ARIADNE: A Dynamic Indoor Signal Map Construction and Localization System

Yiming Ji, Saâd Biaz
Computer Science and Software Engineering
Auburn University
{jiyimin, sbiaz}@auburn.edu

Santosh Pandey, Prathima Agrawal
Electrical and Computer Engineering
Auburn University
{pandesg,pagrawal}@auburn.edu

ABSTRACT

Location determination of mobile users within a building has attracted much attention lately due to its many applications in mobile networking including network intrusion detection problems. However, it is challenging due to the complexities of the indoor radio propagation characteristics exacerbated by the mobility of the user. A common practice is to mechanically generate a table showing the radio signal strength at different known locations in the building. A mobile user's location at an arbitrary point in the building is determined by measuring the signal strength at the location in question and determining the location by referring to the above table using a LMSE (least mean square error) criterion. Obviously, this is a very tedious and time consuming task. This paper proposes a novel and automated location determination method called ARIADNE. Using a two dimensional construction floor plan and only a single actual signal strength measurement, ARIADNE generates an estimated signal strength map comparable to those generated manually by actual measurements. Given the signal measurements for a mobile, a proposed clustering algorithm searches that signal strength map to determine the current mobile's location. The results from ARIADNE are comparable and may even be superior to those from existing localization schemes.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: Network Operations, Miscellaneous

General Terms

Design, Algorithms

Keywords

Localization, indoor radio propagation models, clustering

1. INTRODUCTION

Recently, the demand for wireless communications has grown tremendously. The increasing market for "information anywhere anytime" has been a driving force for increasing advances in mobile wireless communications. Location management and mobility management are critical issues for providing seamless and ubiquitous computing environment for mobile users. For outdoor environments, satellite positioning systems (e.g., the Global Positioning System [1]) offer scalable, efficient, and cost-effective location services that are today available to the large public. Unfortunately, the satellite emitted signals cannot be exploited indoor to effectively determine the location. The design and implementation of a convenient, scalable, and cost-effective indoor location system remains to this day a challenge for the research community.

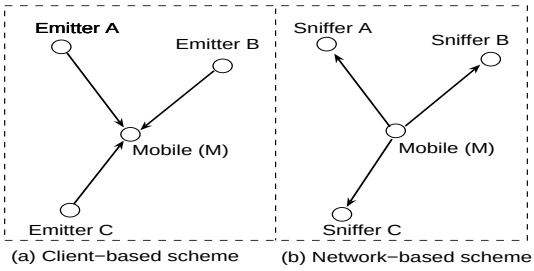

Figure 1: Signal Strength Measurement

Prior to the widespread and popular deployment of RF 802.11 wireless networks, location systems were designed using a specific technology independently from data communication networks. Such location systems exploit infra-red (IR) (*Active Badge* [2]), ultrasound [3, 4, 5], magnetic field [6], or light (cameras) [7]. Such early location systems require specialized hardware used only for the location determination and incur in general a high deployment and maintenance cost. In recent years, the popular success and widespread deployment of RF 802.11 wireless networks enticed many researchers to exploit existing RF 802.11 wireless network infrastructure to build location systems.

Oversimplifying, if the radio propagation signal strength is tightly correlated with the distance between emitter and receiver, then location determination would be a trivial problem that could be solved by one of the following two approaches as illustrated in Figure 1. Figure 1(a) illustrates a *client-based scheme* where three emitters A, B, and C are at known positions. A mobile (client) would listen successively to the three emitters and would measure the signal strength. If the measured signal strength yields the distance from each emitter to the mobile, the location of the mobile reduces to the solution of a simple system of three quadratic equations with two unknowns (assuming the mobile moves in a plan). Note that

MobiSys'06, June 19–22, 2006, Uppsala, Sweden.

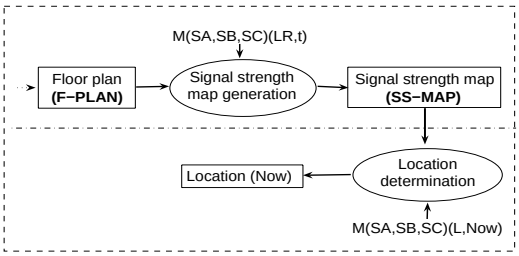

Figure 2: ARIADNE

in this scheme, the client has an active part in the location process: it measures the signals and infers its location. A dual approach, the *network-based scheme*, is illustrated in Figure 1(b): three sniffers at known positions listen to the mobile and measure the signal strength of received packets. It suffices to collect the signal strength measurements (over the network!) from the three sniffers to determine the location using basic calculations. Unfortunately, the relationship between signal strength and distance is not straightforward and is dynamic in nature: even if a mobile does not move, the sniffers in Figure 1(b) will measure the signal strength that varies over time. Moreover, two mobiles that are quite close may generate signals of significantly different strength at the same sniffer. These difficulties make a solution to location determination quite elusive.

In order to address this problem, researchers proposed a two step solution: **First**, establish a signal strength map *SS-MAP* where the signal strength at known and predetermined locations is either manually measured or theoretically estimated, and **Second**, measure signal strength for a mobile at a given location and *SEARCH* the signal strength map *SS-MAP* for the "closest" location that with the best signal strength measurement match. The RADAR [8] system proposed by Bahl and Padmanabhan is exemplary of such an approach: the authors adopt a client-based scheme, collect and, record the radio signal strength received at a mobile (method in Figure 1(a)) from three base stations as a function of location at a selected set of predetermined and known locations. Such records constitute what we call a signal strength map *SS-MAP*. This *measured* signal strength map was used by the authors in two different strategies: (1) the first strategy (they dubbed "empirical method") consists of the mobile sensing the signal strength from the three base stations and *searching* for a record in the measured *SS-MAP* for the best signal strength measurements match; and (2) the second strategy consists of using a simple propagation model to construct an *estimated SS-MAP* that is validated using the *measured SS-MAP*. Estimating is more convenient than measuring a *SS-MAP* especially for a large building. The *estimated SS-MAP* is used the same way as the measured *SS-MAP* in strategy 1. Unfortunately, the authors [8] report that the first strategy (i.e., the "empirical method") outperformed the second strategy that uses the estimated *SS-MAP*.The key weakness of the second strategy is that the radio propagation model results in an estimated *SS-MAP* does not fit well the measured one.

This work proposes a convenient and scalable location system (Please, see Figure 2), dubbed **ARIADNE**.

ARIADNE is a network-based location scheme that requires three sniffers to cover a floor plan as shown in Figure 1(b). Each sniffer listens to a mobile under consideration and measures the average signal strength of packets emitted by the mobile. In the rest of the paper, we refer to the three average signal strength measurements made by the three sniffers as one *signal strength measurement triplet* $M(SA, SB, SC)(L, t)$ where SA, SB, and SC are the signal strengths measured for packets received by respectively sniffers A, B, and C from mobile M at location L and time t. ARIADNE consists of two modules: (1) the first *MAP GENERATION* module estimates a signal strength map **SS-MAP** when given as input a topview CAD floor plan and ***ONE*** signal strength measurement triplet $M(SA, SB, SC)(LR, t)$ for a mobile M located at some reference location LR in the building at some time t, and (2) the second *SEARCH* module determines the location of a mobile M when given as input the estimated signal strength map *SS-MAP* and the current signal strength measurement triplet $M(SA, SB, SC)(L, Now)$ of mobile M at some location L. The contributions of this work address the two modules:

1. **MAP GENERATION**: a sound radio propagation model is developed and validated. The parameters of this model are identified using ray tracing and simulated annealing algorithm. The generation of an accurate signal strength map *SS-MAP* requires only one signal strength measurement triplet.
2. **SEARCH**: A clustering based algorithm is proposed. This clustering based algorithm outperforms, to our knowledge at this time, all search algorithms used so far by the community in the search phase. The accuracy of ARIADNE is, to our knowledge, better than the accuracy reported so far for RF 802.11 based location systems.

Map generation and the location search were extensively tested and validated: simulation results illustrate that signal strength estimates fit well with actual measurements, with a *maximum* average difference around 1.4% of maximum Received Signal Strength Indicator (RSSI), and a *maximum* mean square error MSE around 0.75. With the estimated signal strength map using only *three* sniffers on a floorplan of $45.72m \times 36.57m$, the proposed localization scheme works comparable with most reported localization methods with *maximum* mean error within 3.0 meters and standard deviation below 2.5 meters for a typical office environment.

The remainder of the paper is organized as follows: Section 2 describes previous work done on location estimation over indoor 802.11 networks. Section 3 introduces ARIADNE system. Simulation and experimental comparison are presented in Section 4. Section 5 discusses the performance improvement and mobile user localization. And Section 6 concludes the paper and outlines future research.

2. RELATED WORK

This section will separately survey related work for *map generation* and *location search*.

