Poster Abstract: The Limits of Localization Using RSS

Eiman Elnahraway
eiman@cs.rutgers.edu

Xiaoyan Li
xili@cs.rutgers.edu

Richard P. Martin
rmartin@cs.rutgers.edu

Department of Computer Science
Rutgers University
Piscataway, NJ 08854

ABSTRACT

We characterize the fundamental limits of localization using signal strength in indoor environments. Signal strength approaches are attractive because they are widely applicable to wireless sensor networks and do not require additional localization hardware. We show that although a broad spectrum of algorithms can trade accuracy for precision, none has a significant advantage in localization performance. We found that using commodity 802.11 technology over a range of algorithms, approaches and environments, one can expect a median localization error of 10ft and 97th percentile of 30ft. We present strong evidence that these limitations are fundamental and that they are unlikely to be transcended without fundamentally more complex environmental models or additional localization infrastructure.

Categories and Subject Descriptors

C.2.5 [**Local and Wide-Area Networks**]

General Terms

Algorithms, Measurement, Performance, Design, Experimentation

Keywords

Localization, Wireless Local Area Networks

1. INTRODUCTION

Localizing sensors is necessary for many higher level sensor network functions such as tracking, monitoring and geometric-based routing. Recent years have seen intense research investigating using off-the-shelf radios as a localization infrastructure for sensor nodes. The motivation has been a dual use one: using the same radio hardware for both communication and localization would represent a tremendous savings over deployment of a specific localization infrastructure.

In this work we explore the fundamental limits of localization using signal strength in indoor environments. Such environments are challenging since the radio propagation is much more chaotic than outdoor settings, where signals travel with little obstruction. Exploring the limits of signal strength approaches is important since it tells us the localization performance we can expect without additional hardware in the sensor nodes and base-stations. We use the 802.11 Wireless Local Area Network (WLAN) technology in our study, because of its commodity status. Our results however are applicable to any radio technology where there are considerable environmental effects on the signal propagation.

We compared a wide range of existing *point-based* localization algorithms. We also developed 3 novel algorithms that are *area-based*. That is, the returned localization answer is a possible area (or volume) that might contain the sensor radio rather than a single point. The key property of such algorithms is that they can trade accuracy for precision, where *accuracy* is the likelihood the object is within the area and *precision* is the size of the returned area. In contrast, point-based approaches have difficulty describing such probabilistic trade-offs in a systematic manner. Using accuracy and precision, we were able to quantitatively describe the limits of different localization approaches by observing the impact of increased precision (i.e. less area) on accuracy.

Although examining this accuracy vs. precision tradeoff gives insight into performance limits, such an approach does not help us reason if the observed limitations are fundamental to the algorithm or inherent in the data. Area-based algorithms, however, have an additional critical advantage in their ability to describe localization uncertainty. Therefore, using a Bayesian network, as one of the 3 area-based algorithms we developed, we express the uncertainties arising from these effects in terms of probability density functions (PDFs) that describe the likely position as a function of the observed data and a widely used propagation model.

Our study showed that a broad spectrum of signal-strength based algorithms have similar localization performance. Our results also showed that there is significant uncertainty arising from the data given the Bayesian model. Our conclusion is that these limitations are fundamental and that they are unlikely to be transcended without qualitatively more complex models of the environment or additional hardware above that required for communication.

SenSys'04, November 3–5, 2004, Baltimore,Maryland, USA.
ACM 1-58113-879-2/04/0011.

2. ALGORITHMS

Table 1 summarizes our algorithm menagerie. The algorithms exhibit a broad range of localization techniques, including fingerprinting, signal propagation modeling, maximum likelihood estimation, and Bayesian networks.

3. METRICS

We used a broad range of performance metrics. The traditional localization metric is the "distance error" between the returned position and the true position. However, a problem with this metric is that it does not apply to area-based approaches. We thus introduce metrics appropriate for area-based algorithms.

3.1 Area-Based Metrics

Tile Accuracy. Many of our area-based algorithms describe space as a set of small discrete tiles, rather than as a continuous quantity. Tile accuracy thus refers to the percentage of times the algorithm is able to return the true tile that contains the object. This metric can be somewhat misleading because often, the true tile is close to the returned set, which motivates the next metric.

Distance Accuracy. This is the distance between the true tile and tiles in the returned area. To gauge the distribution of tiles in relation to the true location, we sort all the tiles according to this metric. We then return the distances of the 0^{th} (min), 25^{th}, 50^{th} (median), 75^{th}, and 100^{th} (max) percentiles of the tiles. This metric is somewhat comparable to the traditional metric.

Precision. The overall precision refers to the size of the returned area, i.e., the sq.ft.

