

What Where Wi: An Analysis of Millions of Wi-Fi Access Points

Kipp Jones, Ling Liu

Division of Computer Science and Systems
Georgia Institute of Technology, Atlanta, GA, USA
{kippster, lingliu}@cc.gatech.edu

Abstract— With the growing demand for wireless Internet access and increasing maturity of IEEE 802.11 technologies, wireless networks have sprung up by the millions throughout the world as a popular means for Internet access at homes, in offices and in public areas, such as airports, cafés and coffee shops. An increasingly popular use of IEEE 802.11 networking equipment is to provide wireless ‘hotspots’ as the wireless access points to the Internet. These wireless access points, commonly referred to as WAPs or simply APs, are installed and managed by individuals and businesses in an unregulated manner – allowing anyone to install and operate one of these radio devices using unlicensed radio spectrum. This has allowed literally millions of these APs to become available and ‘visible’ to any interested party who happens to be within range of the radio waves emitted from the device. As the density of these APs increases, these ‘beacons’ can be put into multiple uses. From home networking to wireless positioning to mesh networks, there are more alternative ways for connecting wirelessly as newer, longer-range technologies come to market.

This paper reports an initial study that examines a database of over 5 million wireless access points collected through wardriving by Skyhook Wireless. By performing the analytical study of this data and the information revealed by this data, including the default naming behavior, movement of access points over time, and density of access points, we found that the AP data, coupled with location information, can provide a fertile ground for understanding the “What, Where and Why” of Wi-Fi access points. More importantly, the analysis and mining of this vast and growing collection of AP data can yield important technological, social and economical results.

Keywords: Wi-Fi; Wireless Networks; Access Points; Location Based Services

I. INTRODUCTION

Wireless networks have become increasingly popular in recent years as a means of providing Internet access and the ‘last meter’ connectivity within homes and businesses. These networks allow limited roaming within a designated area such as a home or an office while maintaining connectivity to the Internet. This use of ‘tails’ through the Wi-Fi connectivity to the wired network is the dominant model of wireless Internet access today. We have seen and continue to see new methods of using Wi-Fi emerge such as mobile Voice over IP (VoIP), location based gaming and other services based on a growing array of applications and portable devices [1].

Commercial hotspots – Wi-Fi enabled zones – have sprouted in many places. These access points; in the ubiquitous coffee shop, in airports, in bookstores, are currently providing Internet access to the public. Many of these APs require subscription and payment for the service, while others provide Internet access as some benefit to the public. According to Broadband Wireless Exchange¹, the top hotspot providers now have over 40,000 hotspots worldwide.

The wave of municipal wireless networks like those being rolled out in cities around the country offer another motivation for this study. From high profile efforts in Philadelphia and San Francisco to less publicized effort in places like Moorhead, Minnesota, public enterprise has become extremely interested in the value that a city-wide wireless infrastructure could provide both for the efficiency and the capabilities of the public servants, as well as for the universal access that such an infrastructure could provide to help bridge the digital divide. With the promise of freeing the municipal workers from the tether, providing improved efficiencies in operations, and creating a path for universal access, millions of dollars are being invested in building out wireless connectivity.

Some look at the sea of access points and find commercial value. Companies such as FON² in Spain and WiFiTastic³ in San Jose are out to help individuals monetize their Wi-Fi access points by providing authentication and billing infrastructure that turns ordinary access points into commercial endeavors. In fact, a cottage industry has sprung up dedicated to providing aftermarket modifications to standard access points⁴ in order to facilitate the use for commercial or group access.

Others such as Place Lab [10] and UCSD [2] have shown that the access points need not provide active connectivity to provide value. By using the signal and identity of the myriad access points, value can be obtained by providing services such as positioning information to stationary and mobile users. And unlike systems such as GPS where the location of the beacons are known, the location and signal propagation of these access points can be learned over time

¹ List of top hot spot providers: http://www.bbwxchange.com/top10_wi-fi_hotspot_operators.asp

² FON web site: <http://www.fon.com/>

³ WiFiTastic web site: <http://www.wifitastic.com/>

⁴ Companies such as Sveasoft (<http://www.sveasoft.com/>) provide firmware upgrades for standard access points.

and need not be complete to provide adequate location information.

In this study we analyze a large collection of geolocated access points. The logs that were gathered during the acquisition of the access point information are also examined. We discover a number of ways to analyze and utilize this data beyond its initial purpose. The method and results of this study are explained in detail in the following sections.