2.1 Map Generation

Research on indoor radio propagation is an active field. A study of indoor radio propagation characteristics can be found in [9, 10]. A detailed description of earlier radio propagation models can be found in [11]. Based on the ray tracing technique, several statistical models have been analyzed recently [12, 13, 14]. When considering the *large-scale attenuation model*, most researchers model the radio propagation path loss as a function of the *attenuation exponent* n (Please, see Equation 1), which is two for free space but greater than two for an indoor environment.

$$P(d)[dB] = P(d_0)[dB] - 10 \times n \times log_{10}(\frac{d}{d_0}) \qquad (1)$$

where $P(d)$ is the power at distance d to the transmitter in meters; $P(d_0)$ is the power at a reference distance d_0, usually set to 1.0

meter. n is the attenuation exponent, which is often statistically determined to provide a best fit with measurement readings.

Based on the considered parameters in the radio propagation model, all radio propagation models can grossly be grouped into three categories: (1) Simple attenuation model; (2) Partition model; and (3) Site-specific model.

Simple attenuation model is in the form of Equation 1, and it is the base model for the other models. Hills, Schelegel, and Jenkins [14] used this model as part of an automated design tool to estimate the coverage areas for a set of APs.

Different from the simple attenuation model, the *partition model* reduces the pass loss effect from the attenuation exponent by additional consideration of the attenuation effects from the indoor partitions, like walls and floors. Many successful models belong to this group. A couple of famous examples include Phaiboon's statistical model [13], and *wall attenuation factor* model in RADAR [8]. The RADAR system considers attenuation effects from walls along the direct path between the transmitter and the receiver on the same floor. RADAR's location search in the estimated signal strength map *SS-MAP* yields an average resolution of about 4.3m [8].

Site-specific model is similar to the partition model except that it relates to path loss with site-specific parameters (geometrics, materials, and thickness). Two representative models include Hassan-Ali and Pahlavan's probability model [12], and Lott and Forkel's multi-wall-and-floor model [15]. Compared with the other models, the site-specific model does not depend on special assumption, so it works on most general building environment. However, it is complex and requires detailed site-specific parameters.

All these radio propagation models have the following shortcomings:

1. *Tedious* and extensive measurements are required in order to determine the building-specific attenuation exponent and the attenuation coefficients of indoor partitions.
2. The measurements do not consider the *dynamic behavior* of the indoor radio propagation.
3. They only consider the path loss along the *direct path* between the transmitter and the receiver
4. *Detailed material characteristics* and geometry properties are required if site-specific model is to be used.

The above models are not convenient or scalable in real settings. In contrast, ARIADNE requires only ONE signal strength measurement and a topview floor map to estimate the signal strength map (Please, see Figure 2). If the environment is highly dynamic, ARIADNE can be used to monitor a unique point of measurement and generate on demand an updated signal strength map. Simulated annealing (SA) algorithm is used to dynamically determine the attenuation parameters. Consequently, with realtime measurement, the signal-strength map table could dynamically be built.

2.2 Location Search

As pointed earlier, in order to locate a mobile user inside the building, a simple method is to search the *SS-MAP* for the signal strength of the mobile user. If there is a match in the table, the corresponding location is used to denote the mobile user's position. Otherwise, if an exact match is not obtained, the location with closest signal strength to the measurement is selected as the estimate. A general comparison metric is the least mean square error (LMSE).

$$D = min_{k=1}^{N}\{\frac{1}{n}(\sum_{i=1}^{n}(ss_{m,i,k} - ss_{i,k})^2)^{\frac{1}{2}}\} \qquad (2)$$

where D is the least mean square error, N is the total number of records in the signal strength map table, k denotes the k^{th} record in the *SS-MAP* table; n is the number of sniffers. $ss_{m,i,k}$ denotes measured signal strength at sniffer i of the mobile user, and $ss_{i,k}$ is the signal strength record at a sniffer i in *SS-MAP* table. The *nearest neighbors in signal space* method by Bahl and Padmanabhan [16] is essentially this approach.

A problem with *LMSE* is that two or more very different locations could potentially have same signal strength, thus additional processing must be carried out in order to select a more accurate estimate. Therefore, more advanced localization methods are highly desirable. Prasithsangaree, Krishnamurthy, and Chrysanthis [17] proposed a *closeness elimination scheme*. The main purpose is to find more than three locations from the *SS-MAP* table with signal strength close to the measurement. From these, the three closest positions are selected and their position average is used to denote the estimated location for the mobile user. Similarly, in [18], Pandey *et al.* used the second lowest MSE to assist the estimation. They found that if the LMSE and the second lowest mean square error are physically adjacent, then the middle of their locations yields better estimates.

Hatami and Pahlavan[19] also proposed a modified *LMSE* algorithm, dubbed *prioritized maximum power*. This method sorts measured signal strength in descending order for all sniffers so that a contribution priority of each sniffer in the mapping procedure is obtained. According to the priority, the estimates are restricted to a set of reference points. Then *LMSE* or *closeness elimination scheme* is used to determine the final estimates.

Youssef, Agrawala, and Shankar [20] clustered the positions in the *SS-MAP* with the objective of reducing the computation requirement and to improve estimation accuracy. In the method, the cluster is defined as a set of locations sharing a common set of access points (called *cluster key*). Consequently, the *SS-MAP* is sorted according to the cluster keys. To determine the mobile user's location, a small set of access points (with strongest signal strength mapping) are used to determine a cluster for the most probable location. In [21], Agiwal et al. applied the similar idea in the *LOCATOR* system.

In summary, to locate a mobile user, existing signal strength based localization mechanisms assume precise *SS-MAP* with which the *LMSE* technique (and its simple derivatives) is used to find a match. The cluster algorithm from Youssef *et al.* [20] is used to optimize the computation performance and enhance estimates. ARIADNE proposes another modification of *LMSE*, dubbed *clustering-based search method*. This method is similar to Prasithsangaree *et al.*[17] with the difference that the final positions are not necessarily three. The proposed method is specially designed for imprecise *SS-MAP* tables, and is different from Youssef's clustering approach because this method does not sort or cluster reference positions according to common access points, but instead, it select and cluster a set of candidate positions with smaller mean square errors. The largest cluster is chosen and its center is picked as the location estimate.

2.3 Similar Systems

The RADAR system [8, 16], the closest to ARIADNE, proposes an indoor radio propagation model for localization and tracking. Although the system requires extensive measurements and calibration, the achieved localization performance is not satisfactory. The RADAR radio propagation model does not fully capture the multipath phenomenon as it only considers radio propagation along the direct transmission path.

Similarly, Hatami *et al.* [19] used ray tracing software to gener-

ate a reference signal strength map *SS-MAP*. The proposed system uses five APs deployed in a building of 65×48 meter. To locate a mobile user, two different localization method (*LMSE* and *prioritized maximum power*) are evaluated and compared. The results show that the *LMSE* method provides better estimation performance for users within the building. However, *prioritized maximum power* is less susceptible to reference grid resolution and can achieve better estimates when mobile user resides within the vicinity area outside of the building. The results from Hatami (Figure 3 in [19]) show a relation of 10 meters complementary cumulative positioning error with 54% probability for *prioritized maximum power* method. The research by Hatami [19] mainly focused on localization algorithms targeted at intruder detection. It uses ray tracing software to construct signal strength map *SS-MAP* without introducing the indoor radio propagation model.

Different from these systems, this paper introduces ARIADNE, a new indoor localization system. It contains two modules, the first module is map generation, and it includes a new indoor radio propagation model. The second module is search module, and it presents a clustering-based localization algorithm that works on imprecise radio propagation map tables. The radio propagation model is evaluated by comparing estimates against actual signal strength measurements. This paper reports the localization performance of the ARIADNE system, and further compares the proposed localization algorithm with other existing algorithms.

3. ARIADNE

ARIADNE consists of two modules as illustrated in Figure 2 - namely map generation and search - that are developed in Section 3.1 and Section 3.2, respectively.

3.1 Map Generation Module

Map generation includes multiple steps: Subsection 3.1.1 develops the first step that consists of capturing the characteristics of the floor plan and produce a 3-D model necessary for ray tracing. Subsection 3.1.2 explains how ray tracing is used for the determination of the individual ray contribution to the signal strength on a grid of points. A propagation model is proposed in subsection 3.1.3, and its parameters is solved in Subsection 3.1.4 and Subsection 3.1.5 using simulated annealing.

3.1.1 Floor plan interpretation

The main purpose of the floor plan interpretation is to integrate the geometry acquisition process as an automatic procedure. The major task of the interpretation process is to extract the structural parameters from construction CAD files or floor plan image files.

Structural information is extracted from the picture using basic image processing techniques, in which a picture is denoted as a matrix. Each element in the matrix has a value corresponding to the brightness of the pixel at the corresponding position, which is an integer between 0 and 255. The 0 corresponds to black and 255 to white. If the pixel value of the lines in the picture is denoted by 0, then the grouping of a set of connected 0 value pixels, vertically or horizontally, yields a line.

The wall geometry information is constructed by extending the lines vertically in 2D image with base coordinates and the floor height. By stacking the wall information at each floor, overall structural representation of the building is obtained. Similar to most previous research [22], a wall/floor is modelled as a single plane in the middle. The offset between refracted and incident rays is ignored.

