

## RightSPOT: A Novel Sense of Location for a Smart Personal Object

John Krumm, Gerry Cermak, and Eric Horvitz

Microsoft Research  
Microsoft Corporation  
One Microsoft Way  
Redmond, WA USA  
{jckrumm, gcermak, horvitz}@microsoft.com

**Abstract.** One of the main prerequisites for location-based services is knowledge of location. We present a simple algorithm for computing the location of a device based on signal strengths from FM radio stations. The motivation for this method comes from a new class of smart personal objects that will receive digital data encoded in regular FM radio broadcasts. Given their built-in ability to receive FM and to measure signal strengths, we show how to exploit this ability to measure the device's location. Our algorithm, called RightSPOT, is designed to be robust to manufacturing variations among devices that affect how they measure signal strength. Toward this end, we present a location classification algorithm based not on absolute signal strengths, but on a ranking of signal strengths from multiple FM radio stations. In tests with three devices in six suburban areas, we show that we can correctly infer the device's location about 80% of the time.

### 1 Introduction

One of the promises of ubiquitous computing is to connect users to important information as they move around the world. Our research organization has created a small, low-power device platform named Smart Personal Object Technology (SPOT). The first manifestation of a SPOT device will be a commercially available wristwatch, a prototype of which is shown in

Figure 1. The SPOT device is designed to listen for digitally encoded data such as news stories, weather forecasts, personal messages, traffic updates, and retail directories transmitted on frequency sidebands leased from commercial FM radio stations. The device holds promise for connecting millions of people to valuable notifications and alerts.

The existing method for localizing data for transmission to particular devices is to depend on the limited range of FM radio signals, so that only devices within range of a particular radio tower will get data relevant to that tower's coverage area. Unfortunately, for certain messages, this coarse location resolution is inadequate. Traffic updates, limited time discount offers, and lists of nearby attractions need finer location filtering than that provided by FM radio station coverage areas. One alternative is

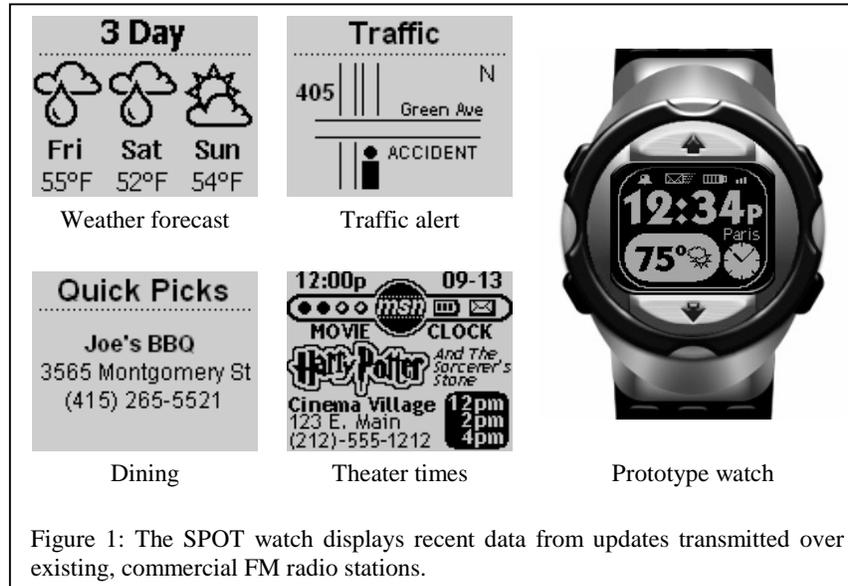


Figure 1: The SPOT watch displays recent data from updates transmitted over existing, commercial FM radio stations.

GPS, but it does not work indoors, adds volume and expense to an already densely packed device, and consumes precious battery power.

Our RightSPOT system embodies a simple method for localizing the device based on signal strengths from existing FM radio stations. The standard SPOT device already contains an FM receiver and the hardware and software necessary to measure signal strength on arbitrary frequencies in the FM band. We show how we can use these signal strengths to localize the device down to a suburb.

The concept of measuring location from radio signal strengths is not new. The RADAR[1] system demonstrated how to localize an 802.11 device based on signal strengths from Wi-Fi access points. Some active badge systems, *e.g.* Krumm *et al.*[2], use radio signal strengths for measuring position. Roos *et al.*[3] presented a probabilistic framework for determining location from radio signal strength. In our case, the attraction of using commercial FM radio is its wide coverage (indoors and outdoors) and the fact that the SPOT device already has hardware and software on board for measuring FM signal strengths. Despite the fact that ubiquitous, commercial radio broadcasts from fixed towers have existed for a long time, we are unaware of other work aimed at trying to use them to infer location.