Algorithm	Abbreviation	Description
	Area-Based	
Simple Point Matching	SPM	Matches the RSS to a tile set using thresholds.
Area Based Probability	ABP-α	Matches the RSS to a tile set probabilistically with confidence bound α%.
Bayesian Network	BN	Returns the most likely tiles using a Bayesian network.
	Point-Based	
Bayesian Point	B1	Returns the most likely point using a Bayesian network.
Averaged Bayesian	B2	Returns the mid-point of the top 2 most likely points.
RADAR [1]	R1	Finds the closest training point based on distance in signal space.
Averaged RADAR	R2	Returns the midpoint of the closest 2 training points in signal space.
Gridded RADAR	GR	Applies RADAR using an interpolated grid.
Highest Probability [2]	P1	Applies likelihood estimation to received signal.
Averaged Highest Probability	P2	Returns the midpoint of the top 2 likelihoods.
Gridded Highest Probability	GP	Applies likelihoods to an interpolated grid.

Table 1: ***All algorithms and variants.***

3.2 Room-level Metrics

Since many indoor sensor-network applications can operate at the level of rooms, we extended the accuracy and precision metrics to operate at the room-level. For area-based algorithms, our approach is to map the returned area into an ordered set of rooms, where the ordering tells the user which room order to try.

Room Accuracy. This corresponds to the percentage of times the true room, where the object is located, is returned in the ordered set of rooms. An important variation of this metric is the $n-$room accuracy, which is the percentage of times the true room is among the top $n-$rooms.

Room Precision. This corresponds to the average number of rooms returned by the algorithm.

4. RESULTS

We characterized the performance of our area-based algorithms using the metrics described above. We then compared their performance with single-location based approaches. To show that our results are not an artifact of a specific floor, we ran our comparisons using measured data from two distinct buildings.

Comparing Area-based Algorithms. Both the number and location of the training fingerprints are expected to impact localization performance. We experimented with different ways of picking training sets depending on the fingerprints' coordinates. We found that as long as the samples are uniformly-distributed, but not necessarily uniformly-spaced, the specific methodology had no measurable effect on our results. The number of samples has an impact, although it was not as strong as we expected.

Our results showed fundamental tradeoff for tile accuracy and precision; any algorithm that improves tile accuracy worsens precision. Characterizing the distance accuracy CDFs for different training sizes showed that area-based algorithms have comparable distance accuracy in the intermediate percentiles (25%, median and 75%); the differences are most pronounced at the edges of the distribution. On the other hand, the algorithms had a wide variety of precisions.

Comparing All Algorithms. Having shown a wide range of area-based algorithms have similar fundamental performance, we expanded our investigation to point-based algorithms. We compared the CDF of the traditional distance error metric for point-based algorithms, along with the CDF of the median percentile for area-based algorithms.

As shown in Figure 1, the key result was the striking similarity of the algorithms. The CDFs had a similar slope, medians around 10-15ft, and long tails after the 97th percentile. Indeed, many CDFs differ by less than a few feet, and there are regions where they cross. The CDFs of the point-based algorithms showed only marginal improvements in performance as a function of sample size with a sufficient sample density. As a general rule of thumb, a relatively sparse sampling with a density of 1/230 ft^2 (every 15ft) was sufficient coverage for all the algorithms. Reasonable performance was obtainable with much less sampling every 20 ft.

Regarding room-level accuracy, we found similar accuracies across many of the algorithms, with the exception of the Bayesian approaches and at low sampling densities. We also found that the top 3-rooms accuracy for area-based algorithms is significant and useful.

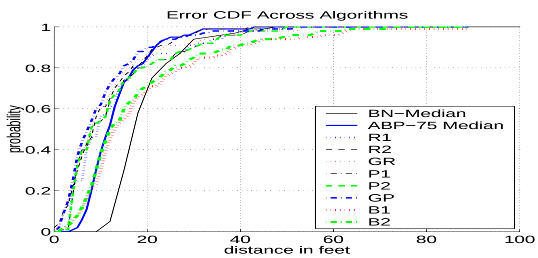

Figure 1: Error CDF across algorithms (115 training samples).

Fundamental Uncertainty. To explore the fundamental uncertainty, we experimented with the BN algorithm since it gives a view of the spatial uncertainty PDF given both measurements and a mathematical model of causal relationships. We generated uncertainty PDFs along both the x, y axes for the testing samples. We found wide distributions in most of the cases, which showed a high degree of uncertainty in the positions. The PDFs from the BN algorithm, along with the very similar performance give very strong evidence that the fundamental uncertainty of all of the algorithms is indeed comparable.

5. CONCLUSION

We characterized the limits of a wide variety of approaches to localization in indoor environments using signal strength and 802.11 technology. We found a median error of 10ft and a 97th percentile of 30ft is an expected bound for the performance of a good algorithm and much sampling. Our results suggest that algorithms based on matching and signal-to-distance functions are unable to capture the myriad of effects on signal propagation. Still, their localization accuracy is significant and useful.

Given that we experimented with large training sets, it is unlikely that additional sampling will increase accuracy. Adding additional hardware and altering the model are the only alternatives.

6. REFERENCES

[1] BAHL, P., AND PADMANABHAN, V. N. RADAR: An In-Building RF-Based User Location and Tracking System. In *INFOCOM* (March 2000).

[2] ROOS, T., MYLLYMAKI, P., AND H.TIRRI. A Statistical Modeling Approach to Location Estimation. *IEEE Transactions on Mobile Computing 1*, 1 (Jan-March 2002).