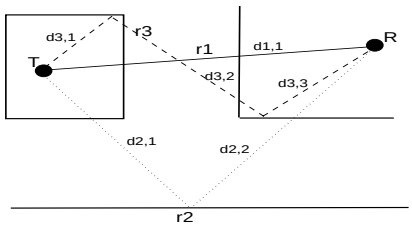

Figure 3: Radio propagation with ray tracing

3.1.2 Ray tracing

Ray tracing (RT) approximates the radio propagation with a finite number of isotropic rays emitted from a transmitting antenna [23]. For an omnidirectional antenna, each ray is assumed to transmit with the same amount of energy at the transmitter, and the energy of the rays will be attenuated at walls or floors due to reflections and transmissions. Ray tracing technique has been widely used to simulate the indoor radio propagation characteristics and to predict the site-specific features of the indoor radio channels [22, 24, 25].

Ray imaging techniques are used to record each ray from the transmitter to the receiver. In the ray imaging technique, the transmitter is assumed to be reflected at each surface around it to produce image transmitters, the reflected rays to the receiver from the real transmitter are considered as direct paths from the mirror images of the true transmitter. Based on *geometrical optics* (GO), each ray from the transmitter to the receiver can be exactly determined. The detailed ray technique is omitted here for lack of space (a good reference can be found in [24, 26, 27]), but instead, several key points of ARIADNE are emphasized.

- Similar to the research by Hassan-Ali and Pahlavan in [12] and Bertoni et al. in [28], the diffraction and scattering effect are neglected in the proposed propagation model because of the minor contribution of the radio in this band;
- Only rays with power above a fixed threshold [29] are considered because highly attenuated rays do not reach the receiver in reality even though a transmission path exists in theory.
- Similar to [12], the multipath power at receiver is determined as the sum of all individual powers regardless of the phase of each path.

Figure 3 depicts a simple scenario where three rays are shown from the transmitter T to the receiver R. Each ray r_i is composed by multiple segments where distance of the j^{th} segment is d_{ij}. Direct path (ray r_1) is denoted by a solid line. The other two paths (ray r_2 and r_3) are indirect and contain reflections. The faint dashed line (ray r_2) has one reflection and dotted line (ray r_3) has two reflections, respectively. The distances traversed by each ray is also depicted in the figure.

3.1.3 Radio propagation model

As explained in Section 3.1.2, the signal power at the receiver is the accumulated multipath power from all individual rays from the same transmitter. For each ray, the attenuation path loss includes three components:

1. The distance-dependent path loss, which is assumed as free space propagation loss;

2. The attenuation due to reflections, which is the product of the reflection coefficient and the total number of reflections from transmitter to the receiver;

3. The attenuation due to transmission, which is the product of the transmission coefficient and the total number of transmission walls.

Consequently, the model is defined as:

$$P = \sum_{i=1}^{N_{r,j}} (P_0 - 20log_{10}(d_i) - \gamma \cdot N_{i,ref} - \alpha \cdot N_{i,trans}) \quad (3)$$

where P is the power (in dB) at receiver, $N_{r,j}$ is the total number of rays received at the receiver j; P_0 is the power (in dB) at a distance of 1 meter; d_i, $N_{i,ref}$, and $N_{i,trans}$ represent the total transmission distance, the total number of reflections and the total number of (wall) transmissions of the i^{th} ray, respectively. γ is the reflection coefficient, and α is the transmission coefficient.

In Figure 3, the transmission distances for three rays (r_1, r_2, and r_3) are $d_{1,1}$, $d_{2,1} + d_{2,2}$, and $d_{3,1} + d_{3,2} + d_{3,3}$, respectively. Ray r_2 has one reflection, and ray r_3 has two reflections. All three rays have two wall transmissions.When starting from transmitter T, all three rays are assumed to hold the same amount of power. With different transmission conditions, the final signal power of each individual ray observed at the receiver R are different. And the overall signal power at the receiver R is the sum of the powers from all received rays.

The site specific parameters (N_{ray}, d_i, $N_{i,ref}$, and $N_{i,trans}$) in Equation 3 can be obtained directly from ray tracing as described in the Section 3.1.2. The other three parameters (P_0, γ, and α), in other similar research, are usually derived from tedious measurements. ARIADNE does not require extensive on site measurements. Instead, simulated annealing (SA) technique is used to determine optimal values for the three parameters of the proposed model. ONE reference measurement only is required.

3.1.4 Parameters Estimation

To estimate the radio propagation parameters (reference power of the ray P_0, reflection coefficient γ, and transmission coefficient α), some measurements at reference positions inside the building are needed. If a maximum of n reference measurements are available, a linear system of $Ax = b$ (derived from equation 3) can be used to determine the three unknowns $x = [P_0\ \gamma\ \alpha]^T$.

To solve the linear equations, the method of least squares could be used. However, it is difficult. As stated earlier, only rays with power above certain threshold are considered in the radio propagation model. Or in other words, from ray tracing simulation, a maximum number of N rays may exist, theoretically, from the transmitter to the receiver. In reality, only n $(n < N)$ rays are actually received because of the different attenuation along each individual path. Since some rays are too weak to contribute the energy at receiver, they must be eliminated from the linear system. Such an elimination process is very difficult at this stage because of the lack of the energy information (again, the *Chicken and Egg Dilemma*). In this research, we use simulated annealing algorithm to search the optimal value of $x = [P_0,\ \gamma,\ \alpha]^T$.

3.1.5 Simulated Annealing Search Algorithm

Simulated Annealing (SA) [30, 31] is a method used to search for a minimum in a general system. It is based on the process of the way a metal cools down to the optimal state (the annealing process). SA's major advantage is an ability of a random search which not only accepts changes that decrease objective function, but also some changes that increase it. Thus, SA method can achieve global optimization without getting trapped at a local minima [32].

The original Metropolis scheme [30] indicates that an initial state of a thermodynamic system is chosen at energy E and a desired temperature T. Holding at that temperature T, the initial configuration is perturbed and the change in energy dE is computed. Applying *Monte Carlo sampling techniques*, the physical annealing process is modelled successfully by computer simulation methods. A convenient formula can be borrowed from thermodynamics:

$$\widehat{P(E)} = exp(-\frac{E}{kT}) \quad (4)$$

which expresses the annealing probability $\widehat{P(E)}$ of a change on energy E at temperature T, where k is Boltzmann's constant.

Given initial values of $x = [P_0\ \gamma\ \alpha]^T$ at a temperature T, the power of each individual ray can be computed (Equation 3). (The initial values can be any positive numbers, however, better values will minimize the search time. Generally, better values can be derived from literature.) Neglecting those rays with power below the threshold, and summing the powers of all others, yield the multipath power at the receiver. The least minimum squared error (Equation 2) allows the comparison of the power estimates fitness with the measurements, and henceforth the adjustment of the parameters of x accordingly.

To adjust the parameters, a random movement is generated by adding a deviate from the Cauchy distribution to each parameter of $x = [P_0\ \gamma\ \alpha]^T$:

$$x_{i+1} = x_i + T \cdot tan(\widehat{P}),\ i = 1, 2, 3 \quad (5)$$

The cooling schedule for the temperature *T* can use a simple method similar to [31]:

$$T_{i+1} = a \cdot T_i,\ a \in (0, 1) \quad (6)$$

Consequently, the Simulated annealing search algorithm can be detailed below:

1) Define initial values for $x = [P_0\ \gamma\ \alpha]^T$.

2) Define the temperature, T_{max} for highest temperature and T_{min} for the cooling down value;

3) Calculate the annealing probability from Equation 4;

4) Update the displacement for the parameters using Equation 5;

5) Calculate the fitness between the estimates and the measurements using equation 2: if a better agreement is obtained, keep the displacement from the above step; else, keep the displacement with certain probability;

6) Update the temperature *T* by equation 6, and repeat steps 3, 4, and 5 until $T < T_{min}$ or specified minimum errors is achieved.

Simulated annealing method can effectively estimate parameter triple $x = [P_0\ \gamma\ \alpha]^T$ with only ONE reference measurement.

3.2 Search module: Clustering-based Search Algorithm

To locate a mobile user, the current user's signal strength measurement triplet is searched from the signal strength map *SS-MAP* for a match. Currently, most search algorithms are based on the *LMSE* and select a single location as the estimate. This method works if a detailed and precise *SS-MAP* for the building is available. As indicated in many papers, the signal strength is observed

to be very dynamic at different measurement times, and to collect a fine-grid signal strength map is time-consuming for large scale building deployments. Consequently, the *LMSE* method will not generate optimal estimates in most circumstances. Therefore, it is difficult to make a decision if this method is to be used exclusively.

ARIADNE proposes a clustering-based search algorithm for the indoor localization of a mobile user based on the following findings:

- ARIADNE constructs fine-grid signal strength map *SS-MAP* based on the radio propagation model and the site-specific geometry of the building;
- The *SS-MAP* from the propagation model provides real-time estimates without further human intervention. However, it is only a close fit to the measurements, or in other words, it is imprecise and small estimation errors are expected for some locations;
- Consequently, *LMSE* may result multiple possible locations in the SS-MAP table, or the unique location corresponding to the LMSE is not necessarily the right position;
- If a set of positions (corresponding to low mean square error with respect to a predetermined threshold) is selected, the positions can be grouped into several clusters. The largest cluster will generally have higher probability to contain the true position for the mobile user.
- The location estimates with the clustering-based search method may provide larger errors for some positions, however, the overall estimation error gets lowered and the confidence is improved.

