Statistical Learning Strategies for RF-based Indoor Device-Free Passive Localization

Chenren Xu
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
lendlice@winlab.rutgers.edu

Bernhard Firner
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
bfirner@winlab.rutgers.edu

Yanyong Zhang
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
yyzhang@winlab.rutgers.edu

Richard Howard
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
reh@winlab.rutgers.edu

Jun Li
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
jonjunli@winlab.rutgers.edu

Abstract

In this paper, we present the design, implementation and evaluation of a RF-based device-free passive localization strategy using active RFID nodes. Patterns of the measured power on multiple radio links are used to determine the location of a person in a room in a home environment. We develop an adaptive algorithm and training technique to minimize multi-path effects. With experimental deployment in a 5×8 meters room, we demonstrate that our system can successfully localize an individual to a 30-inch grid square with an 97.2% accuracy and 0.36 meters average error distance.

Categories and Subject Descriptors

C.3 [**Special-Purpose and Application-Based Systems**]: Real-time and embedded systems

General Terms

Algorithm, Experimentation, Measurement

Keywords

RFID, Device-free Passive Localization

1 Introduction

As wireless sensor networks become more pervasive, sensor-based context-aware applications are becoming increasingly available. These systems can automatically collect contextual information in the home or work environment, and this information can improve automation, increase safety, and reduce labor costs. One of the fundamental functions of such applications is localizing people and objects.

Radio frequency based localization techniques have received much attention due to their ubiquity and low cost. RADAR [1] localizes people in indoor environments by requiring people to carry wireless transmitters. To eliminate this limitation, Youssef [3] announces the availability of device-free passive (DfP) localization. In other words, the person can be tracked without carrying a transmitter. Nevertheless, DfP localization in an indoor environment is a very challenging problem, mainly due to the well-known "multipath" effect, which is caused by the reflection and diffraction of the RF signal from objects in the environment. We address this by formulating the localization problem as a linear classification problem, and validate our formulation through extensive experimentation in a one-bedroom apartment.

2 Localization Strategies

Due to multi-path in a cluttered environment, a subject moving across the line-of-sight of a radio link can affect the RF signals in a complicated and unpredictable fashion. The details of multi-path depend upon the position and composition of all objects in the deployed region, and therefore precise calculations are impossible. Hence, we use training data to characterize the room from a statistical perspective. In our approach, we first slice a deployed region into equal-sized cells, and then we localize a subject to a cell. For this purpose, we obtain the training data by collecting the RF signal of each radio link when the subject moves around within each of these cells. Based on this training information, we can determine the cell with the maximum likelihood of containing the subject. We treat all the possible vectors for ra-

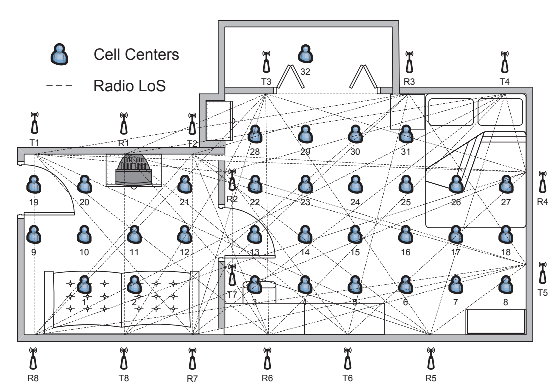

Figure 1. Experimental Topology. The one-bedroom deployment region is partitioned into 32 cells. Eight transmitters and eight receivers are deployed.

SenSys'11, November 1–4, 2011, Seattle, WA, USA.
ACM 978-1-4503-0718-5/11/11

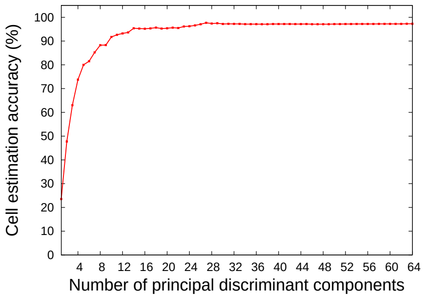

Figure 2. Localization accuracy versus the number of most important principal discriminant components.

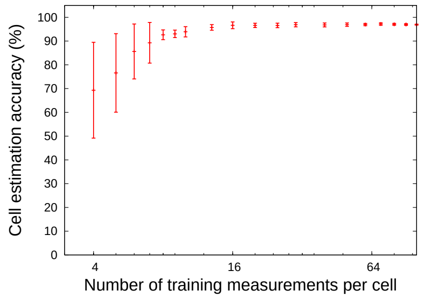

Figure 3. Localization accuracy with 95% confidence interval error bar versus the number of training measurements.

