Infrastructure-based location estimation in WLAN networks

Sachin Ganu[1]
WINLAB, Rutgers University
73 Brett Rd
Piscataway, NJ 08854
Email: sachin@winlab.rutgers.edu

A. S. Krishnakumar
Avaya Labs Research
233 Mt. Airy Rd.
Basking Ridge, NJ 07090
Email: ask@avaya.com

P. Krishnan
Avaya Labs Research
233 Mt. Airy Rd.
Basking Ridge, NJ 07090
Email: pk@avaya.com

Abstract

This paper focuses on fundamental system deployment aspects of location estimation in 802.11-based wireless networks. We concentrate on adaptable infrastructure-based approaches, where sniffers measure received signal strength from clients to locate them. Our implementation experience and experimental results show that sniffer-based location estimation is feasible and works well provided some important rules are followed. By studying data over a 6-month period, we observe that adaptation of models is necessary for good location estimation, and we show that our techniques enable location estimation with minimal profiling. We also present an intriguing client-assisted approach for location estimation where a client, APs and sniffers collaborate to locate a terminal in enterprise (infrastructure-mode) wireless networks.

1. Introduction

Many signal strength (SS)-based techniques have been proposed for location estimation in wireless networks, especially 802.11-based networks. While several researchers have concentrated on techniques for location estimation and improvements to them [3][4][6][7][8][11][12], there has been little work done in studying the systems aspects of how such techniques will actually be deployed. In particular, it has been implicitly assumed that the deployments will be *client-based*, where (i) the SS models are built by profiling the site, using measurements from visible access points (APs), and (ii) for location estimation, a client reports back SS measurements from visible access points, and these measurements are compared against the SS model to locate it. However, many enterprises, especially for ease of management (i.e., provisioning, security, deployment and maintenance) would prefer an infrastructure-based deployment where simple sniffers monitor client activity and measure the SS of transmissions from clients. With reducing chipset prices, low-cost sniffers can be built, and they can also provide a platform for supporting other monitoring services like security, intrusion detection, and QoS measurements. The use of an infrastructure-based (i.e., sniffer-based) deployment does raise several interesting options and issues that we study here.

In this paper we present *Palantir*, an infrastructure-based monitoring and location estimation system. We articulate the opportunities and challenges raised by such a sniffer-based deployment, and present our implementation insights and experimental results. A fundamental question we ask is: Is the use of sniffers for location estimation adequate, and how does it compare to client-based location estimation? We make some interesting experimental observations on the system-level reciprocity of SS measurements and their impact on location estimation, compare client-based and sniffer-based location estimation, and study the impact of the location of sniffers on location estimation.

Additionally, most of the recent work in location estimation techniques requires a substantial amount of site profiling to build their SS models [3][6][7][11]. Changing radio environments at a site [5][7] due to environmental, building and occupancy conditions affect signal propagation models and may require frequent re-profiling. Infrastructure changes, e.g., adding/moving access points also adversely impact signal strength models. Environments like warehouses and malls are especially dynamic. Our experiments reported in this paper show the need for adaptation of location estimation SS models even in seemingly static environments. Techniques for location estimation that do minimal or no profiling [5] will be particularly useful in this context; our work described in this paper supports such architectures. We also motivate and present a new client-assisted approach for location estimation, where client terminals, APs, and sniffers collaborate to locate the clients in an (infrastructure-mode) wireless network.

1.1. Related Work

Prior techniques for model building in location estimation include work where each point is mapped to a SS (Signal Strength) vector [6] or a SS probability distribution [11]. For matching a SS vector to a model, nearest neighbor [6] or probabilistic techniques [11] are used. The necessity of model adaptation was identified in [5][7], and [5][9] considered methods of simplifying model building. Providing client libraries and APIs to support client-based deployments was studied in [7].