The clustering-based search algorithm is a two-phase search algorithm. The first phase is named as data collection and cluster preparation phase, and it is introduced in Subsection 3.2.1, where a set of candidate locations with lower mean square error within the threshold are selected and preprocessed with the purpose to neglect isolated locations from the set. The second phase is clustering phase, and it is presented in Subsection 3.2.2, where the remaining candidate locations are grouped into several clusters and the center of the largest cluster is chosen as the final estimate.

3.2.1 Data Collection and Cluster Preparation Phase

In the Data Collection and Cluster Preparation phase, the current signal strength measurement triplet $M(SA, SB, SC)(L, Now)$ of mobile M at some location L is compared with all records from the estimated *SS-MAP*. In stead of selecting only a single location for estimation, ARIADNE select a set of candidate locations according to a predetermined mean square error (MSE) threshold.

Because of the imprecise nature of the estimated *SS-MAP*, some of the selected candidate locations may be scattered around the floor plan. In order to prepare the candidate locations for clustering, the scattered or isolated (unlikely) location points must be detected and omitted from the set of candidate locations. The isolated position is characterized by a larger distance from its location to all other candidate locations. For example, if there are total N candidate locations in a selected location set, let x_i and x_j be two location clusters with m and n candidate locations, respectively, $m, n \in [1, N]$, and $m + n \leq N$; and let $d_{i,j}$ be the minimum pairwise distance from any member instances of these two clusters.

$$d_{i,j} = min(dist(x_{i,r}, x_{j,t})) \quad (7)$$

where r and t represent the position instance in cluster x_i and x_j, respectively; $1 \leq r \leq m$, and $1 \leq t \leq n$. If the candidate location

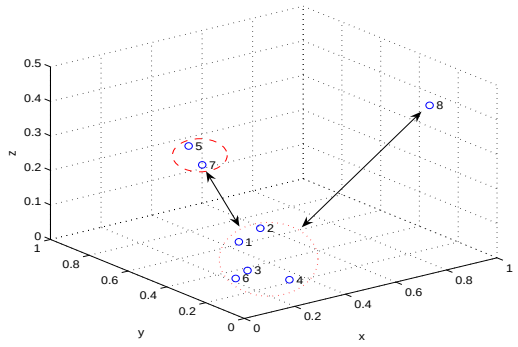

Figure 4: Position isolation example

cluster x_i has larger distance $d_{i,j}$ to every other clusters, it means that cluster x_i is an isolated cluster. If further this cluster contains much smaller populations, it may be omitted from the candidate location set.

Figure 4 shows an example of a set of positions in space. In the figure, position 8 is isolated from all others; positions 5 and 7 are close to each other and they may be treated as one group which is again separated from others. Figure 5 gives the distance information between (group) positions for the data set in Figure 4. If positions of {1,2,3,4,6} are grouped into one cluster, and positions {5,7} form a second cluster, then the minimum distance between these two clusters is 0.3340. If positions of $\{1 \sim 7\}$ are to be grouped into a bigger cluster, and the position 8 is another group, then the minimum distance between them is 0.8311. For the data set in the example, positions of {5,7, and 8} may be neglected during this preparation phase.

3.2.2 Clustering Phase

After the cluster preparation phase, most of the remaining positions have neighbors close to them. Consequently, the main purpose of the clustering phase is to determine the intrinsic grouping of the set for these positions, and to select the right cluster for the estimates.

To group the set of points in space, two common methods are available. The first one is an hierarchical clustering method, and the second one is K-clustering method.

- The hierarchical clustering method [33] produces a hierarchy tree structure of the original data set. The leaves are individual elements and internal nodes are sub-clusters. Each level of the tree represents a partition of the original data set of several sub-clusters. Figure 5 is an example of the hierarchical clustering method.
- K-clustering method searches the best k cluster centroids, and partition the data set by assigning each point to its nearest centroid. K-means clustering [34] is one of the most common K-clustering algorithm.

The clustering procedure is more observable if hierarchical structure of the original data set is obtained (Figure 5). If a minimum of three neighbors are selected ({1,2,6}), it translates to the *closeness elimination scheme* addressed by Prasithsangaree in [17]. And if only two neighbors are chosen ({1,2}), it is the *two closest neighboring scheme* by Pandey [18]. However it is difficult to determine the exact number of neighboring positions that should be selected in

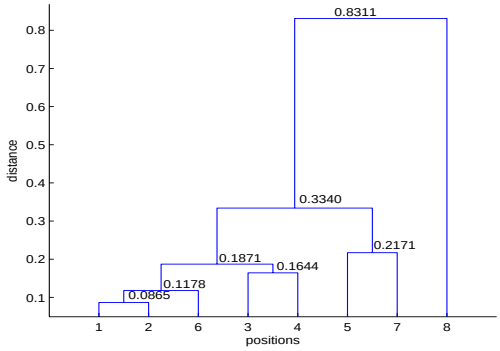

Figure 5: Distance between objects

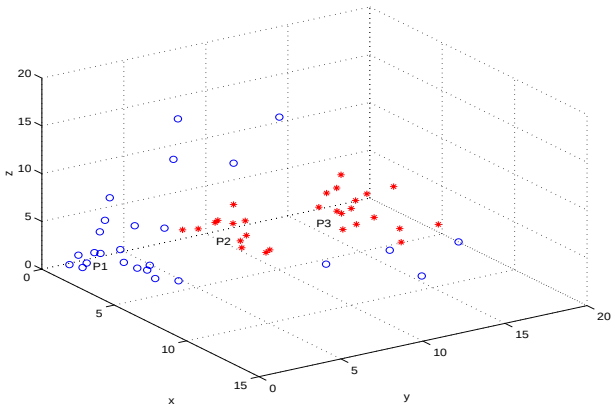

Figure 6: Decision making in tracking mobile user

order to obtain an optimal estimation. Hence, the idea of clustering is extended to incorporate larger number of estimates available by ray tracing. The K-means clustering algorithm is an unsupervised learning method. It starts with randomly selected cluster centers, and the final clustering performance is sensitive to them. Similarly, it is difficult to decide the optimal number of clusters for any given data set.

Unfortunately, there are no general theoretical solutions for these difficulties. This study heuristically explores these problems through simulations. It is found that the strict selection of fixed number of neighboring positions using hierarchical clustering algorithm does not yield satisfying results in most cases. On the contrary, after neglecting the isolated positions in the preparation phase, the k-means algorithm generally yields better estimates.

To determine the actual number of clusters, we select one that provides better separation of the original data as defined in Equation 8.

$$D_c = min(\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{j,i} - \overline{x}_{i,ctr})^2)^{1/2}) \quad (8)$$

where D_c is the total distance of all locations to the respective cluster centers, a small value represents better separation for the determined number of cluster. N is the predetermined number of clusters; M is the maximum number of position points in i^{th} cluster; $x_{j,i}$ denotes the coordinates of j^{th} position in i^{th} cluster; and $\overline{x}_{i,ctr}$ is the center coordinates of i^{th} cluster.

The sensitivity of the selection to initial cluster centers is minimized by running the clustering algorithm multiple times and an averaging result is used.

3.3 Mobility Analysis

To monitor a mobile user, a common method is to maintain and analyze a sliding window with a series of samples of signal strength from the mobile user within a time period. At each signal strength record, the stationary localization method is used individually. Then all location information is connected to give the user's position on a continuous basis.

ARIADNE exploits a similar idea as Abhijit, Ellis, and Fan [35] to track a mobile user in the building. The method is based on the fact that a mobile user does not move arbitrarily within the building. Instead, there is correlation between the current position with the previous location, for example, the distance between two continuous locations can not exceed certain limit. In this clustering-based search algorithm, the largest cluster is selected to denote the possible location of a stationary user. If the clustering history is used, the decision criteria may additionally be restricted with the history information. This may produce better estimates in reality. An example scenario is shown in Figure 6.

In Figure 6, the mobile user's previous location is denoted by P_1. Two candidate current positions are denoted by P_2 and P_3. If (stationary) clustering-based search algorithm is to be used, the location of P_3 should be selected because of the larger population in the clustering group. However, the distance between P_1 and P_3 is beyond the reasonable limit for the mobile user inside the building within the sampling time period. Consequently, the center at P_2 is chosen as the current location of the mobile user.

One shortcoming for this method is that it requires additional memory to maintain a history record. Moreover, the position estimation performance relies on the correctness of the history record. If the initial estimation is not correct, the subsequent position estimation tends to generate unacceptable results. Consequently, a mechanism to automatically reinitiate the localization algorithm is required.

In this research, two principles are used to restart the localization algorithm:

1. If the history constraint results in a cluster with significant smaller populations according to a predetermined threshold;

2. If the history constraint results in a position on the other side of a partition;

4. SIMULATION AND EXPERIMENT

4.1 Experiment Setup

Figure 7 shows the floor plan ($45.72m \times 36.57m$) of the building used for this study. Three sniffers A, B, and C are deployed inside the building with sniffers A and C deployed close to west and east boundaries respectively. Sniffer B is slightly south of the center of the building. Each sniffer was implemented on an IBM T30 ThinkPad running RedHat 9 operating system. Through the AP, sniffers connected with the global monitor, which is used to process the signal strength data. The global monitor also stores the signal strength tables and estimates the users' current location based on the signal strength readings by sniffers A, B, and C.

