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.2.1 [Computer System Organization]: Network Architecture and Design - Distributed networks, Wireless communications;

General Terms
Algorithms, Design, Experimentation

Keywords
Indoor localization, 802.11, radio propagation models

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.

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 the 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 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 in 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, 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.

The remainder of the paper is organized as follows: Section 2 describes previous work done on location estimation over in-
door 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 [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.

$$P(d)[dB] = P(d_0)[dB] - 10 \times n \times \log_{10} \left( \frac{d}{d_0} \right)$$  \hspace{1cm} (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. With point-by-point measurement, Hills et al. report that an attenuation exponent of 2.60 yields the best fit in the buildings on the Carnegie Mellon University campus. A difference of 3.0 dB between the measurements and estimates is achieved in most cases.

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]. Phaiboon’s model considers multiple floor environment. The test results show the estimated signal strength from the partition model agrees better than that of simple attenuation model [13]. In contrast, 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]. Hassan-Ali et al. compared the estimated signal strength with measurements using a probability model. The results of mean error of 2.77dB and standard deviation of 2.87dB are obtained [12]. 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).

$$D = \min_{k=1}^{N} \left\{ \frac{1}{n} \left( \sum_{i=1}^{n} (ss_{m,i} - ss_i)^2 \right)^{\frac{1}{2}} \right\}$$  \hspace{1cm} (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 sniffer records, $ss_{m,i}$ denotes measured signal strength at sniffer $i$ of the mobile user, and $ss_i$ 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 approach because it 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 generate 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. He 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 include 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. While original construction CAD drawings provide detailed information for the whole building, ARIADNE, however, only requires the floor plan for each floor of a building.

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 modeled as a single plane in the middle. The offset between refracted and incident ray is ignored.

Figure 3 shows a site floor plan where ARIADNE was tested. The floor is about 150×120 ft.. Figures 4(a) and 4(b) display respectively the 2D topview and the 3D view with height information extracted from the original plan. Geometry information of the walls and floor/ceiling is stored in a database table for ray tracing. The floor plan in Figure 3 includes about 958 walls in total.

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 is widely used to simulate the radio propagation in indoor environment [24, 25, 22].

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.

- The diffraction and scattering effect are neglected in the proposed propagation model because of the minor contribution of the radio in this band [28, 12];
- 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.
- The multipath power at receiver is determined as the sum of all individual powers regardless of the phase of each path [12].

Figure 5 depicts a simple scenario where three rays are shown from the transmitter $T$ to the receiver $R$. Each ray $r_j$ 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:

\[ P = \sum_{i=1}^{N_{r,j}} (P_0 - 20 \log_{10}(d_i) - \gamma \cdot N_{i,ref} - \alpha \cdot N_{i,trans}) \] (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. \( \gamma \) is the reflection coefficient, and \( \alpha \) is the transmission coefficient.

In Figure 5, 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 (\( P_0, \gamma \), and \( \alpha \)), 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 measurement only is required.

### 3.1.4 Parameters Estimation

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

Let \( P_j \) denote a reference measurement at location \( j \) (0 < \( j \leq n \)), and let \( N_{r,j}, N_{j,ref}, \) and \( N_{j,trans} \) be the total number of rays received by the receiver, total number of reflections and total number of transmissions, separately, matrix \( A \) and \( b \) are defined as:

\[
A = \begin{bmatrix}
N_{r,1} - \sum_{i=1}^{N_{r,1}} (N_{i,ref}) & - \sum_{i=1}^{N_{r,1}} (N_{i,trans}) \\
\vdots & \vdots \\
N_{r,j} - \sum_{i=1}^{N_{r,j}} (N_{i,ref}) & - \sum_{i=1}^{N_{r,j}} (N_{i,trans}) \\
\vdots & \vdots \\
N_{r,n} - \sum_{i=1}^{N_{r,n}} (N_{n,ref}) & - \sum_{i=1}^{N_{r,n}} (N_{n,trans})
\end{bmatrix} \] (4)

\[
b = \begin{bmatrix}
P_0 + 20 \sum_{i=1}^{N_{r,1}} (\log_{10}(d_i)) \\
\vdots \\
P_0 + 20 \sum_{i=1}^{N_{r,n}} (\log_{10}(d_n))
\end{bmatrix}
\] (5)