Developing the understanding of how these wireless networks are being used in different geographical regions can guide not only on how to carry out more efficient hotspot deployments and network design [10], but also on how to proceed and leverage the existing investment [4]. Some most representative questions include:

- How many access points are present and what are their characteristics?
- How to conduct a taxonomic analysis of network properties, such as how many open networks versus closed ones are there? Which vendor is selling the most of APs?
- What types of wireless networks can be designed for legitimate public use (open access) and what their performance would be?
- What should be considered as non-legitimate use of the network and how can we prevent the networks from misuse or abuse.
- How to assess the coverage of a particular network?
- How to assess the saturation of the spectrum?

In this paper, we report our initial analysis of over 5 millions of geolocated access points and the scanning logs associated with these access points, addressing some of the questions listed above. Our statistical analysis shows that there is significant information contained in such a large collection of APs, and the AP data can be linked with other information sources to create additional value towards developing the understanding of the many characteristics of wireless networks in general. We conjecture that the knowledge of the current infrastructure and the improvement of our network models will help increase the effective use of the networks.

The rest of the paper proceeds as follows. We first give an overview of the data set and the process that was used to gather this data. Then we describe the analytical results of the study, focusing on a taxonomic analysis of a selection of network properties, including how many access points are present in a given geographical region and what are their characteristics, how many open networks are there versus closed ones, which vendor is selling the most of APs. We also provide a short discussion on the related work before concluding the paper. The paper ends with a discussion of the areas for further consideration and research.

II. APPROACH

To conduct this study, we acquired the rights to analyze a data set provided by Skyhook Wireless. This data set is collected by a fleet of drivers that systematically drive urban areas to scan for 802.11 Wi-Fi access points.

A version of this paper has been submitted to IEEE Portable 2007.

The data set was gathered by Skyhook Wireless during the time between April 2004 and December 2005. This data corresponds to the systematic scanning in some 75 cities throughout the United States as shown in Fig. 1.



Figure 1. Skyhook Wireless⁶ coverage areas. Cities in red are in progress while cities in blue have been completed.

The process of gathering this data is often referred to as ‘wardriving’, a term derived from wardialing which was popularized by the movie WarGames⁷. Wardriving is the act of locating wireless access points through the use of wireless scanning equipment within a moving vehicle.

Traditional wardriving is ad-hoc and resultant datasets are composed of numerous passes by many drivers. In these instances, the decision of which routes to drive is often made with respect to the types of roads – major roads are driven more often than minor roads. When this data is used to calculate the location of the access point, it is common to see the ‘weight’ of the major roads unduly influence the derived location of the access point. This effect is referred to as ‘arterial bias’ and is represented in Fig. 2.

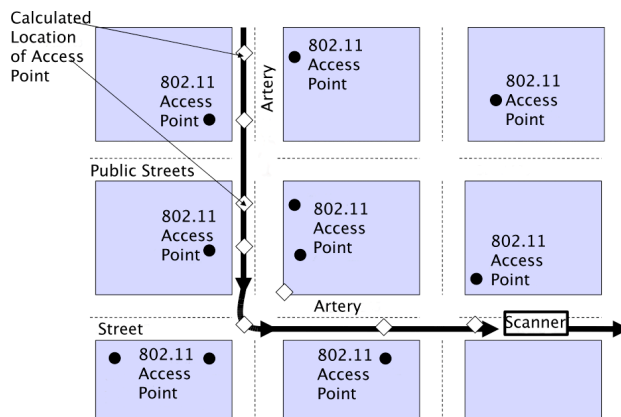


Figure 2. The effect of arterial bias on location estimation.

However, routes for Skyhook Wireless drivers are determined such that this arterial bias is virtually eliminated, increasing the accuracy of the resultant location estimation.

Efficient routing for obtaining the data can be modeled as a Chinese Postman problem. The Chinese Postman problem is defined as finding the shortest route in a network that

⁶ Data provided by Skyhook Wireless, map retrieved January 2006. <http://www.skyhookwireless.com>

⁷ For further information on wardriving see Wikipedia, <http://en.wikipedia.org/wiki/Wardriving>

traverses each edge. In this case, the network is the road system and we desire to find the most efficient route that traverses each segment of the roadway. While this problem has been shown to be NP-complete [14], Fredrickson [6] analyzes several approximation algorithms that provide worst-case bounds as low as $\frac{3}{2}$. By ensuring that readings are gathered from as many angles as possible, a more accurate estimation of the source can be calculated. This effect is illustrated in Fig. 3.