Infrastructure-based approaches using sniffers have been proposed [10][16][17], but there is little work comparing them with client-based approaches, or understanding the issues associated with their use. In [10], the authors discuss a sniffer-based approach for location in prison environments concentrating on the RF hardware used. In particular, the

[1] Portions of this work were done when S. Ganu was visiting Avaya Labs.

nature of their environment was specialized, and there was little discussion of issues like channel hopping, number of sniffers and their location, profiling and adaptation issues, comparing client- and sniffer-based approaches, etc., that is the main focus of this paper. We are also not aware of any investigation into mixing client-provided information with sniffed information for location estimation.

In Section 2, we present our Palantir system, discussing the design of our sniffers in detail. We then present our location estimation experimental results in Section 3. Our experiments use data from a 6-month period. In Section 4, we introduce our model of client-assisted location estimation and study its benefits.

2. The Palantir System

We designed and implemented *Palantir*, an infrastructure-based monitoring and location estimation system. The Palantir system uses sniffers to monitor information about clients, and uses the monitored information for location estimation and security. In this paper, we concentrate on the location estimation aspect of Palantir, which is based on received SS.

The design of the Palantir system is motivated by the LEASE architecture [5] that uses sniffers and emitters in a new way for location estimation. The LEASE architecture and method requires minimal profiling and automatically adapts the SS model used for location estimation when the environment changes. The concentration in this paper is on the sniffers used to detect clients and the SS of the transmissions from the clients. Available sniffing tools (software) [14] are mostly monolithic, client-based and do not provide any location estimation or remote monitoring capabilities. While a basic wireless sniffer is quite easy to build, the aspects of using a sniffer for location estimation are intriguing as described below in Section 2.1. Our location estimation strategy is described in Section 2.2.

2.1. Sniffers in Palantir

The sniffers are the main component in Palantir, and are built on a single board computer platform with a dual Ethernet interface and a PCMCIA slot for a wireless card [15] that allows easy deployment. The sniffers operate in a passive scanning mode and sense the wireless medium on all or predetermined channels. They listen for communication from wireless terminals and record and timestamp information. In particular, the sniffers capture the management, control and data frames and decode information such as the MAC address, SSID etc. which is present in the frame header, and also extract the received SS from its wireless interface card. Currently, we do not decrypt the payload and only look at information that is unencrypted in the header. The measurements are sent by each sniffer to a centralized database, and used for security assessments and location estimation. The sniffer's wireless interface is entirely passive, and all communication with the database is through its Ethernet interface. The sniffers could either be co-located with AP's or at other positions based on the availability of Ethernet jacks and power outlets. Some of the issues we encountered while deploying sniffers and using them for location estimation are discussed below.

2.1.1. Asymmetry of signal strength

While performing experiments with Palantir, we found that there exists some asymmetry while measuring SS between two devices. Our experiment here involved a sniffer co-located with an AP. A client measured the received SS from the AP (using probe request responses) while the co-located sniffer measured the SS from the client (using packets received), and the difference, ΔSS, was computed. (Both client and AP were transmitting at the same power level.) With different locations of the client, we observed that ΔSS could be as large as 10 dBm. Some of this discrepancy may be attributed to the non-simultaneous measurement of SS in the two directions. This will still leave a residual asymmetry and any system must be resilient to this variation. (We also experimented with interchanging the wireless cards between the client and the sniffer, and using different types of cards; in all cases, we observed asymmetry.) We refer to this as *system-level* asymmetry of SS. It is not clear if this asymmetry affects location estimation, which is typically a function of many signal strength measurements, and we analyze that later in Sections 3.2 and 3.3.

2.1.2. Number of packets received

An important consideration in sniffer design is to ensure that "typical" client activity can be sniffed. (This is a new issue when compared to a client-based system.) More importantly, since typical enterprises would use several channels (e.g., the three non-overlapping channels in 802.11b), and sniffers need to measure SS from all radio-visible clients, they must scan more than one channel. The number of packets received at the sniffer depends on the sweep rate of the sniffers and the dwell time on each channel. Note that for general sniffing, it is sufficient to see one packet from a client; for location estimation, seeing more helps a lot given the vagaries of SS behavior [6][7][12].

