. and Petrick, A. *IEEE 802.11 Handbook: A Designer's Companion.* Standards Information Network IEEE Press. 1999

[7] Messier, A., Robinson, J. and Pahlavan, K. *Performance Monitoring of a Wireless Campus Area Network.* Proceedings of the 22nd IEEE Conference on Local Computer Networks (LCN '97). 1997. p. 232-238

[8] Tang, D. and Baker M. *Analysis of a Metropolitan-Area Wireless Network.* Proceedings of the Fifth Annual ACM/IEEE International Conference on Mobile Computing and Networking. August, 1999. p. 13-23.

[9] Tang, D. and Baker M. *Analysis of a Local-Area Wireless Network.* Proceedings of the Sixth Annual ACM/IEEE International Conference on Mobile Computing and Networking. August, 2000.

[10] Taylor, W. *Re: WIRELESS: background info.* E-mail to author. 11 May 2001.

[11] Zeltserman, D. *A Practical Guide to SNMPv3 and Network Management.* Prentice-Hall, Inc. 1999.

A Privacy Statement

Dartmouth is installing a wireless computer network for students, faculty and staff to connect to the network from anywhere on campus. We are studying the traffic on the wireless network to understand how this network is used. To ensure user privacy, this study will maintain user anonymity and examine only the header information of data passing through the network. Header information specifies the type of information that is being transferred over the network. This specifically excludes the contents of the data, such as, usernames, passwords, filenames and files.

We will never release any data that will identify any user or their specific data.

FAQ

1. If I am connected to the wireless network:
   (a) Do you know who I am?
       - No. We are only monitoring hardware addresses for statistical purposes, never will a portable device be linked to its user.

   (b) Do you know where I am?
       - No. We will log information about which devices (referenced by hardware addresses) are connected to the network at particular times. This information will be used to track the movement of devices. More concretely, it will be known that a particular device was connected to a particular access point in the network at a particular time but the owner of that device will be unknown.

   (c) Will there be a log containing the files I have downloaded?
       - No. The filenames and content of the information you transfer on the network is referred to as the payload. Our devices never examine the payload.

   (d) If I log onto Blitzmail, or other account will my password be stored?

    - No. Passwords and usernames are not monitored or saved.

   (e) I use the internet make credit card purchases, will you see my credit card number?

    - No. Information such as credit card numbers will not be monitored.

   (f) Do you know where I’ve been browsing?

    - The study may log hostnames for statistical purposes. However, sites will never be associated with users. We will monitor the length of the time a user has been using the wireless network for any activity, including browsing. However, as mentioned the user will remain totally anonymous.

   (g) Will these devices hinder the speed of the network?

    - Minimally. We will create a minimal amount of traffic (at most 0.3% of the network bandwidth) to transfer statistical data. The wireless network will never be affected.

2. How are devices being monitored?

   - We will measure statistics about the traffic on all of the wireless access points, to measure how many devices are connected, how much data is transferred, and the like. We will monitor selected access points in more detail, using a network monitoring device installed near the access point. Those monitors will record all of the network ”headers”, but none of the ”payload”. That means these devices will record what types of application (e.g., web browser, email) is being used, but not record any of the pages downloaded, mail transferred, usernames, passwords, or other private information.
   - Also, whenever a user connects to or disconnects from the network we will make a record of the event in a log. The entry will have the user’s hardware address so the user still will still remain anonymous.