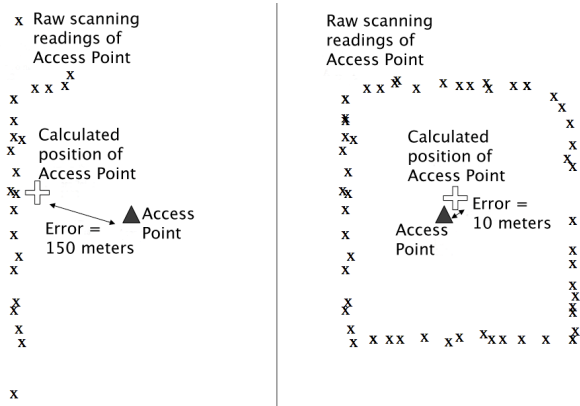


Figure 3. Reducing arterial bias by traversing all roads.

The data is logged using software called Bertha, a proprietary scanning software package from Skyhook Wireless. The software runs on custom configured mobile devices connected to a standard GPS device via serial or Bluetooth communications. During scanning, no connections are established to the access points.

The software utilizes commercial access points to automate the upload of scanned data to a central server. Upon upload, the scanning data is processed to produce the correlation of each access point with its GPS location and signal strength information. The specifics of the algorithm for this calculation is beyond the scope of this paper; however, a number of methods and algorithms have been developed ranging from simple triangulation of signals to more complex hierarchical Bayesian sensor models [12].

The system measures the signal strength and gathers access point information from the radio signal produced by each AP. For each access point, this includes multiple records that include its name or Service Set Identifier (SSID), the Media Access Control (MAC) address and the timestamp when the AP was scanned. Unfortunately, the dataset does not include channel or security setting information that would be useful for a number of studies.

Concurrent with the logging of this data, the geolocation in the form of latitude, longitude, number of satellites, and error is captured using GPS.

The system also tracks the ‘movement’ or change in calculated location for each access point. For purposes of this study, the resulting processed dataset as well as the movement dataset were analyzed at two different points in time. This data was stored in a standard relational database of MySQL version 5.0.19 using the MyISAM engine. Table 1 describes the available tables and the relevant fields within each table.

A version of this paper has been submitted to IEEE Portable 2007.

Table 1. Data tables and relevant fields.

Table Name	Description	Fields
CentralAP	Unique access points and derived location	MAC, SSID, Latitude, Longitude, Date, Scanner Key
ChangeAP	Location adjustments	MAC, Date, Previous Latitude / Longitude, New Latitude / Longitude, Distance
RawScanningLogs	AP scan records	MAC, Date, Latitude, Longitude, Signal Strength, Scanner Key
ScannerGpsLogs	GPS logs	Scanner Key, Latitude, Longitude, Satellites, Date

III. RESULTS

The results from this study include a set of statistical measures as well as a set of tools, which we will use to continue the analysis and mining of the growing amount of AP data collected by Skyhook Wireless on the subject. Results were calculated using a number of methods including database queries, java programs, spreadsheets, GIS tools, and mash-ups with Google Maps.

During the analysis two different datasets were used based on snapshots of the data at a particular time. The first dataset included 3,571,212 access points (Dataset 1) while the second included 5,660,428 access points (Dataset 2).

It is estimated that there are in over 40 million access points deployed in the United States [ref?]. Assuming this estimate is correct, the data sets would represent approximately 9% and 14% of the deployed access points respectively.

Fig. 4 provides a visualization of over 100,000 scanned access points in the North Atlanta region, nearly blanketing certain areas of the city. This image provides a visceral ‘sense’ of how thoroughly wireless networks have penetrated our urban and suburban society.

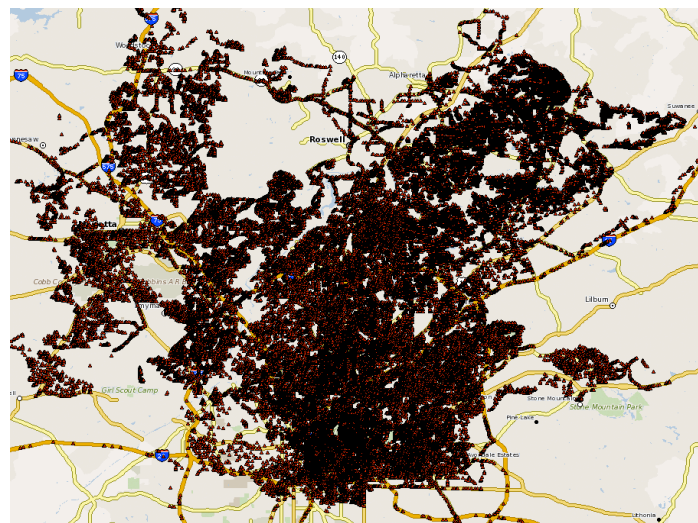


Figure 4. All access points in North Atlanta region.