To study the issue of number of packets seen, we conducted a trace-driven simulation. We captured time-stamped packets from one HTTP and one email transaction; typical activities a mobile client would perform. (We did this for a VPN and non-VPN access from the client, considering current wireless deployment architectures [18], and observed similar results.) The transactions were short, and completed in approximately 4s (for sending one short email) and 10s (for http; here we made a request to cnn.com with all embedded images). Assuming that the sniffer was at a random channel at the start of the transaction, and the client was at a fixed channel, we computed the average and minimum packets seen (over several choices for the random channel) for the respective transactions. Representative results are summarized in Figure 1.

We observe that (i) when only three channels (e.g., the non-overlapping channels in 802.11) are used and need to be sniffed for location estimation, typical transactions (like email and http) provide enough packets (measured as the average and minimum number of packets seen) for SS measurement for an appropriate choice of channel dwell time, and (ii) if all 11 channels need to be sniffed, a specific transaction might not get recorded by the sniffers. However, for a "chatty" terminal, the probability of the sniffer missing the terminal's traffic will decrease substantially with the number of transactions.

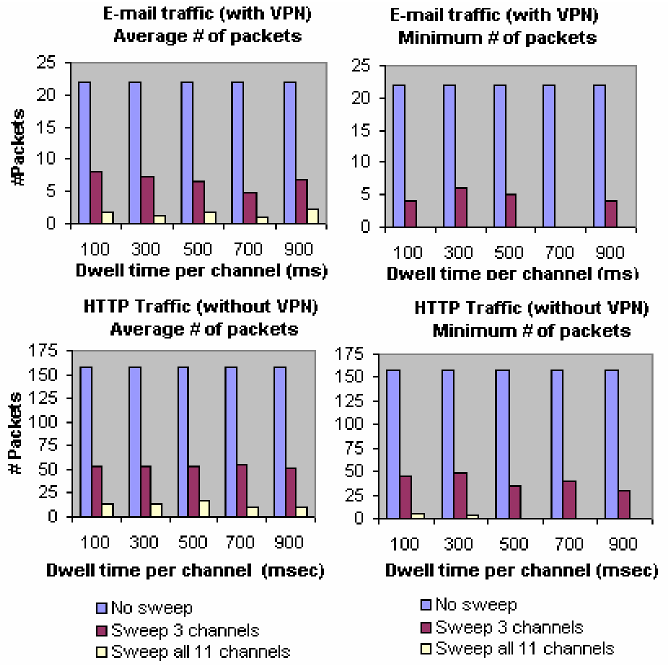

Figure 1 Average and Minimum packets received for E-mail and HTTP traffic

Motivated by our observations, in our current deployment (where only the three non-overlapping channels are used in the enterprise), we configured the sniffers to sweep at the rate of three channels per second, giving a dwell time on each channel of 333 ms. We also have the ability to determine the IP address of most clients (e.g., by querying appropriate management information bases, MIBs) and eliciting (in most cases) packets from the clients in response to a `ping`. Similar actions are possible at the MAC layer also.

2.1.3. Channel estimation

Our implementation allows sniffers to be configured to sweep through all channels, a few channels or dwell on a particular channel. The results from Figure 1 motivate this aspect, since having to scan on many channels may result in not receiving sufficient packets from a client, unless all the sniffers in radio-range of a terminal are tuned to the channel in which the client is transmitting. This requires us to know the channel on which a client is transmitting. Channel information can be determined based on the class of clients targeted.

"Associated" clients. The channel information is already embedded in the beacons (both in infrastructure and ad-hoc mode) [2]. A terminal is on the same channel as the AP with which it is associated. Thus, we can determine the channel for the terminal.

Other clients. In this case, the terminal is not associated with an AP (e.g., is operating in ad-hoc mode, or is a rogue of some sort). Since the adjacent channels of 802.11b are overlapping [2], we observed that the sniffers are able to receive and decode packets correctly even ±1 channel on either side of the actual channel on which the terminal dwells. In this case, the sniffer uses a heuristic of estimating the client's channel to be the one on which it records the strongest signal strength. We observed that in most cases our heuristic works; however, and interestingly, it is not perfect (i.e., the strongest signal is occasionally seen from an adjacent channel). We hope to study this issue in more detail later. Our current work emphasizes associated clients.