To solve the linear equations, the method of least squares could be used. However, there is a problem. As it is stated in Section 3.1.2, only the rays with power above a 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 \( n < N \) 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 in the system of Equations 4 and 5. Such an elimination process is difficult and time consuming because of the lack of the energy information for each individual ray at this stage. Alternatively, a simulated annealing search algorithm can be used to find 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 optimum 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 connection between SA algorithm and mathematical minimization was first introduced by Pincus [33]. Later, Kirkpatrick, Gelatt, and Vecchi [34] provided the basis of an optimization technique for combinatorial (and other) problems.

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 modeled successfully by computer simulation methods. A convenient formula can be borrowed from thermodynamics:

\[
P(E) = \exp\left(-\frac{E}{kT}\right)
\] (6)

which expresses the annealing probability \( P(E) \) 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 pow-
ers 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 \( x = [P_0 \gamma \alpha]^T \):

\[
x_{i+1} = x_i + T \cdot \tan(\hat{P}), \quad i = 1, 2, 3
\]

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

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

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, \( T_{\text{max}} \) for highest temperature and \( T_{\text{min}} \) for the cooling down value;
3) Calculate the annealing probability from Equation 6;
4) Update the displacement for the parameters using Equation 7;
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 8, and repeat steps 3, 4, and 5 until \( T < T_{\text{min}} \) or specified minimum errors is achieved.

Simulated Annealing method can effectively estimate parameter pair \( x = [P_0 \gamma \alpha]^T \) with only one reference measurement. Generally, more reference measurements are assumed to provide better average parameter estimates. This is not true as shown in Section 4.3.2.

### 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 hit. 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 time, and to collect a fine-grid signal strength map is time consuming for large scale building deployments. Consequently, \( LMSE \) method may give good position estimates in certain cases, but unacceptable results for many others. 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 select 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 may form several clusters and the largest cluster generally contains the true position of 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 this 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 \( x_1 \) and \( x_2 \) 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.

\[
d_{i,j} = \min(\text{dist}(x_i,r, x_j,t))
\]

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 cluster \( x_1 \) has larger distance \( d_{i,j} \) to every other clusters, it means that cluster \( x_1 \) is an isolated cluster.
If further this cluster contains much smaller populations, it may be omitted from the candidate location set.

Figure 6 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 7 gives the distance information between (group) positions for the data set in Figure 6. 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 [35] 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 7 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 [36] 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 7). 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 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 cluster 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, it is found that a number of 3 or 4 clusters with 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 10.

\[
D_c = \min \left( \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{M} (x_{j,i} - \overline{x}_{i,ctr})^2 \right)^{1/2}
\]  

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 [37] 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 8.

In Figure 8, 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 prefect, the following position estimation tends to generate unacceptable results. Consequently, a mechanism to automatically reinitialize 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 9 shows the floor plan 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. Sniffers are connected to a computer, called global monitor, that processes and analysis 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 reported by sniffers $A$, $B$, and $C$.

The signal strengths were collected at 30 different locations in order to validate the ARIADNE radio propagation model. A Toshiba laptop with Linksys WAP 11 wireless card was used for 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 are marked in Figure 9 with faint dots as marked from 1 to 30. The ‘×’ denotes grid positions in estimated SS-MAP table. The same series of measurements at the 30 locations were repeated on 6 different days noted in the following as Dayi. ARIADNE radio propagation model (see Section 3.1.3) is evaluated using the 30 signal strength measurement triplets collected in the building in 6 different days. Note that ONE of the 30 signal strength triplets is randomly selected to estimate ARIADNE radio propagation model and 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.