A. Access Point Naming and Default Settings

Access point naming analysis aims at understanding and identifying the different behaviors in the method that users name their access points. Through examining both datasets, we observe that approximately 44% of the studied access points retain their default factory names. Fig. 5 below is an example of the naming statistics showing the top 25 names by frequency.

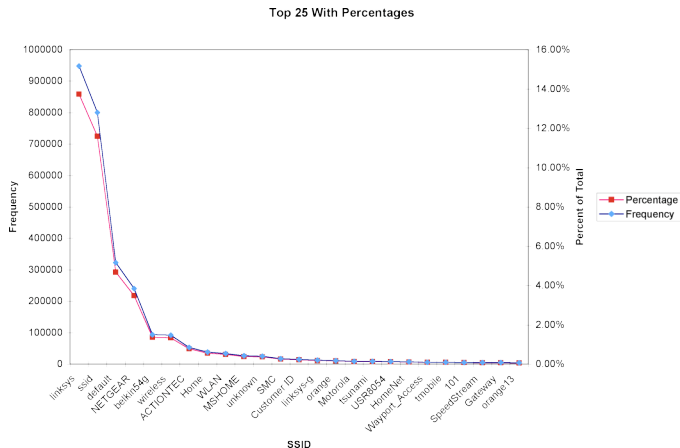


Figure 5. Top SSID frequency and percentage of total

Fig. 6 depicts the same area as shown in Fig. 4 but includes only those access points which retained the default ‘Linksys’ SSID configuration. As can be noted, there is a reduction in total number of access points, but there remains a substantial installed base of access points with default configurations.

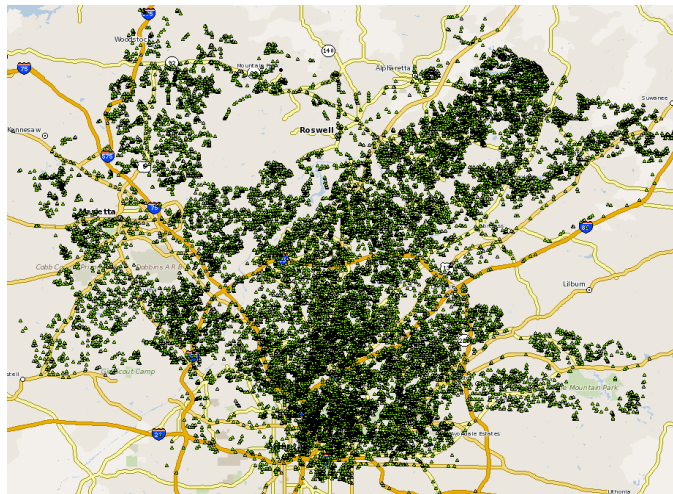


Figure 6. Access points in Atlanta with Linksys default settings.

The use of default settings has been noted in other studies such as those produced by Bychkovsky et al in [4] where nearly 75% of the APs had default security settings (not secured) while nearly 50% had default SSID names. Our current dataset does not provide specific information on the security settings of the access points scanned, but we can infer the possible default security settings based on default SSID names. If we were to use the same ratio as previous studies have uncovered, it may indicate that as many as 66% of the access points remain unsecured. This information

A version of this paper has been submitted to IEEE Portable 2007.

indicates that manufacturers would do well to ensure that sufficient security and privacy is provided as a default configuration.

B. Manufacturer Information

Manufacturer information regarding the deployed APs was analyzed using two measures. First, the default naming characteristics provides a baseline measure that indicates the relative number of access points that maintain their default settings. Second, an analysis of the MAC address was performed to correlate these assigned addresses with the manufacturer information. The results, shown in Fig. 7, indicate the market leadership of Linksys (a Division of Cisco Systems, Inc.) in the Wi-Fi marketplace. Of the 5,660,428 access points in Dataset 2, we found 38% of them belonged to Linksys. In addition, the top 20 manufacturers represent 96% of the access points in the database.