2.1.4. Location of the Sniffers

The location of sniffers will likely be dictated by location of power outlets and/or Ethernet jacks (since power over Ethernet is supported in our sniffers). However, if there are several choices for sniffer placement, which ones should be chosen? Co-locating sniffers with APs is a definite possibility, but is there a benefit in not doing so? We investigate this experimentally in Sections 3.4 and 4. We note that APs are usually deployed for coverage with some (minimal) overlap, and not to ensure location estimation which typically requires a view of multiple APs at any point on the floor. We have also observed sites with APs deployed linearly, leading to obvious ambiguity in location estimation. We expect that administrators would prefer to deploy passive sniffers rather than additional APs. While deploying sniffers "far away" from APs could technically lead to some packets getting lost due to hidden station problems, we have not noticed in our experiments any discernible issues due to this.

2.2. Location Estimation via SS in Palantir

The location estimation technique in Palantir is based on the one in [5], and is summarized in this section. Specifically, the floor of the building is divided into (3ft × 3ft) grids, the received SS data from the profiled points at each sniffer is smoothed and local bivariate interpolation (using Akima splines [1]) is used to estimate the SS at the center of each grid. Putting together the estimates for all sniffers provides a SS vector at each grid center, or the *model*. A client's received SS at each sniffer is compared against this model using nearest neighbor search [6] to get the estimated location of the client. The technique allows for an SS model to be built with very few profiled points; these profiled points are expected to be uniformly distributed on the floor of the site. We also peg signal strength from below at a small value *s* (-92dBm), i.e., the absence of SS is interpreted as a SS of *s*. We note that in LEASE [5], by deploying stationary emitters at the points where profiling is desired, the most current SS model can be adaptively and automatically built.

3. Location Estimation Experiments and Results

3.1. Experimental setup

The experiments reported in this paper were performed at a site we refer to as BR over the course of 6 months. Site BR measures 225ft × 144ft and has five deployed APs (the diamonds in Figure 2). We used five sniffers in our experiments, and a client (IPAQ running Linux). We experimented with two different sniffer deployment locations (co-located with APs, shown using the squares, and a "diamond" configuration, shown using the triangles in Figure 2). We experimented with SS data collected using client-and sniffer-based profiling, where the profiling was done in the "open" areas of the building. Note that the diamond configuration has two sets of three collinear sniffers; in normal deployments, such collinearity will be avoided, but in our experiments we gain valuable insight with such a scenario.

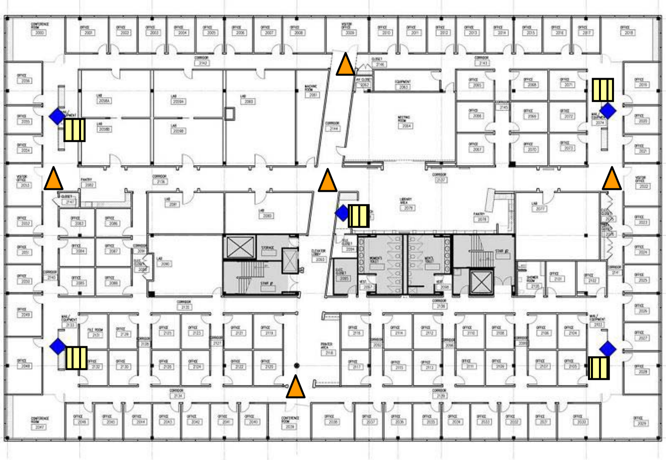

Figure 2 Site BR, and placement of APs and sniffers.