In order to validate the proposed indoor radio propagation model, the signal strengths were collected at 30 different locations. A Toshiba laptop with Linksys WAP 11 wireless card was used for

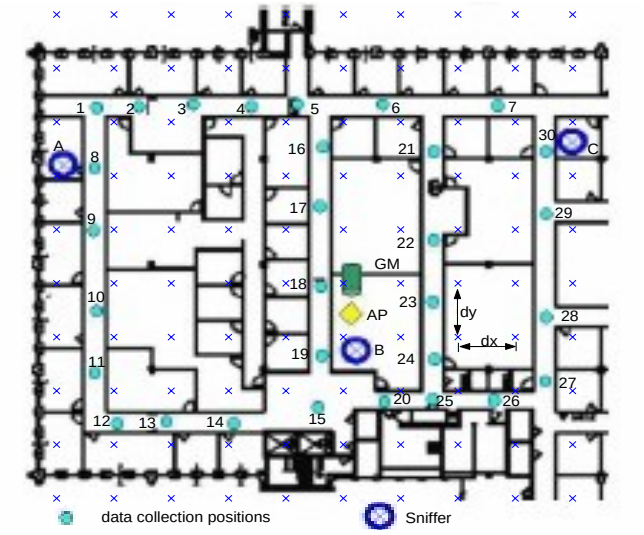

Figure 7: Floor plan with sniffers and data collection positions ('×': reference positions in *SS-MAP* table)

Table 1: Measured signal strength indicator variability over days

	Day1-2	Day1-3	Day1-4	Day1-5	Day1-6
Max dif	5.48	6.77	6.30	7.18	7.66
Avg dif	2.21	2.79	2.46	3.02	2.81
MSE	0.48	0.58	0.55	0.65	0.63

data collection. At each location, about 100 sample packets were emitted at an interval of 0.5 seconds and measured at sniffers A, B, and C. The positions of data collection (and validation) are marked in Figure 7 with faint dots as marked from 1 to 30. The '×' denotes grid positions in estimated *SS-MAP* table. Note that these data validation positions are all in hallway because of the administration reasons, however, none of the validation positions is *directly* connected with a sniffer without a wall partition.

The same series of measurements at the 30 locations were repeated on 6 different days noted in the following as Day_i. ARIADNE radio propagation model (see Section 3.1.3) is evaluated using the 30 signal strength measurement triplets collected in the building for all 6 days. Note that ONE of the 30 signal strength triplets is randomly selected to estimate site specific parameters in the ARIADNE radio propagation model and then to compute an estimated *SS-MAP* table for a specific set of grid positions. Given the estimated *SS-MAP* and the measured signal strength triplet $M(SA, SB, SC)(L, Now)$ for a mobile at some location L, different search localization algorithms were evaluated.

Because of the variation of the environment and the multi-path characteristics, the signal strength measurement changes over time. However, for a single signal strength measurement at a position during the short measurement time-slot, the variation of the measurement value is relatively small. The signal strength measurement results presented in this paper are the average of all 100 samples within the measurement time period (i.e. $0.5s$). The next section additionally compares the signal strength measurements over different days.

4.2 Signal Strength Measurements

The signal strength was collected on six different days at the data collection positions indicated in Figure 7. . Taking Day_1 as reference, Table 1 reports the variability of the signal strength from day to day: each column Day_{1-i} displays the variability between Day_1 and Day_i. The variability is captured using the maximum and average difference, and the mean square error MSE respectively defined as:

$$\begin{cases} \text{max diff} & = max_{i=1}^{n}\{abs(SS_{Day_j,i} - SS_{Day_k,i})\} \\ \text{average} & = mean_{i=1}^{n}\{abs(SS_{Day_j,i} - SS_{Day_k,i})\} \\ MSE & = \frac{1}{n}(\sum_{i=1}^{n}(SS_{Day_j,i} - SS_{Day_k,i})^2)^{\frac{1}{2}} \end{cases} \quad (9)$$

where $SS_{Day_j,i}$ and $SS_{Day_k,i}$ represent the signal strength measurements at location i on j^{th} day Day_j and k^{th} day Day_k respectively; n is the number of SS triplets ($n = 30$).

Table 1 illustrates the dynamic nature of the indoor signal strength over time. Such a variability shows that any search localization method on a static signal strength map *SS-MAP* will in general perform poorly.

4.3 Radio Propagation Model Validation

The parameters $[P_0, \gamma, \alpha]$ of ARIADNE radio propagation model are estimated for a given day Day_i using ONE randomly selected signal strength measurement triplet among the 30 triplets from day Day_i. A signal strength measurement triplet is *randomly selected* because a potential user of ARIADNE may take the reference measurement anywhere in the building. ARIADNE radio propagation model is then evaluated using the 30 signal strength measurement triplets from the same day Day_i. The mean square error MSE between estimates and measurements is used to evaluate the the model fitness.

The influence of the maximum number of reflections and the maximum number of traversed walls on the accuracy of ARIADNE radio propagation model was investigated:

1. As the *number of reflections* increases, the attenuation of the signal increases. After some number of reflections, the contribution of power becomes negligible. Based on the simulations, taking into account more than 3 reflections induces heavy computations without any improvement of the precision or accuracy. This conclusion concurs with other researchers (Valenzuela, Fortune, and Ling[36]), the maximum number of reflections is set to 3;

2. A ray path will traverse at most a limited number of walls. As the ray traverses walls on a direct path, it weakens in power. We call the *maximum number of transmission walls* MW the number of walls traversed before the ray "dies". From simulations in the building considered here, no improvement in precision or accuracy is achieved for MW over 20. Results are reported here for MW taking values 15 and 20 for comparison.

4.3.1 Simulation Results

Extensive simulations were carried out, and the average results of all simulation runs are compared against the 30 signal strength measurement triplets. For each test run, ONE signal strength measurement triplet is randomly selected as a reference among the 30 measurement triplets. Good agreement is obtained between the estimated signal strength map and the measured one.

Typical comparison results are shown in Figure 8 that consists of three plots respectively for sniffers A, B, and C. For each plot, the x-axis represents the 30 positions from the data collection and

Table 2: Radio propagation model verification, maximum RSSI=255

Estimation vs. Measurement		Sniffer A		Sniffer B		Sniffer C	
		Wall:15	Wall:20	Wall:15	Wall:20	Wall:15	Wall:20
2 reflections	Max difference	8.2913	8.3820	15.3067	15.3385	9.4413	7.9786
	Average difference	3.3916	3.3487	5.3976	6.3310	3.9688	3.1910
	MSE	0.7472	0.7472	1.2309	1.3504	0.8525	0.6976
3 reflections	Max difference	8.4514	8.8607	15.4559	13.9691	8.7444	8.2063
	Average difference	3.5720	3.5842	6.0425	6.1516	3.5482	3.1853
	MSE	0.7723	0.7711	1.3343	1.2967	0.7710	0.6886

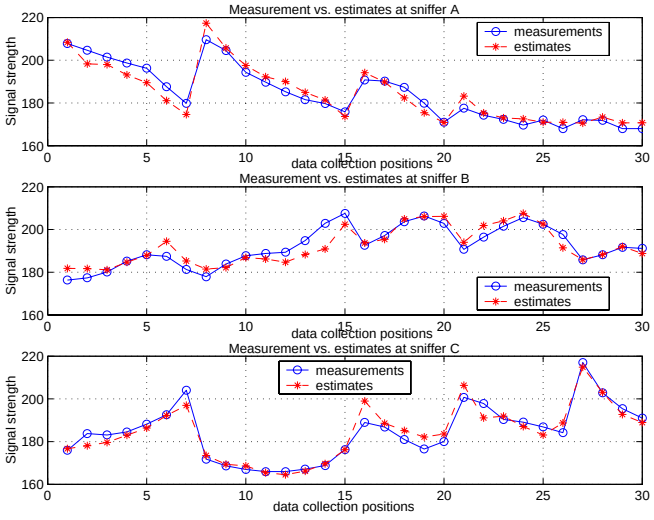

Figure 8: Estimation results and comparison with measurements at data collection positions

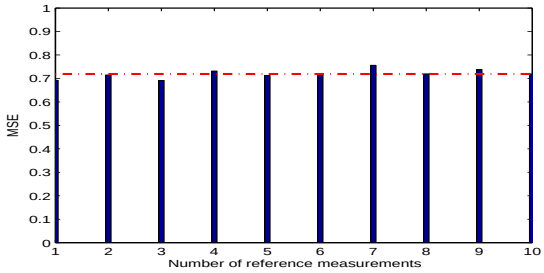

Figure 9: MSE vs. number of reference measurements

the y-axis denotes the signal strength measured as received signal strength indicator (RSSI). The maximum number of reflections is 2 and the maximum number of transmission walls transmission is 20. The points with symbol '$\ominus$' are the signal strength measurements, and the points with symbol '$\times$' are the estimates.