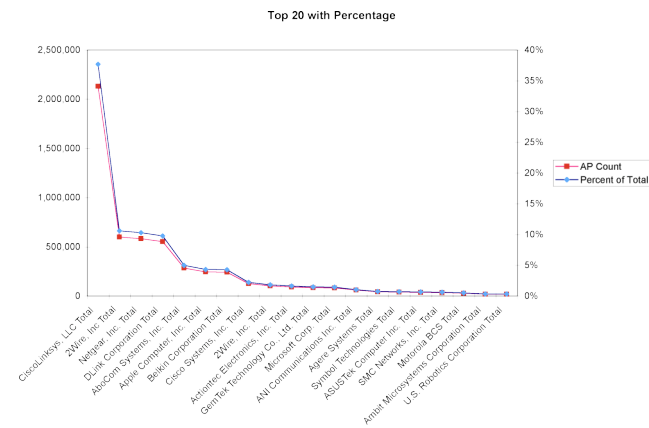


Figure 7. Manufacturer based on MAC address.

Note that Fig. 5 and Fig. 7 are closely related as would be expected based on the manufacturer and the default SSIDs for the access points.

C. Access Point Stability

One unforeseen use for which the growing Wi-Fi infrastructure is being used is to provide a location service similar to GPS. Skyhook Wireless provides such a service. Based on the location of the access points a user with a Wi-Fi enabled device can determine their location via the Wireless Positioning Service (WPS).

One key measure that is important for the performance guarantee of the WPS system in general is the stability of the access points. In the cases where access points were actively being moved (i.e., person moving across town or elsewhere), the WPS system could be at risk of providing faulty or degraded information.

In order to measure the motion behavior of APs, the Skyhook system captures the old and new location calculations for each access point over time. This data includes the date, the previous latitude/longitude, and the new latitude/longitude. This motion data was analyzed to determine the relative frequency of movement of the access points. Of the 3,571,212 APs in Dataset 1 that were measured, only 864 (.02%) had moved further than 1 kilometer. This observation leads to several conclusions:

1. There is insufficient longitudinal data to provide a full and accurate study of the motion of APs over time. This is due to the fact that re-scans of many of the areas had not been accomplished at the time of this study.
2. Of those that had been re-scanned, they do not move frequently.
3. Most of the motion measured was due to refinement of the position of the AP caused by the introduction of additional scan data or the improvement of the location calculation algorithm.

As more data is gathered and areas of cities are re-scanned, this access point stability study will become more meaningful from a statistical perspective. The initial findings are encouraging. In addition, systems should also rely on signals from more than one access point to reduce the effect of access point motion.

D. Access Point Density

In this section we explore the density of access points, aiming at better understanding of the current state of wireless deployment and exploring what the future wireless landscape will look like.

In particular, the density of access points can play an important role in how or whether individual access points can cooperate with each other. In certain scenarios, it has been suggested that access points, even those owned by individuals, could cooperate in a mesh network in which they use each other as a blanket of connectivity rather than relying solely on the individual Internet connections for each and every access point. For this to work there must be sufficient density of access points such that the majority of APs are able to not just ‘see’ several other access points, but must have sufficient signal strength to perform reliable connectivity between them [5].

It should be noted that there is a difference between being able to detect the radio signal from an access point and the ability to actually perform the network connection necessary to transport data. Certain services may be performed without the need for actual connectivity and thus can deal with a lower density of access points. For example, the Wireless Positioning Service does not require connectivity, providing a nominal coverage of 200 meters radius for each access point⁸.

Of course the signal propagation depends on a large variety of variables from building material, antennas, power, and other obstacles that may attenuate the signal [9]. In addition, newer wireless standards such as 802.11n and WiMax (802.16) have the ability to travel much farther and still maintain their ability to provide connectivity. Nevertheless, we focus on standard 802.11b/g (standard Wi-Fi) due to the proliferation of already deployed devices⁹.

We assume that the nominal range for standard 802.11b/g is 100 meters¹⁰. For simplicity, we use a basic calculation, and assume uniform circular coverage for each AP, noting that each square kilometer of area would require approximately 33 access points. This is quite different than the purpose built networks being constructed for metro-scale Wi-Fi networks which utilize specialized radio and antenna equipment to reduce the hardware requirements.

Using the database linked with Google Maps we can quickly determine the access point density of any particular area. Table 1 provides a sample of these density measurements based on rough bounding boxes of a given area.