When profiling, we took several measurements at different points on the floor, and at every point the following steps were taken:

- Measure the signal strength from the access points at the client and record it as a *client reading* (*x, y*, client ss-vector)
- Generate and send a short burst of packets from the client. The sniffers then record the signal strength from this client and report to a database with a timestamp. This comprises the *sniffer reading* (*x, y*, sniffer ss vector)

We broadly grouped the data sets into three categories: *Set A* (original data) had only client readings, *Set B* (data taken about 4 months later) had client and sniffer readings with sniffers co-located with APs, and *Set C* (data taken about 5 months later) had client and sniffer readings with the sniffers in the diamond configuration. Unless otherwise specified (e.g., as in Section 3.5), each result represents experiments using data from one set; we note that the conclusions presented here largely hold independent of the data set used. In all cases, the data used for testing location estimation was always different from the one used to build the model. (The technique for model building and location estimation was summarized in Section 2.2.) In our results presented below, the estimation error is the median error (chosen for easy comparison with prior work), and the error is depicted as a function of the number of profiled points. In effect, the points chosen to build the model were chosen as close to be as uniformly spaced as possible on the floor, from the data in the model-building set. A *client-based model* refers to a model built using client readings, and a *sniffer-based model* refers to a model built using sniffer readings. Similarly, the tests are referred to as *client-based tests* and *sniffer-based tests* depending on the readings used.

Several experiments were run using different combinations of the collected data. The following sub-sections describe in detail the various experiments performed and their results.

3.2. Comparing Client and Sniffer-based approaches

We compared the performance of a client-based approach against a sniffer-based approach for location estimation; the sniffers were co-located with the APs. As shown in Figure 3, we see that both client and sniffer-based approaches perform comparably. The result is quite interesting when taken in conjunction with our observations from Section 2.1.1, and raises the question of whether interchanging the SS models (i.e., using a client-based model for sniffer-based testing) would work. We study this below in Section 3.3.

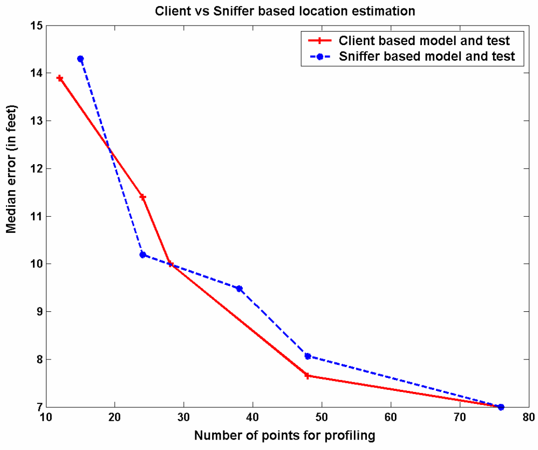

Figure 3 Client vs. sniffer-based location estimation; Data: *Set B*.

We would like to re-iterate here a result from [5], also shown by Figure 3 that the technique from Section 2.2 provides very good location estimation in absolute terms with very little profiling. In particular, we note that the median error of 14.3-10.3 ft with 15-24 profiled points is better than reported in prior work [6] using comparable profiling. Note also that the area of site BR (approx. 32000 sq. ft.) is much larger than the site in [6] (approx. 10500 sq. ft.).

3.3. Asymmetric profiling and testing

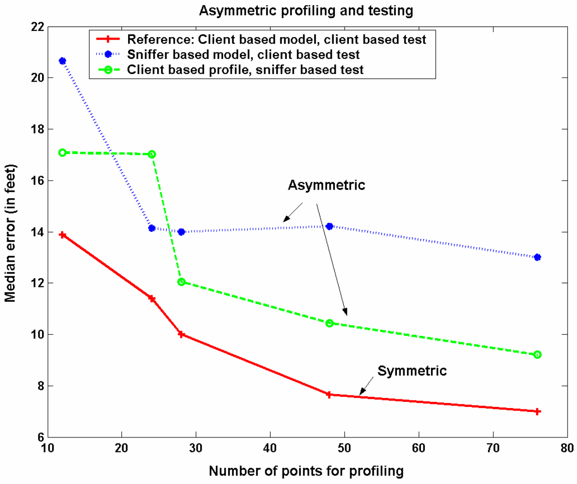

Figure 4 Mixing client- and sniffer-based profiling and location estimation; data: *Set B*