Table 2 reports for Day_1 the difference between estimates and measurements and illustrates the impact on ARIADNE accuracy of the maximum number of reflections (2 or 3) and the maximum number of transmission walls (15 or 20). Each simulation run uses as reference *one* signal strength measurement triplet *randomly* selected within the 30 data collection positions from Day_1. Each number reported in the table is the result averaged over more than 20 simulation runs.

Table 2 shows that results are quite similar for 2 and 3 reflections. This shows that higher order reflection rays marginally affect power estimation accuracy. This conclusion agrees with Valenzuela *et al.*'s results [36]. Consequently in the following test runs, rays are restricted to at most of two reflections.

Table 2 illustrates that the maximum number of transmission walls of 15 or 20 yield close results.

However, Table 2 suggests an interesting observation: estimates for sniffer B are not as good as estimates for sniffers A and C. The difference is systematic and significant. It was observed that the wireless card at sniffer B always provides lower readings than those of sniffer A and C. Despite this, the search localization on the estimated signal strength map *SS-MAP* still works well because the received signal strength from a mobile user at sniffer B is always comparable with B's former readings. So no calibration was done to produce measurement data used in this work.

Finally, Table 2 illustrates that ARIADNE radio propagation model yields good estimates (the *maximum* signal strength difference is within $3\% \sim 5\%$ of the maximum RSSI, see Figure 8): It is worth noting that the variability over days (please, see Table 1) is quite close to the difference between estimates and measurements.

4.3.2 *Number of Necessary Reference Measurements And Location Dependency*

In Section 3.1.5, the simulated annealing searching algorithm is used to estimate the parameter triple of $[P_0, \gamma, \alpha]$ using one reference signal strength measurement triplet. This section addresses the question whether multiple reference measurement triplets would yield estimates that are closer to measurements. The answer is surprising: one reference measurement triplet will yield estimates as good as estimates from 2, 3, or 10 reference measurement triplets.

Figure 9 confirms the findings. The x-axis denotes the number of reference signal strength measurement triplets used for estimating the parameters of the ARIADNE radio propagation model, and the y-axis represents the mean square error MSE of signal strength between measurements and estimates. For each run for x references, x references are *randomly* selected to be used to estimate the radio propagation model and construct the signal strength map. Each point on Figure 9 is an average over 20 runs.

To evaluate the impact of the location of the reference measurement on the performance of the signal strength estimates, the signal strength map is constructed using each individual 30 collected signal strength measurements. Figure 10 provides the average minimum squared error (MSE) for sniffers A, B, and C when using as reference measurement one of the 30 collected signal strength measurements. The x-axis is the location number where the signal strength measurement was made. The y-axis is the average MSE over all sniffers. Results point out that lowest MSE is obtained with

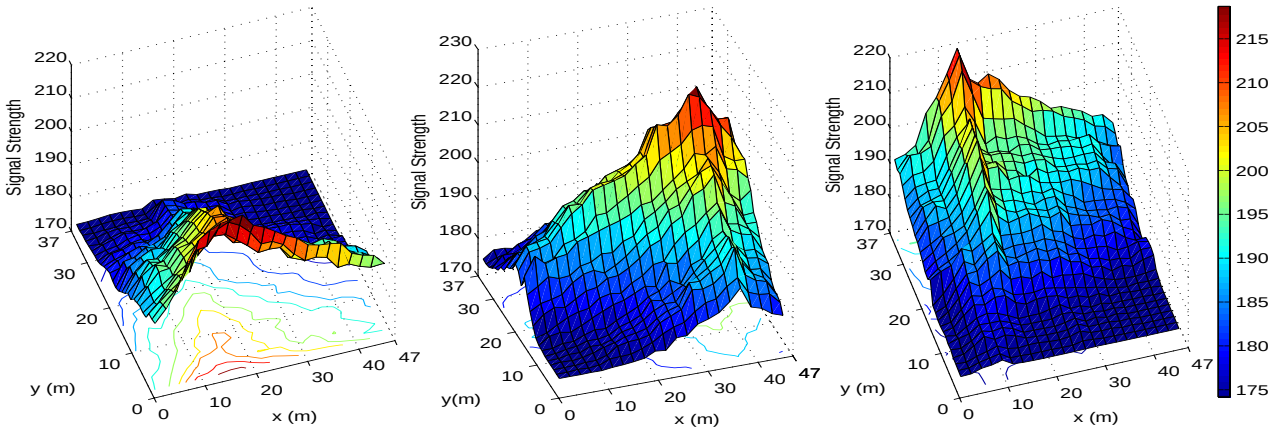

Figure 11: Signal strength shaded surface at sniffers A, B, and C

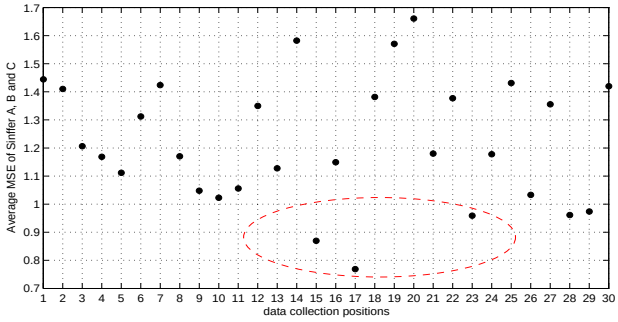

Figure 10: Reference Measurement Selection

signal strength measurements 15, 17, and 23. These measurements were made, as shown in Figure 7, close to the center of the building. This appears to suggest that the reference measurement should be made at the center of gravity of the sniffers.

Note however, for a building with non-uniform walls (i.e. different construction materials and thickness), or with spatially different (human) occupation, it is not appropriate to only select a single measurement at a reference position. In this situation, it is necessary to first identify all representative regions for the building such that each region has homogeneous construction material and in uniform space. Thus a regional SS-MAP can be constructed in advance, and the overall building based SS-MAP can then be stitched.

It is worth noting that ARIADNE radio propagation model is quite accurate even though the sniffers are not deployed in an optimal fashion. The next section addresses the poor deployment of sniffers through a coverage analysis.

4.3.3 Coverage Analysis

This study does not have access to the sniffers and thus cannot modify sniffers' deployment. This work just exploits a data set collected in the work [18]. A gross coverage analysis shows that the sniffers are not optimally deployed.

Figure 11 provides a coverage map for each sniffer (A, B, and C). In each 3D plot, the xy plan is the building floor and the z-axis is the estimated signal strength (RSSI). As the reader can observe, a considerable portion of the area is dark for each sniffer. These dark areas are quite flat as there is not much difference in the signal strength for locations in this area. Figure 11 raises a hope that better results can be achieved with a better deployment of sniffers.

Alternatively, more sniffers could be deployed to provide better coverage for the building. This way, only the three best signal strength measurements are selected for future applications, i.e. if the received signal at one monitor is too weak, the sniffer will not report the signal strength measurement to the global monitor (GM).

4.4 Localization Performance

ARIADNE radio propagation model constructs an imprecise signal strength map *SS-MAP* on a grid of locations. To locate a mobile M "sniffed" as $M(SA, SB, SC)(L)$ at a location L, the signal strength map *SS-MAP* must be *searched* for a best match. This section evaluates the Least Minimum Squared Error, the multiple nearest neighbors, and the proposed clustering-based search techniques. The impact of the grid resolution is also evaluated.

4.4.1 LMSE, multiple nearest neighbors & clustering

This section evaluates the localization performance of *LMSE*, *multiple nearest neighbors*, and *clustering-based* techniques. As explained in Section 2.2 and Section 3.2, *LMSE* picks only the position with LMSE to the "sniffed" signal strength of the mobile user. The scheme of multiple nearest neighbors (*nearest neighbors in signal space, closeness elimination scheme*), as the name suggests, selects multiple closest neighbors and computes the average of these neighbors' positions for the estimates. This work evaluates both 2 and 3 nearest neighbors. The *clustering-based* search method works similarly to multiple nearest neighbors technique, however, it is more flexible in that it does not restrict or fix the number of neighbors. Instead, the clustering-based search algorithm selects a set of candidate positions with signal strength close to the mobile user, and group these positions in space into multiple clusters. Then the algorithm picks the center of the largest cluster as the estimate.

A signal strength map *SS-MAP* is built over a grid of known locations (reference points) based on the proposed radio propagation model. The horizontal and vertical distances between reference

Table 3: Localization performance of six experimental measurements

	Clustering		LMSE		2-N		3-N	
	err	std	err	std	err	std	err	std
Day1	2.8372	2.4304	2.7442	2.0349	3.7355	2.9256	3.5412	3.0458
Day2	2.5330	2.2388	3.5297	2.3543	4.5651	3.5070	4.1878	3.2926
Day3	2.7076	2.1568	3.7510	2.6856	4.1948	2.7037	4.0549	2.6667
Day4	2.9063	2.4727	2.9170	2.5019	2.7875	2.5861	2.6399	2.6080
Day5	3.0004	2.5388	3.6431	2.1429	4.3931	2.5808	3.9705	2.5022
Day6	3.1074	1.7975	3.0704	1.7990	3.5920	2.1638	3.5151	2.1886
Avg	2.8487	2.2725	3.2759	2.2531	3.8780	2.7445	3.6516	2.7173

Table 4: Grid resolution on the performance of Localization

	0.75×1.5 m		1.5×1.5 m		3.0×3.0 m	
	err	std	err	std	err	std
Avg	2.8487	2.2725	3.2861	2.0494	3.4481	3.5999

points are 0.75 meter and 1.5 meter, respectively (see Figure 7). Six different *SS-MAP*s for the 6 days Day_i were constructed to evaluate the localization performance of the three strategies.