Table 2. Access Point Density

Region	Area (km ²)	Access Points	Density (APs/km ²)
U.S.	9,166,600	5,615,451	0.6
Las Vegas	240	26,069	109
Kansas City	270	29,438	109
Atlanta	460	65,364	142
San Francisco	213	69,502	326
Seattle	165	64,923	395
Boston	225	164,072	729
Manhattan	105	194,651	1,854

As observed in the density statistic analysis, major metropolitan areas are well above the 33 AP/km² that we noted above. This is especially true as you focus on high-density population areas, with Manhattan, for example, having a density of over 1,800 access points per square kilometer.

Given these sample points, it would appear that there are several areas in the U.S. that would be able to support new models of use for the already deployed access points. Of course there are many technical, security, and business issues that stand in the way of these alternate connection models, but it is interesting to note that deployment of infrastructure is not the obstacle.

E. AP Density and Demographics

Once the access points are geographically mapped they can be combined with demographic information. For example, we can examine the density of access points within census tracts and compare this with the population or household income. This may help city planners in understanding and planning deployment of municipal wireless networks.

Through the use of MapServer¹¹ combined with the TIGER (Topologically Integrated Geographic Encoding and Referencing system) data provided by the US Census

⁸ Personal communication with Farshid Alizadeh, VP of Technology Development and Research for Skyhook Wireless.

⁹ For more on the 802 series of standards, see: <http://grouper.ieee.org/groups/802/>

A version of this paper has been submitted to IEEE Portable 2007.

¹⁰ Based on numerous sources including:

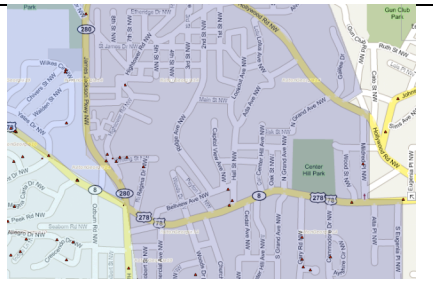
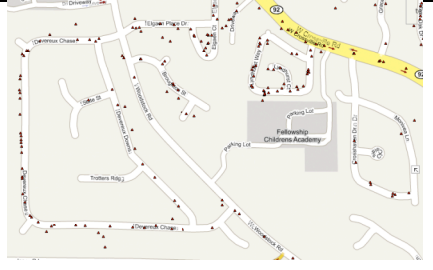
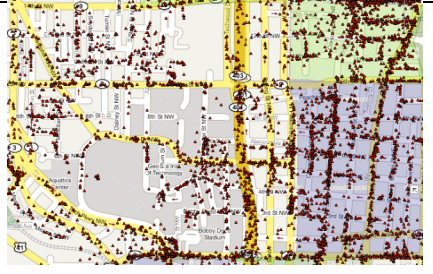
<http://www.designnews.com/article/CA272261.html>

¹¹ Open source GIS tools <http://ka-map.maptools.org/> and <http://mapserver.gis.umn.edu/>

Bureau¹² we are able to visualize the density of access points in particular regions.

Table 3 shows three areas in the Atlanta region examined for understanding of the correlation between AP density and demographics. For each area, a representative census tract was examined and included in the table. Red triangles in the images depict the estimated location of specific access points. Population and household income relate to specific sample census tracts within the area depicted and do not relate to the total area shown.

Table 3. Access Points and Demographics

Region	Grove Park	
APs	46	
Sample Tract	86.01	
Population	5,811	
Household Income¹³	\$18,051	
Region	Roswell	
APs	176	
Sample Tract	114.07	
Population	9,456	
Household Income	\$79,364	
Region	Midtown	
APs	2780	
Sample Tract	12	
Population	4,216	
Household Income	\$40,654	

As one would expect, we see a higher density of access points in areas of higher household income, presumably the constituents have more money available to spend on computing equipment and services.

We also note the heavy density in areas such as Midtown Atlanta, near the Georgia Tech campus, CNN, and a couple of other organizations. This is likely due to a number of factors including the age and proclivity to use technology as well as the density of businesses in the region.

It is important to note that these are preliminary results. There are many other demographic parameters that could be studied relative to the location and density of access points to better understand the adoption of wireless technologies and to guide the design of next generation of wireless networks..

F. Security and Privacy

There has been much written about the security risks [1][3][8][13] associated with wireless networking. Research to date has focused on two primary topics: protecting access to these wireless networks and maintaining the privacy of the users and the data on these networks. There is a potential for information leakage that is accessible even for those not interested in directly accessing the network. These wireless access points emit a certain amount of information that can be combined with the physical location and potentially used for other purposes.