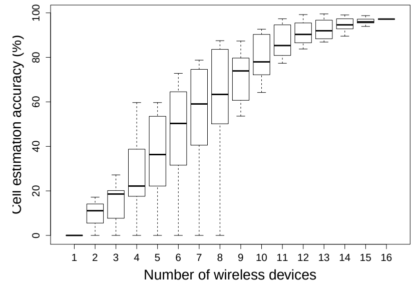

Figure 4. Localization accuracy boxplot versus the number of RF devices that fail to work from all the combinations of RFID devices.

dio signal strength from all the radio links when a subject is located in a cell as a class. We treat each class as a multivariate Gaussian, construct a multi-class training dataset, and use Linear Discriminant Analysis (LDA) as our classification algorithm to solve the indoor localization problem. We use RSS (Received Signal Strength) as our quantitative measurements of a radio link's signal strength. Suppose we have L independent radio links and K cells in the indoor region in which we intend to localize a subject. In the training phase, for each cell k, we collect a set of RSS values with the subject present in this cell. When the subject is in cell k, we collect the RSS measurements from all the links as $[x_{k,1},\ldots,x_{k,L}]$, which composes a RSS vector, $\mathbf{x_k}$ for cell k. At the end of the training phase, a K-class classifier is built based on these training data, and subsequently used to classify the testing subject with unknown class label based on LDA algorithm in the testing phase.

LDA aims to find a linear combination of features which characterize or separate two or more classes [2]. We assume the density function of each class k is multivariate Gaussian $f_k(x)$ and derive the objective function $argmax_k\, f_k(x)$. When we collect the training data for each cell, we make a subject move randomly within the cell rather than stand still. In this way, we average the variance caused by multi-path effect and subject's different orientations. In the testing phase, we make another subject appear randomly in a cell within the deployed region. We collect the RSS data, plug in the data into the K-class classifier and find the class label with the maximum value from objective functions among all the classes. We estimate that cell contains the subject.

3 Experimental Results

Our experimental setup consists of a host PC serving as the system manager and eight RFID transmitters and eight RFID receivers providing 64 radio links. The radios operate at 433.1 MHz. Each transmitter broadcasts a packet with its unique id every 100 milliseconds. The receivers receive the packets, extract the RSS values and forward them to the host PC over USB for data collection and analysis.

The deployment takes place in a one-bedroom apartment with the total area of 5×8 meters, which contains furniture as shown in Figure 1. The room is spatially divided into 32 cells (size of each cell is 30×30 inch). In the training phase, the first author stands in each of these cells and makes 100 RSS measurements for all 64 links. Then, in the testing phase, we have another subject (with different height and weight from the first author) stand in a random cell with a random orientation, which is repeated independently 3200 times. We consider a test successful if the estimated cell is the same as the occupied cell. To evaluate localization performance, we define localization accuracy as the success rate among all the tests, and average error distance as the average distance between the actual location and the center of the estimated cell. Applying our LDA algorithm, we successfully localize an subject to a 30-inch grid square with an 97.2% accuracy and 0.36 meters average error distance.

Furthermore, we have also studied the system performance when we try to reduce the computational overhead, decrease the training time, and with missing devices. Through eigen-decomposition, we select the most important principal discriminant components from both the training data and testing data, and show our localization accuracy in Figure 2. We find that if we are willing to reduce the localization accuracy from 97% to 90%, then choosing the first 10 of 64 discriminant components will be sufficient. Reducing training complexity is important. In Figure 3, we plot the localization accuracy with 95% confidence interval as a function of the number of training measurements per cell. We observe that with only 8 of 100 training measurements in each cell, we can achieve 90% localization accuracy. Minimizing hardware is important as well. We analyze our data by selectively removing all the combinations of devices with corresponding radio links, and show the localization accuracy boxplot as a function of number of RFID devices fail to work in Figure 4. We find that our system can even achieve as high as 90% localization accuracy even if 3 transmitters and 3 receivers are missing.

4 References

[1] P. Bahl and V. Padmanabhan. Radar: an in-building rf-based user location and tracking system. In *IEEE INFOCOM*, 2000.

[2] T. Hastie, R. Tibshirani, and J. H. Friedman. *The Elements of Statistical Learning*. Springer, 2nd edition, 2003.

[3] M. Youssef, M. Mah, and A. Agrawala. Challenges: device-free passive localization for wireless environments. In *ACM MobiCom*, 2007.