With the "sniffed" signal strength triplet as an input, the signal strength map *SS-MAP* is searched using *LMSE*, *multiple nearest neighbors*, or *clustering-based* techniques. Table 3 summarizes the error distance in meters between the real location and the estimated location. The error and standard deviation are reported for each search method for the six days. The last row provides an overall average of the six days. The *clustering-based* localization algorithm in general outperforms all other techniques. For a grid positions of 0.75×1.5 $meter$ apart, and for the floor plan in Figure 7, clustering-based method gives the position estimation with average error of 2.8487 meters. The estimation with *clustering-based* is respectively 14.99%, 36.13% and 28.18% closer than with other techniques.

Note that the localization performance reported in the table 3 is based on a dynamically estimated SS-MAP over a grid of positions with resolution of 0.75×1.5; and the system has only *three* sniffers that are not optimally deployed inside the building (Please, see Section 4.3.3). Thus we anticipate that better localization performance can be obtained if the system is with optimal configurations and finer reference grid resolution (as will be discussed in the next section).

4.4.2 *Impact on Grid Resolution between Reference Points*

This section addresses the question whether a finer resolution grid signal strength map would yield better accuracy for the *cluster-based* search technique. A simple example is provided to explain why finer grid resolution yields better results and simulations confirm this in Table 4. Table 4 provides the error and standard deviation for 3 different grid resolutions (in meter): 0.75×1.5, 1.5×1.5, and 3.0×3.0.

Lower estimation error at finer grid resolution is due to more candidate points in the vicinity of the true position. Figure 12 illustrates the impact of grid resolution. In Figure 12, the true position of the mobile user is denoted by $\diamond$ at point D. If coarse grid resolution is to be used (reference positions at cross points of solid lines), a set of four points ($1 \sim 4$ with sign $\otimes$ in the figure) may be selected as the final cluster group of estimates. The center of this

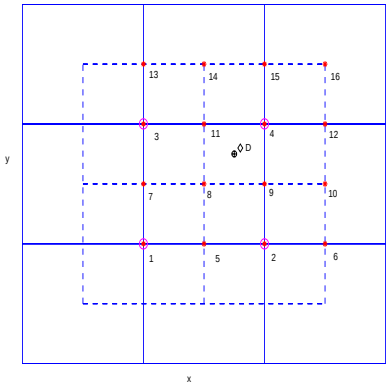

Figure 12: Grid resolution on the performance of clustering localization

cluster is given at position 8. Alternatively, if finer grid resolution is to be used (reference positions at cross points of both solid and dashed lines), then all points in the figure ($1 \sim 16$ with sign * in the figure) may be included in the final cluster. And the center will be in position $\oplus$, which is much closer to the true location of the mobile user. Hence, fine grid resolution yields better localization performance.

4.5 Dynamic SS-MAP Update

A 'stationary emitter' as described in [37] judiciously positioned in the building can be used to be periodically "sniffed" to provide the reference signal strength measurement to dynamically generate a real time signal strength map *SS-MAP*. Such reference device would capture dynamic changes in the environment. The question is whether the map can be computed fast enough to take into account swift changes in the building.

The signal strength map *SS-MAP* table consists a grid of known locations in which the values of corresponding signal strength are stored. Generally, a higher resolution table is required in order to obtain better location estimation. Higher grid resolution induces more computation for the construction of the ray tracing from each point to a set of receivers (APs or sniffers). For example, the approximate time taken to run the ray tracing for 30 positions and 3 sniffers is about 2 hours on a machine of x86 family processor at 1.4GHz with 256 MB physical memory. However, a building floor plan rarely changes. Therefore, ray tracing can be processed once and ray information can be stored in advance. The stored ray information can be fed to ARIADNE to generate a dynamic signal strength map. Given the ray information, the construction of a *SS-MAP* table with 300 points and 3 sniffers takes less than two

Table 5: Mobile vs. static localization

	Mobile	Static
Path 1-7	2.0798	2.8558
Path 8-11	1.7476	2.6384
Path 16-19,15	2.0742	2.3065
Path 21-25	3.3953	4.9681
Path 27-30	1.3847	1.7620
Avg	2.1363	2.9061

minutes. So, a dynamic realtime signal strength map is possible as long as structure conditions in a building remain stable.

5. DISCUSSION

5.1 Mobile User Localization

If a user is mobile, better accuracy can be achieved due to geometric constraints and physical limits (maximum speed, movement patterns along corridors, and users do not step on walls unless drunk!). As in RADAR, mobility helps improve accuracy. Assun a mobile user is moving along a corridor in Figure 7. The distan limit of the mobile user between two continuous locations with a sampling period is no more than 5.0 meters. Experiments wi mobile users were conducted in this work. Table 5 provides the tl localization performance for stationary and mobile users.

The path information in Table 5 is corresponding to the data cc lection positions along the corridor in Figure 7. For example, *Pa 1-7* denotes the scenario of a mobile user moving from position to position 7 along the corridor. The numeric value in the table the average estimation error in meters for all data collection p sitions along the path. The bottom row is the overall localizatic performance for both cases.

5.2 Precision Improvement Method

To improve the localization performance, the following thre problems are important:

1. Accurate signal strength readings from all sniffers: In this work, the readings from sniffer B contain system error (see Section 4.3.1). The performance of the proposed localization scheme should improve if accurate signal strength readings are available;

2. Optimal sniffer deployment: In Section 4.3.3, it is found that deployment of the three sniffers is not optimal. Specially, coverage insensitive areas do exist at corners in the studied floor of the building. To improve this, sniffers A and C may be placed a little closer to the center in y direction (see figure 11);

3. The number of deployed sniffers: In [38], Haeberlen *et al.* indicated that the deployment of more sniffers improves the localization accuracy. Thus most existing indoor systems present decent localization performance with special settings. For example, Ladd et al. [39] deployed 14 sniffers over a floor of $\sim 60m \times 35m$ and reported 1.5 meters accuracy with about 70% probability. Haeberlen et al. [38] used 33 sniffers for a three-story building of 12,000 square meters and estimated with 95% accuracy over 60% cells; and Bahl and Padmanabhan in [16] used 5 sniffers over a plan of $43.96m \times 21.84m$ and obtained the mean error between 3.0m to 5.93m with 90% probability.

The experiment presented in this paper uses *only* three sniffers, and the localization is based on the dynamically *estimated* (instead of measurement) signal strength map table. Therefore, much better localization performance should be achievable for the ARIADNE system when more sniffers are optimally deployed. Also, it shows that the estimation precision actually depends on many site-specific parameters, thus it is inappropriate to evaluate a system only based on the best performance from a specific experiment.

5.3 System Adaptability

To explore the system's adaptability over other environments, the ARIADNE system was also deployed in a basement building (Please, see Figure 13) at Auburn University. As depicted in the figure, room 107, 109, 112 and 111 are classrooms, room 101 and 110 are computer labs, and the rest rooms are offices. Typical offices usually include computers,bookshelves, and metal cabinets. Different from other tested buildings, the experimental basement building also contains a set of construction columns as well as storage/utility rooms. In the experiment, three to five sniffers were deployed to study the sniffer deployment strategy.

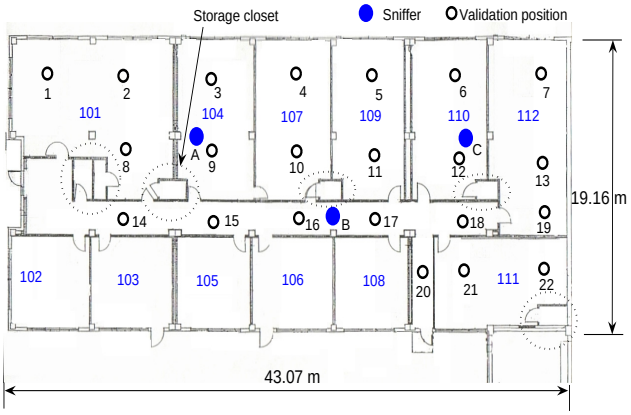

Figure 13: ARIADNE over a basement Building

The experiment was carried out continuously for four months in Auburn University. For all evaluation positions, the estimated signal strength from the proposed indoor radio propagation model fits quite well with measurements. A typical comparison result is given in Table 6, where the mean square error (*MSE*), the standard deviation (*std*), and the maximum relative difference (*Max_dif*) are provided for all three deployed sniffers. The maximum relative difference is calculated as the maximum difference between estimates SS_{est} and the measurements SS_{mea} divided by the maximum measured signal strength values:

$$Max\_dif = \frac{max(abs(SS_{est} - SS_{mea}))}{max(SS_{mea})} \quad (10)$$

With the estimated signal strength map table, the clustering based searching algorithm was used to dynamically search a location for the mobile. With various sniffers deployment strategy using only three sniffers, the localization error of $2.69m \sim 4.22m$ was obtained for the experimental basement building. When more sniffers were deployed, the localization performance improved significantly: the deployment of four sniffers provides an error around $2.30m$, and five deployed sniffers gives an error within $1.92m$.