We determined the location estimation accuracy when a client-based model is used for sniffer-based testing and vice-versa. In this case (data from set *B*), the sniffers and APs were co-located. As shown in Figure 4, we found that in mixing the two techniques, there is a discernible degradation in the performance. This seems reasonable when taken in conjunction with the observation from Section 2.1.1 that from a "system" point of view, SS measurements are not always reciprocal. Figure 4 suggests that symmetry is recommended in approaches for profiling and location estimation; i.e., use a client (sniffer)-based model when locating using a client (sniffer)-based approach.

3.4. Location estimation errors vs. location of sniffers

As noted in Section 2.1.4, where we locate sniffers could be an important issue, and we may not co-locate sniffers with APs depending on AP placement and coverage. In this context, we placed the sniffers in two different configurations, the co-located and diamond configurations (See Section 3.1). Table 1 shows "visibility", i.e., the percentage of profiled points seen by *k* sniffers, as a function of *k,* the number of sniffers that see points.

Case	***Number of sniffers, k***					
	0	**1**	**2**	**3**	**4**	**5**
Co-located	0	13.2	18.4	46.0	22.4	0
Diamond	0	0	6.8	20.2	33.7	39.3

Table 1 Percentage of profiled points seen by *k* sniffers.

We observe that the diamond configuration seems to have more visibility in that more profiled points are seen by more sniffers. Figure 5 shows the sniffer-based location estimation results for the two configurations.

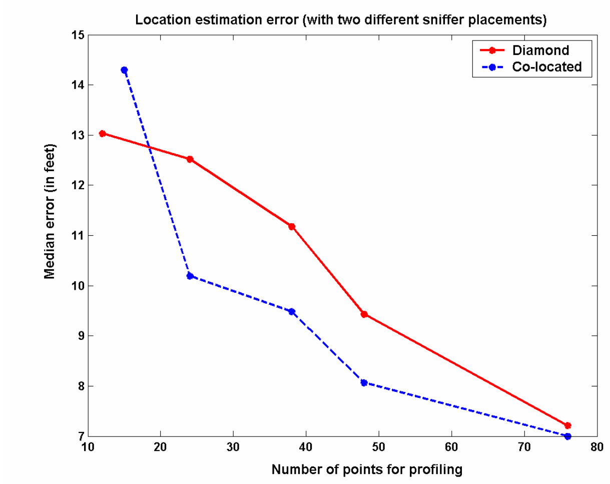

Figure 5 Performance with two different sniffer placements; data: *Set B* (co-located), *Set C* (diamond).

We find that the co-located configuration performs a little better than the diamond configuration in our experiment, likely due to the obvious ambiguity in some estimates introduced by the collinear sniffers in the diamond configuration. We note that more sniffers seeing a point need not always directly translate to better location estimation. The non-co-located (diamond) configuration, however, has additional benefits that we explore later in Section 4.

3.5. A Case for Adaptation: Using an older Model

We used client readings from our three sets of data for this experiment. We used a model built using an earlier data set against test points collected more recently. As shown in Figure 6, we found that the median error results are substantially higher for the cases where the tests were performed using models that were built much earlier, especially when the number of points used for profiling is larger. Interestingly, when using an older profile, the error seems to "level out" and not go down with increased profiling.

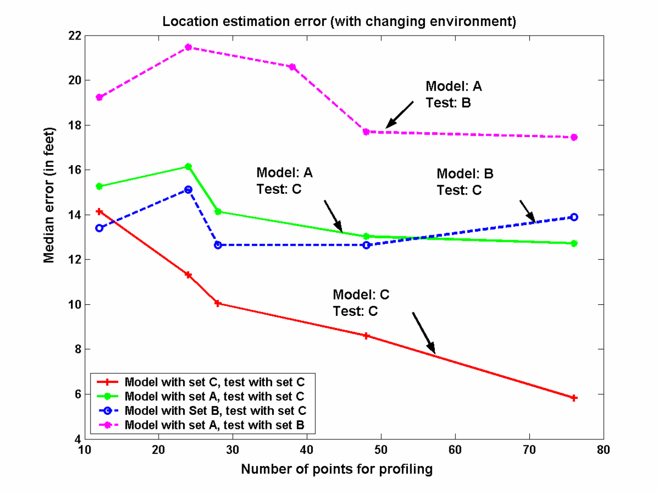

Figure 6 Using an older profile for location estimation. Data set A > B > C, where ">" means "older than".