## 2 Measuring Signal Strength

RightSPOT uses a vector of radio signal strengths taken from different frequencies to identify location. Each time a location is to be inferred, the device scans through a set of FM frequencies and records the signal strength of each one.

A standard SPOT device must be able to scan FM radio stations and measure signal

strength in order to find a sufficiently powerful one transmitting SPOT data. The “radio signal strength indicator” (rssi) comes from an analogue-to-digital converter (ADC) in the device. The raw digital measurements from each frequency are scaled and then averaged over 20 readings for 13 milliseconds. The ADC and associated circuitry are not carefully calibrated to measure rssi in any certain units nor to be consistent from device to device.

The expected inconsistency among devices for measures of rssi provides a challenge, as such variations complicate attempts to generalize for reuse a single mapping between signal strengths and locations. One solution to this problem is to specially calibrate each device using a source of variable, known FM transmission strengths. The result of one of these tests is shown in Figure 2, showing how the rssi readings of a particular device vary with a known transmitted signal strength. The data for this test was taken in a Faraday cage, and the procedure was deemed too costly for mass production.

Another solution is to train each device after its purchase at different locations, gathering signal strength vectors in known places to be used in the same device later. Regular consumers would likely not tolerate such a training regimen.

In addition to manufacturing variations, signal strengths are also affected by the watch’s orientation, its surroundings, and the adjustment of the wrist band which serves as the radio antenna. It would be nearly impossible to anticipate all these variable factors affecting absolute signal strength. If we could anticipate absolute signal strengths, a probabilistic framework like that of Roos *et al.*[3] would be appropriate. However, given the impracticality of discovering each device’s response characteristics, we pursued an alternative method of comparing signal strengths.

Rather than depend on absolute signal strength as an indicator of position, Right-SPOT employs a *ranking* of a set of radio stations in terms of their measured rssi as described in the next section. While we cannot depend on the devices to give consis-

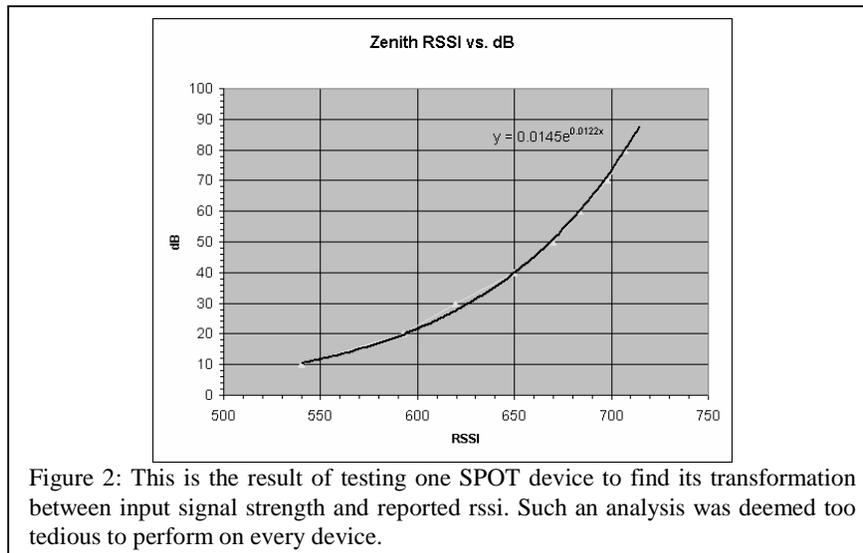


Figure 2: This is the result of testing one SPOT device to find its transformation between input signal strength and reported rssi. Such an analysis was deemed too tedious to perform on every device.

tent absolute signal strengths, they are designed such that the relationship between input signal strength and output rssi is at least monotonically increasing, such that an increase in input signal strength translates to an increase in measured rssi. We leverage this monotonicity in the localization procedure.