However, exploration of this topic did not yield any particularly critical security issue derived from the data. Perhaps the ability to locate unsecured wireless access points based on the default settings of the SSID could lead one to more easily locate these access points, but as the analyzed data does not contain the security settings for the access points¹⁵, this would seem a more difficult approach than simply wardriving and locating open networks.

The density information could provide ‘hints’ to those with nefarious motives to concentrate in particular areas; but, again, this seems a stretch and would appear to be less valuable compared to the work it would require to obtain the data.

Another opportunity would rely on the exposure of exploits for a particular HotSpot provider. In this scenario, knowing the location of the providers’ access points could help leverage a known exploit. A database such as the one studied in this report could provide such a resource, but again this does not seem to provide for a scalable attack.

We argue that, while people looking to utilize an open network may use the information emitted, it does not appear that additional value for evildoers comes from the aggregation of this data nor does the data appear to provide for a scalable attack due to the necessity of physical proximity to the access points.

While the aggregation of the location data does not appear to represent much security or privacy concern, we are compelled to reiterate the concern regarding default configuration. It is feasible that as many as 26 million access points remain in an unsecured configuration.

The use of this location information to track devices that use the WPS service or subsequently use the location information from other Location Based Services remains a security and privacy concern and is on our agenda for further investigation.

G. Infrastructure and Deployment

As noted, the density of access points in urban and suburban areas appears to present opportunities for new network models. Several of these models are in commercial or research experimentation today, including:

¹⁵ However, in previous research these two data have been closely related, see <http://www.extremetech.com/article2/0,1697,1152933,00.asp> for example.

¹² U.S. Census Bureau Geography: <http://www.census.gov/geo/www/>

¹³ Income and population data obtained from the U. S. Census Bureau (<http://factfinder.census.gov>).

- Social sharing of networks such as within a neighborhood, apartment, or other structure.
- Commercial aggregation of private access points such as services being marketed by FON and WiFiTastic.
- Open Wi-Fi networks for vehicular Internet access [4].
- Extension of municipal wireless networks by allowing private access points to be ‘linked’ into the municipal grid to extend the signal reach.

This last point is of particular interest as it is a model that turns a personal asset into a true community benefit. There remain many questions about the viability of such a model, from cost, to acceptance, to interference, to abusive behavior.

The applicability of this model, however, has implications in international wireless network deployment. In fact, the One Laptop Per Child¹⁶ (OLPC) project relies on a similar model, utilizing a multi-hop mesh technology to provide Internet access by traversing nodes until it reaches an external gateway. To meet this need, still others are working on overlay networks to provide the infrastructure for Wi-Fi access; the Green WiFi project is an example of this effort¹⁷.

IV. CONCLUSIONS

This study is the first to look at the information available from a large-scale database of geolocated access points, and provides a glimpse into the value of the data. Our initial analytical results show that statistical mining of this data and the information revealed by this data (such as the default naming behavior, movement of access points over time, and density of access points) can yield important technological, social and economical results. Concretely, the findings of this analysis can provide technological guidelines for the design of wireless network systems that are more efficient, more scalable, and more reliable. The results have also suggested both social and economical benefits of open networks. For example, the data set has provided a glimpse into the market dynamics by examining the actual statistics of manufacturer data of the access points that have been deployed as well as their location of deployment. Further, the behaviors of the people who install the access points (whether business or individual) can yield some interesting results, including the default configuration habits of those users. In addition, we show that the geographic information can be exploited to understand the wireless infrastructure at large by exploring network characteristics such as access point density, demographic propensities, and signal propagation behavior. We believe that many areas in the U.S. would be able to support new models of use for the already deployed wireless access points.

Our research will continue along several dimensions, aiming at continuing to improve our understanding of this rapidly growing and changing infrastructure of wireless networks and the applications that it can support. The main areas that we are currently exploring for further research include.

Improve location accuracy. We continue to analyze the growing data set – currently over 12 million APs, 300 million GPS readings and 4 billion signal strength records, and explore methods to improve the location accuracy and confidence. One proposed approach is to use geographic constraints to improve both the original access point positioning calculation as well as the resulting end-user location calculation. Accuracy as well as measures of confidence in accuracy will improve usability of location-based applications.

Demographic Analysis. Further analysis of the AP density is currently underway by utilizing the GIS system to determine density by census tract. This analysis will provide further detail and introduce the ability to correlate demographic details to the access point distribution.

Extending Dataset collected. Gathering additional parameters such as security settings and signal frequency can enable even further analysis and help to understand the potential level of signal interference and provide insight into methods to reduce the conflict and improve future deployments.