This seems to indicate that the radio model does change over time. The result may be considered a little surprising, since our site is not very dynamic (unlike a mall, restaurant, or warehouse, for example), but does see changes in its occupancy amount at different times. We conclude that the model needs to adapt to changes in the environment in order to have reasonable accuracy in location estimation. Our techniques facilitate easier adaptation since they require minimal profiling without compromising location estimation accuracy.

4. Client-Assisted Location Estimation

In this section, we study an interesting twist to the location estimation deployment problem. Consider that we profile and build an SS model using the sniffer readings and also for the APs, using the client readings. As described earlier in this paper, the sniffers report the SS measured from client transmissions. Additionally, assume that the client also reports the SS seen from the APs (e.g., using appropriate APIs [7]). How much additional benefit is gained by using the information from the clients? We refer to this architecture in which the sniffers locate clients, but where the client helps in its location by providing SS readings from visible APs as *client-assisted location estimation.* Clearly, client assistance has particular relevance when the sniffers are not co-located with the APs.

To experimentally study the benefit of client-assistance, we took the data from *Set C* and melded the client and sniffer data appropriately. In particular, we created a vector of size 10 for each profiled point corresponding to the 5 sniffer readings and the 5 client-based readings. We built a model using this data set and tested against an appropriate test set, where each point in the test set also had a SS vector of size 10, with 5 sniffer readings and 5 client-based readings.

Figure 7 shows the results of the experiment. We see that there is a significant advantage in client-assistance, with an improvement in median estimation error of a few feet.

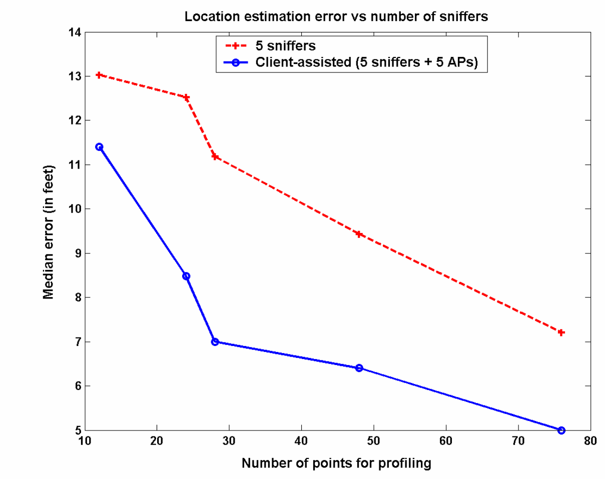

Figure 7 Client-assisted location estimation; data: *Set C.*

Figure 7 can also be interpreted somewhat loosely as the improvement observed with increasing the number of sniffers (or the number of APs if one were to use a fully client-based deployment.) In particular, client-assistance with 5 APs and 5 sniffers is (using the intuitions from Sections 3.2 and 3.3) almost like having 5 sniffers co-located with the APs in addition to the existing 5 sniffers. We note some interesting aspects though. Many applications would likely require location estimation that is adequately provided by just the sniffers (5, in our case), and this can be achieved with *no* client changes or involvement. For other applications that require more precision, client-assistance can be used. Sniffers provide monitoring benefits in addition to location estimation, and will likely be deployed in enterprises. Hence, client-assisted infrastructure-based location estimation architecture will provide the benefits of both worlds.

While the result from Figure 7 may at first glance seem to contradict an observation in [7] where the authors suggest that there is little advantage in going over 3 APs, we note that all location estimation results are dependent (at least) on the size of the site. While we also expect diminishing returns with more sniffers, we believe that with larger sites, increasing the number of sniffers proportionally might help.