### 3 RightSPOT Location Algorithm

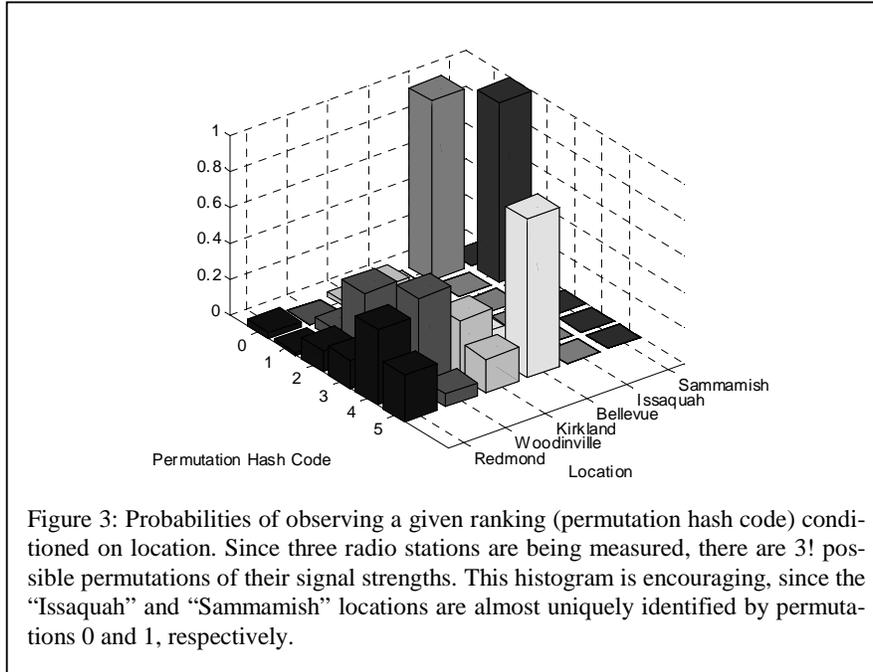
The RightSPOT algorithm infers the location of the device by scanning a list of  $n$  FM radio frequencies,  $\underline{f} = (f_1, f_2, \dots, f_n)$  resulting in a corresponding vector of measured signal strengths  $\underline{s} = (s_1, s_2, \dots, s_n)$ . In our tests, we varied the number of radio stations  $n$  from 2 to 9. A sort routine is used to compute a rank vector of the signal strengths,  $\underline{r} = (r_1, r_2, \dots, r_n)$  in ascending order, where each  $r_i$  gives the rank of the corresponding  $s_i$  in  $\underline{s}$ . For example, if the signal strength vector were  $\underline{s} = (12, 40, 38, 10)$ , the corresponding rank vector would be  $\underline{r} = (2, 4, 3, 1)$ . We note that the rank vector is insensitive to any monotonically increasing function of the elements of  $\underline{s}$ , which makes the algorithm robust to variations in how different devices measure signal strength.

For  $n$  radio stations, there are  $n!$  possible rank vectors, which are the permutations of the integers  $1, 2, \dots, n$ . Each rank vector can be mapped to an integer  $R \in \{0, 1, \dots, n!-1\}$  using a mixed-radix representation of the integers as described by Knuth[4]. Thus we generate a unique hash code for each permutation of signal strengths.

Our classification scheme was motivated by our assumption that different locations will show different relative signal strengths. Ideally each location would map to a single, unique value of  $R$ . In reality, due to noise, caused by such factors as the local tilt and position of the SPOT antenna and local electromagnetic effects cause by buildings and terrain, each location produces a distribution of different  $R$ 's. For training the system, we bring a SPOT device to each of  $L$  locations, gathering hash codes  $R_i^{(l)}$ , where  $l = 1, 2, \dots, L$  indexes over the locations and  $i = 1, 2, \dots, N_l$  indexes over the hash codes observed at location  $l$ . For each location, we can construct a normalized histogram of the hash codes to approximate the discrete probability distribution of hash codes seen at that point,  $p(R|l)$ . An example of these normalized histograms for six locations and three frequencies is shown in Figure 3.

Given the observation likelihoods  $p(R|l)$ , and an observation  $R^*$ , we can compute the probability of being in any of the  $L$  locations using Bayes rule:

$$p(l|R^*) = \frac{p(R^*|l)p(l)}{\sum_{l'=1}^L p(R^*|l')p(l')} \quad (1)$$



Here  $p(l)$  is the *a priori* probability of being at location  $l$ . As we have no prior knowledge of the device’s location, we assume a uniform distribution, setting  $p(l) = 1/L$ . Rather than compute likelihoods, we can directly compare nonnormalized posteriors. The Bayes classifier specifies taking the class with the maximum *a posteriori* probability, *i.e.*

$$l^* = \arg \max_{l=1..L} p(l|R^*) = \arg \max_{l=1..L} p(R^*|l) \quad (2)$$

Algorithmically, this means that for an observation  $R^*$ , we consult the normalized histogram (*e.g.* Figure 3), look up the values of  $p(R^*|l)$  over the full range of locations  $l \in \{1, 2, \dots, L\}$ , and take the location  $l$  with the largest value of  $p(R^*|l)$ .