Table 6: ARIADNE indoor radio signal strength estimation for the basement building

	Sniffer A			Sniffer B			Sniffer C		
	err	std	max_dif	err	std	max_dif	err	std	max_dif
Average	0.8273	0.0851	7.09%	0.9139	0.1454	8.20%	0.8945	0.1930	7.30%

6. CONCLUSION AND FUTURE WORK

This paper introduced a new and automated signal strength estimation tool called ARIADNE. The radio propagation model derived in this paper enables the creation of the signal strength map for an entire building with minimal manual intervention. The scalable algorithm generates a signal strength map with high accuracy and thus can be easily deployed for generating these maps for indoor premises. The time varying nature of the propagation characteristics of the wireless channel poses problems to a signal strength table created manually. This is because, even though such a table is accurate at a given instant of time, it can be rendered useless at another instant. The map generation module presented in this paper enables the creation, on a demand basis, of a signal strength map automatically and almost instantaneously. The resulting map is comparable in accuracy to that of a signal strength map manually generated at that instant of time.

Moreover, on the search module, a clustering-based localization algorithm is developed to search the inaccurate signal strength map *SS-MAP*. The authors argue that the signal strength map, even measurement based, is discrete and is not accurate because of the measurement errors and time-variant property of the radio propagation channel. Consequently, search algorithms such as *LMSE* and *nearest neighbors*, are needed. However, they do not perform as well as the proposed clustering-based algorithm. Simulation results validate the algorithms and procedures used in ARIADNE. In addition, if position history information is to be used, the localization performance for a mobile user is significantly improved.

All actual measurements in this paper are mainly based on our previous work. Results from ARIADNE indicate that these sniffers were not optimally placed. In the client-based signal strength estimation technique, positions of the sniffers play an important role in determining the accuracy and performance of the results. Future work will study these problems in more detail.

7. REFERENCES

[1] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Springer-Verlag, 4th edition, 1997.

[2] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The Active badge Location Systems. *ACM Transactions on Information Systems*, 10(1):91–102, Jan. 1992.

[3] R. Ward, A. Jones, and A. Hopper. A New Location Technique for the Active Office. *IEEE Personal Communications*, 4(5):42–47, Oct. 1997.

[4] A. Ward. *Sensor-driven Computing*. PhD thesis, Cambridge University, U.K., May 1999.

[5] N.B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. *Proceedings of MOBICOM'00*, Aug. 2000.

[6] Ascension technology corporation, http://www.ascension-tech.com.

[7] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale, and S. Shafer. Multi-camera multi-person tracking for EasyLiving. In *Third IEEE International Workshop on Visual Surveillance*, Dublin, Ireland, July 2000.

[8] P. Bahl and V. Padmanabhan. RADAR: An In-Building RF-Based User Location and Tracking System. *Proc. IEEE Infocom 200*, pages 775–784, 2000.

[9] Homayoun Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7), July 1993.

[10] D. Molkdar. Review on radio propagation into and within buildings. *Microwaves, Antennas and Propagation*, 138:61–73, Feb 1991.

[11] Theodore S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, 1996.

[12] M. Hassan-Ali and K. Pahlavan. A new statistical model for site-specific indoor radio propagation prediction based on geometric optics and geometric probability. *Wireless Communications, IEEE Transactions on*, 1:112–124, Jan. 2002.

[13] Supachai Phaiboon. An Empirically Based Path Loss Model for Indoor Wireless Channels in Laboratory Building. In *Proceedings of IEEE TENCON'02*, 2002.

[14] Alex Hills, Jon Schlegel, and Ben Jenkins. Estimating Signal Strengths in the Design of an Indoor Wireless Network. *IEEE Trans. on Wireless Communications*, 3(1), Jan. 2004.

[15] M. Lott and I. Forkel. A multi-wall-and-floor model for indoor radio propagation. *Vehicular Technology Conference, IEEE*, 1:464–468, May 2001.

[16] Paramvir Bahl, V.N. Padmanabhan, and A. Balachandran. Enhancements to the RADAR User Location and Tracking System. Technical report, Microsoft Research, WA 98052, Feb. 2000.

[17] P. Prasithsangaree, P. Krishnamurthy, and P.K. Chrysanthis. On indoor position location with wireless LANs. In *Personal, Indoor and Mobile Radio Communications, 2002. The 13th IEEE International Symposium on*, pages 720–724, Sept. 2002.

[18] Santosh Pandey, B. Kim, F. Anjum, and P. Agrawal. Client assisted location data acquisition scheme for secure enterprise wireless networks. In *IEEE WCNC 2005*, 2005.

[19] Ahmad Hatami and Kaveh Pahlavan. In_building Intruder Detection for WLAN Access. *Position Location and Navigation Symposium, 2004, PLANS 2004*, pages 592–597, April 2004.

[20] M.A. Youssef, A. Agrawala, and A. Udaya Shankar. WLAN location determination via clustering and probability distributions. *Pervasive Computing and Communications, 2003. (PerCom 2003). Proceedings of the First IEEE International Conference on*, March 2003.

[21] Ankur Agiwal, Parakram Khandpur, and Huzur Saran. Locator: location estimation system for wireless lans. In *Proceedings of the 2nd ACM international workshop on Wireless mobile applications and services on WLAN hotspots*, Oct. 2004.

[22] Zhong Ji, Bing-Hong Li, Hao-XIng Wang, Hsing-Yi Chen, and Tapan K. Sarkar. Efficient Ray-Tracing Methods for Propagation Prediction for Indoor Wireless Communications. In *IEEE Antennas and Propatation Magazine*, volume 43, April 2001.

[23] H. Kim and H. Ling. Electromagnetic scattering from an inhomogeneous object by ray tracing. *IEEE Trans. Antennas Propagat.*, 40:517–525, May 1992.

[24] Reinaldo A. Valenzuela. Ray tracing prediction of indoor radio propagation. In *Personal, Indoor and Mobile Radio Communications, 1994.* 5th IEEE International Symposium on Wireless Networks - Catching the Mobile Future., Sept. 1994.

[25] Chang-Fa Yang, Boau-Cheng Wu, and Chuen-Jyi Ko. A Ray-Tracing Method for Modeling Indoor Wave Propagation and Penetration. *IEEE Transaction on Antennas and Propagation*, 46(6), June 1998.

[26] A. Falsafi, K. Pahlavan, and G. Yang. Transmission Techniques for Radio LAN's - A Comparative Performance Evaluation Using Ray Tracing. *IEEE J. on Sel. Areas in Comm.*, 14(3):477–491, April 1996.

[27] M. Lott. On the performance of an Advanced 3D Ray Tracing Method. In *Proc. of European Wireless & ITG Mobile Communication*, Oct. 1999.

[28] Henry L. Bertoni, Walter Honcharenko, L.R. Maciel, and H. Xia. UHF propagation prediction for wireless personal communications. *Proc. IEEE*, 82:1333–1359, Sept. 1994.

[29] G. German, Q. Spencer, L. Swindlehurst, and R. Valenzuela. Wireless indoor channel modeling: statistical agreement of ray tracing simulations and channel sounding measurements. *Acoustics, Speech, and Signal Processing, 2001. Proceedings, IEEE Int. Conf. on*, 4, May 2001.

[30] N. Metropolis, A.W. Rosenbluth, M. N. Rosenbluth, A.H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1958.

[31] K.A. Dowsland. Simulated annealing. In C. R. Reeves, editor, *Modern Heuristic Techniques for Combinatorial Problems, chapter 2. McGraw-Hill Book Company, Berkshire*, 1995.

[32] S. Kirkpatrick, C. D. Gelatt, and M.P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598), May 1983.

[33] S.C. Johnson. Hierarchical Clustering Schemes. *Psychometrika, 2:241-254*, 1967.

[34] J. B. MacQueen. Some Methods for classification and Analysis of Multivariate Observations. In *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, pages 1:281–297. Berkeley, University of California Press, 1967.

[35] Vijay Abhijit, Carla Ellis, and Xiaobo Fan. Experiences with an Inbuilding Tracking System. In *PIMRC*, 2003.

[36] R.A. Valenzuela, S. Fortune, and J. Ling. Indoor propagation prediction accuracy and speed versus number of reflections in image-based 3-D ray-tracing. *48th IEEE Vehicular Technology Conferernce*, 1:539–543, May 1998.

[37] S. Ganu, A.S. Krishnakumar, and P. Krishnan. Infrastructure-based Location Estimation in WLAN Networks. *IEEE Wireless Communications and Networking Conference (WCNC 2004)*, 1:465–470, March 2004.

[38] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach, and L. E. Kavraki. Practical Robust Localization over Large-Scale 802.11 Wireless Networks. In *MobiCom*, Sept 2004.

[39] Andrew M. Ladd, Kostas E. Bekris, Algis Pl Rudys, G. Marceau, Lydia E. Kavraki, and Dan S. Wallach. Robotics-Based Location Sensing Using Wireless Ehternet. In *MobiCom'02*, Sept. 2002.