5. Conclusion

In this paper, we have studied practical issues in infrastructure-based deployment for location estimation in WLAN networks. Using sniffers to monitor clients and a signal-strength based model for location estimation, we have discussed several issues in sniffer implementation, including reciprocity of signal strength, number of packets expected to be seen by the sniffers, location of sniffers, etc. Through detailed experiments using data collected over a period of 6 months, we have demonstrated that a sniffer-based approach to location estimation is both feasible and desirable, provided certain rules (described in the paper) are followed. We have seen good location estimation with minimal profiling. We have also shown that having a sniffer-based approach where the sniffers are not co-located with APs, and where clients assist in their location enables an interesting client-assisted location estimation strategy that provides good estimates of a terminal's location.

6. References

[1] H. Akima, "A new method of Interpolation and Smooth Curve Fitting based on Local Procedures," *Journal of the ACM, Vol. 17, No. 4,*October 1970, pp 589-602.

[2] IEEE 802.11b Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, http://standards.ieee.org/getieee802/802.11.html

[3] Andrew M. Ladd, Kostas E. Bekris, Algis Rudys, Guillaume Marceau, Lydia E. Kavraki, Dan S., "Robotics-Based Location Sensing using Wireless Ethernet," *The Eighth ACM MOBICOM Conf.*, September 2002.

[4] P. Prasithsangaree, P. Krishnamurthy, P. K. Chrysanthis, "On Indoor Position Location With Wireless LANs, " *The 13th IEEE PIMRC Conference,* September 2002.

[5] P. Krishnan, A.S. Krishnakumar, Wen-Hua Ju, C. Mallows, S. Ganu, "A System for LEASE: System for Location Estimation Assisted by Stationary Emitters for Indoor RF Wireless Networks", *Proceedings of IEEE Infocom 2004, Hong Kong, March 2004.*

[6] P. Bahl, V.N. Padmanabhan, "RADAR: An In-Building RF-based User Location and Tracking System," *Proceedings of IEEE Infocom 2000, Tel Aviv, Israel,* March 2000.

[7] P. Bahl, V.N. Padmanabhan, A. Balachandran, "Enhancements to the RADAR User Location and Tracking System," *Microsoft Research Technical Report*, February 2000.

[8] S. Saha, K. Chaudhuri, D. Sanghi, P. Bhagwat, "Location Determination of a Mobile Device using IEEE 802.11 Access Point Signals," *IEEE WCNC Conference*, 2003.

[9] A. Smailagic, D.P. Siewiorek, J. Anhalt, D. Kogan, Y. Wang, " Location Sensing and Privacy in a Context Aware Computing Environment," *Pervasive Computing,* 2001.

[10] T. W. Christ, P.A. Godwin, "A Prison Guard Duress Alarm Location System," *Proc. IEEE International Carnahan Conference on Security Technology,* October 1993.

[11] Moustafa Youssef, Ashok Agrawala, A. Udaya Shankar, "WLAN Location Determination via Clustering and Probability Distributions," *IEEE Intnl. Conf. on Pervasive Comp. and Comm. (PerCom) 2003*, March 23-26, 2003.

[12] Moustafa Youssef, Ashok Agrawala, "Small-Scale Compensation for WLAN Location Determination Systems," *IEEE WCNC Conference,* March 16-20, 2003.

[13] James, G., and Hastie, T. "Generalizations of the Bias/Variance Decomposition for Prediction Error," *Technical Report,* Department of Statistics, Stanford University, Stanford, CA.

[14] Kismet, http://www.kismetwireless.net

[15] Soekris Engineering, http://www.soekris.com

[16] Air Defense, http://www.airdefense.net

[17] Newbury Networks, http://www.newburynetworks.com

[18] Intel White Paper, "VPN and WEP: Wireless 802.11security in a corporate environment", http://www.intel.com/eBusiness/pdf/it/wp021306.pdf