## 4 Results

To test the RightSPOT algorithm, we chose three SPOT watches at random from our laboratory’s store of test devices. We made no effort to choose watches that gave consistent signal strengths, and we made no effort to calibrate the watches with respect to an absolute signal source nor with respect to each other. This simulates the eventual production environment, considering the economic infeasibility of calibrating each device.

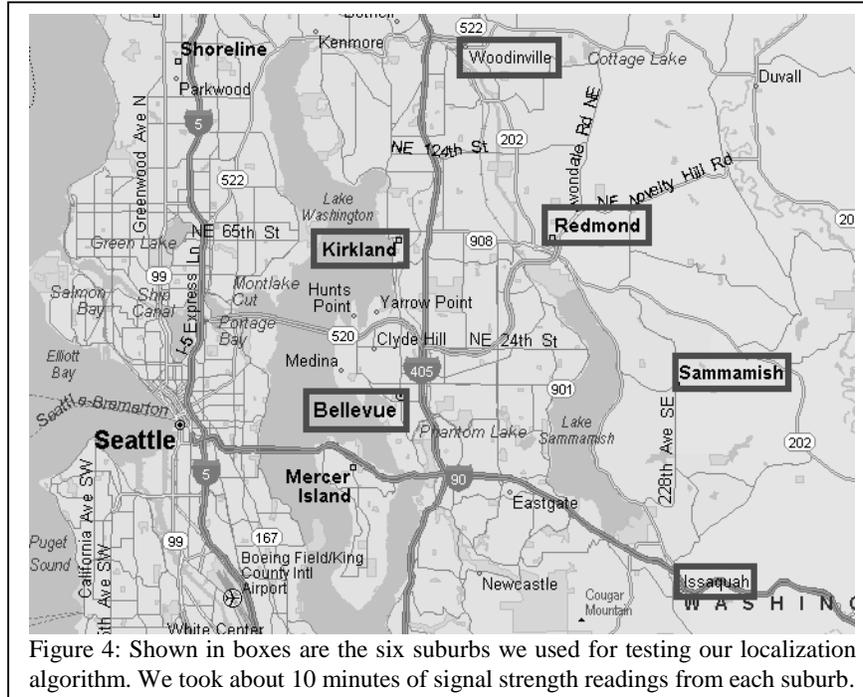


Figure 4: Shown in boxes are the six suburbs we used for testing our localization algorithm. We took about 10 minutes of signal strength readings from each suburb.

We programmed each device to measure signal strengths of 32 different local FM radio stations. In order to download the signal strength in to a PC, the watch had to be mounted in its charger stand, which also provides an RS232 port for data transfer. The RS232 port, in turn, is only active when the charger stand is powered. Thus we put the chargers and watches in the back seat of a car and powered the chargers from the car's battery. We drove to six different suburbs in our area, logging all 32 signal strengths once per second. In each suburb, we took an average of about 620 readings (~10 minutes) while driving around the suburb's retail core. We chose the retail core as compelling applications of this technology will involve retail businesses sending out time-sensitive offers and listings of local attractions. The six test suburbs are shown in Figure 4.

The raw rssi data from the devices is noisy, as shown in Figure 5, so we applied a windowed median filter to the data, replacing each rssi with the median of itself and the preceding 29 unfiltered values.

We tested by alternately picking data from one of the three devices as the basis for the normalized histograms and testing with data from the other two. This was a more realistic test than merely testing each device against itself, because ultimately all the devices will have to depend on one pre-programmed set of histograms for determining their location.

In an effort to minimize the storage and computational burden of location determination, we also experimented with using a much reduced subset of the 32 recorded