Access Point Coverage Models. With the more than 4 billion scan records, it would be possible to construct more detailed models of the propagation pattern for each access point. Coupling this model with GIS information could lead to an even better understanding of the coverage of existing and yet to be deployed networks. For example, deployment planning of these nodes could be improved by creating more accurate signal propagation models. Linking the scanned data from existing Wi-Fi access points with knowledge of land use and land cover (LULC) data can extend our knowledge and model of Wi-Fi signal propagation [10].

Current propagation models can be tested against the raw scan data that has been acquired by Skyhook Wireless. These models can be further refined by introducing elements from the GIS model (elevation, structures, ground cover) to create a more accurate representation of the signal footprint.

ACKNOWLEDGMENT:

This research is partially supported by Skyhook Wireless, Georgia Tech Tennenbaum Institute, grants from NSF SGER, NSF CyberTrust, NSF CSR, NSF ITR, IBM SUR grant, and IBM Faculty Award. Portions of the described system are covered by patent pending receipt number 15528274.

REFERENCES

- [1] Balachandran, A., Woelker, G.M., and Bahl, P. (2003). Wireless hotspots: current challenges and future directions. In Proceedings of WMASH 2003, pp 1-9, Sept. 2003.
- [2] Battiti, R., Lo Cigno, R., Sabel, M., et al. (2005). Wireless LANs: From WarChalking to open access networks. *Mobile Networks & Applications* 10 (3): 275-287 JUN 2005.
- [3] Borisov, Nikita, Goldberg, Ian, and Wagner, David (2001). Intercepting mobile communications: The insecurity of 802.11. In Proceedings of MOBICOM 2001, 2001. <http://citeseer.csail.mit.edu/article/borisov01intercepting.html>.
- [4] Bychkovsky, V., Hull, B., Miu, A., Balakrishnan, H., and Madden, S. (2006). A measurement study of vehicular Internet access using In Situ Wi-Fi Networks. *MobiCom '06*, September 24-29, 2006.
- [5] Byers, S., Kormann, D. (2003). 802.11B Access Point Mapping. *Communications of the ACM*, May 2003, Vol.46, No. 5.

¹⁶ The One Laptop per Child project is focused on providing computing and information technology to children in developing countries (<http://www.laptop.org/>).

¹⁷ The Green WiFi project uses solar powered Wi-Fi access points to provide a ‘self-healing’ grid network (<http://www.green-wifi.org/>).

A version of this paper has been submitted to IEEE Portable 2007.

- [6] Frederickson, G. N. (1979). Approximation Algorithms for Some Postman Problems. *J. ACM* 26, 3 (Jul. 1979), 538-554.
- [7] Griswold, W. G., Boyer, R., Brown, S. W., Truong, T. M., Bhasker, E., Jay, G. R., and Shapiro, R. B. (2002). Activecampus - sustaining educational communities through mobile technology. Technical Report CS2002-0714, University of California, San Diego, Department of Computer Science and Engineering, July 2002.
- [8] Gruteser M., Grunwald, D. (2004). A methodological assessment of location privacy risks in wireless hotspot networks. *Lecture Notes in Computer Science* 2802: 10-24 2004.
- [9] Henderson, T., Kotz, D. and Abyzov, I. (2004). The changing usage of a mature campus-wide wireless network. *Proceeding of MobiCom 2004*. pp187-201. Sept. 2004.
- [10] Kirner, J. L. and Anderson, H. R. (1998). The application of land use cover data to wireless communication system design. In *Proceedings of the ESRI User Conference*, 1998.
- [11] LaMarca, A. et al., (2005). Place Lab: Device Positioning Using Radio Beacons in the Wild. *Proc. 3rd Int'l Conf. Pervasive Computing (Pervasive 05)*, LNCS 3468, Springer, 2005, pp. 116-133.
- [12] Letchner, J., Fox, D. and LaMarca, A. (2005). Large-Scale Localization from Wireless Signal Strength. *Proceedings of the National Conference on Artificial Intelligence (AAAI 2005)*.
- [13] Mishra, A., Petroni, N. L., Arbaugh, W. A., et al. (2004). Security issues in IEEE 802.11 wireless local area networks: a survey. *Wireless Communications & Mobile Computing*, 4 (8): 821-833, December 2004.
- [14] Papadimitriou, C. H. (1967). On the complexity of edge traversing. *JACM* 23, 3 (July 1976), 544-554.