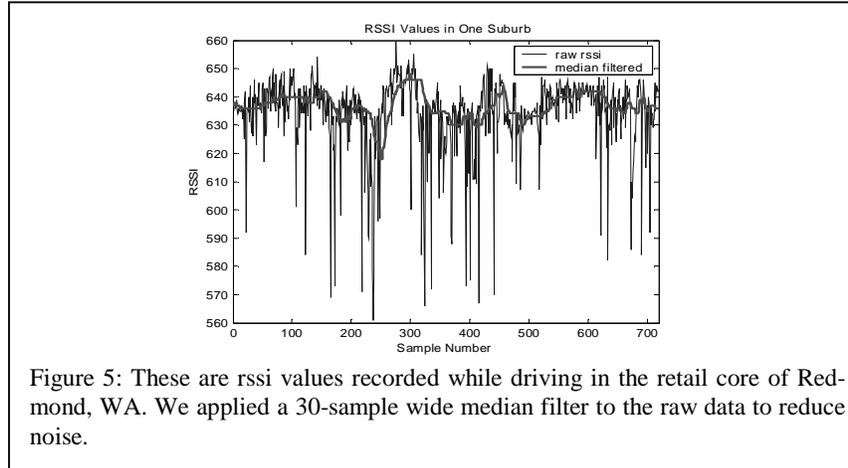


Figure 5: These are rssi values recorded while driving in the retail core of Redmond, WA. We applied a 30-sample wide median filter to the raw data to reduce noise.

radio stations. For each test of  $n$  radio stations, we examined the set of  $\binom{32}{n}$  different combinations of stations to use for classifying location. The results in terms of classification accuracy are shown in Table 1. The reported accuracy is the fraction of correct inferences made over all the tests where one device was held out for making histograms and the other two used for testing. For  $n \leq 5$  we could exhaustively test all possible combinations. For  $n > 5$  we tested a random subset of 10,000 combinations. This explains why the classify accuracy goes down when moving from  $n = 5$  to  $n = 6$  stations: we likely did not find the best combination in the random subset for  $n = 6$ . The best classification accuracy is 81.7% using  $n = 8$  radio stations.

We were pleasantly surprised at how well the simple RightSPOT algorithm could achieve this level of location accuracy using transmitters and receivers that were not intended for providing location information.

## 5 Conclusions

Our study demonstrates the feasibility of using existing FM radio signals to localize a device down to a suburb using a simple algorithm. For the specific focus of our work on supporting SPOT devices, we leveraged existing hardware to measure FM radio signal strengths, so the capability for localization only requires the addition of a small amount of software to an existing configuration. Different devices measure signal strengths differently, and signal strengths are also affected by many other variables. Our Bayesian classification algorithm does not use absolute signal strengths, but instead uses a ranking of signal strengths to help ensure robustness across devices and other variables. We made our tests realistic by making a random choice of devices, by bypassing calibration, and by not testing the individual devices against themselves.

Number of Radio Stations = $n$	Best Subset of FM Radio Stations	Classification Accuracy
2	{KPLU, KEXP}	34.3%
3	{KWFJ, KEXP, KVTI}	61.6%
4	{KWFJ, KEXP, KSER, KVTI}	72.9%
5	{KEXP, KSER, KVTI, KBCS, KLSY}	77.6%
6	{KVTI, KGHP, KRWM, KLSY, KEXP, KSER}	76.2%
7	{KBCS, KEXP, KUBE, KBSG, KVTI, KSER, KLSY}	79.7%
8	{KVTI, KSER, KBCS, KJR, KNHC, KEXP, KBSG, KUBE}	81.7%
9	{KVTI, KRXY, KPLU, KJR, KBCS, KLSY, KUBE, KSER, KEXP}	76.8%

Table 1: Classification accuracy generally increases by including more radio stations in the analysis.

As we move forward with this research, we will study the question of how to anticipate signal strength characteristics of different locations without actually visiting them. One possible solution is to use a commercially available signal strength simulator like RadioSoft's ComStudy software. Such a simulation could also help determine a good subset of radio stations to listen to for best localization. Another interesting problem is how to determine which radio station histograms to transmit to a device, especially if we can make no *a priori* assumptions about where in the world it might be. Finally, we might realize a boost in classification accuracy if we were to smooth position inferences over time, adhering to constraints about how fast devices are expected to move between locations.

## References

1. Bahl, P. and V.N. Padmanabhan. *RADAR: An In-Building RF-Based Location and Tracking System*. in *IEEE INFOCOM 2000*. 2000. Tel-Aviv, Israel.
2. Krumm, J., L. Williams, and G. Smith. *SmartMoveX on a Graph -- An Inexpensive Active Badge Tracker*. in *UbiComp 2002*. 2002. Goteborg, Sweden: Springer.
3. Roos, T., P. Myllymaki, and H. Tirri, *A Statistical Modeling Approach to Location Estimation*. *IEEE Transactions on Mobile Computing*, 2002. **1**(1): p. 59-69.
4. Knuth, D., *Seminumerical Algorithms*, in *The Art of Computer Programming*. 1981, Addison-Wesley: Reading, Massachusetts. p. 64 (Algorithm P).