4.2 Measurements

The signal strength was collected on six different days at the data collection positions indicated in Figure 9. Taking Day1
as reference. Table 1 reports the variability of the signal strength from day to day: each column Day1-⋯ displays the variability between Day1 and Dayy. The variability is captured using the maximum and average difference, and the mean square error MSE respectively defined as:

\[
\begin{align*}
\text{max diff} & = \max_{i=1}^{n} \{ \text{abs}(SS_{\text{Day},i} - SS_{\text{Day},i}) \} \\
\text{average} & = \text{mean}_{i=1}^{n} \{ \text{abs}(SS_{\text{Day},i} - SS_{\text{Day},i}) \} \\
\text{MSE} & = \frac{1}{n} \left( \sum_{i=1}^{n} (SS_{\text{Day},i} - SS_{\text{Day},i})^2 \right)^{\frac{1}{2}}
\end{align*}
\]

where \( SS_{\text{Day},i} \) and \( SS_{\text{Day},i} \) represent the signal strength measurements at location \( i \) on \( j \)th day \( \text{Day}_j \) and \( k \)th day \( \text{Day}_k \) respectively; \( n \) is the number of signal strength triplets (here 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 \( \text{Day}_j \) using ONE randomly selected signal strength measurement triplet among the 30 triplets from day \( \text{Day}_j \). 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 \( \text{Day}_j \). The mean square error MSE between estimates and measurements is used to evaluate 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[38]), the maximum number of reflections is set to 3;

2. In the building considered here, a direct path will traverse at most between 15 and 20 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, no improvement in precision or accuracy is achieved for \( MW \) over 20. Results are reported here for \( MW \) taking values 15 and 20.

### 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 between the estimated signal strength map and the measured one.

Typical results are shown in Figure 10 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 the \( y \)-axis denotes the signal strength measured as received signal strength indicator (RSSI). Each point is an average result of 5 test runs (each test run uses a different randomly selected measured reference triplet). The maximum number of reflections is 2 and the maximum number of transmission walls transmission is 20. The points with symbol ‘⊙’ are the signal strength measurements, and the points with symbol ‘×’ are the estimates.

Table 2 reports for \( \text{Day}_1 \) 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 \( \text{Day}_1 \). Each number reported in the table is the result average 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 conclu-

<table>
<thead>
<tr>
<th></th>
<th>Day1-2</th>
<th>Day1-3</th>
<th>Day1-4</th>
<th>Day1-5</th>
<th>Day1-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max difference</td>
<td>5.48</td>
<td>0.77</td>
<td>6.30</td>
<td>7.18</td>
<td>7.66</td>
</tr>
<tr>
<td>Average difference</td>
<td>2.21</td>
<td>2.79</td>
<td>2.46</td>
<td>3.02</td>
<td>2.81</td>
</tr>
<tr>
<td>MSE</td>
<td>0.48</td>
<td>0.58</td>
<td>0.55</td>
<td>0.65</td>
<td>0.63</td>
</tr>
</tbody>
</table>
Table 2: Radio propagation model verification, maximum RSSI=255

<table>
<thead>
<tr>
<th>Estimation vs. Measurement</th>
<th>Sniffer A</th>
<th>Sniffer B</th>
<th>Sniffer C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average difference</td>
<td>3.3916</td>
<td>3.4387</td>
<td>5.3976</td>
</tr>
<tr>
<td>MSE</td>
<td>0.7472</td>
<td>0.7472</td>
<td>1.2309</td>
</tr>
<tr>
<td>3 reflections Max difference</td>
<td>8.4514</td>
<td>8.5607</td>
<td>15.4559</td>
</tr>
<tr>
<td>MSE</td>
<td>0.7723</td>
<td>0.7711</td>
<td>1.3343</td>
</tr>
</tbody>
</table>

Figure 11: MSE vs. number of reference measurements

Figure 12: Reference Measurement Selection

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 pair of \([F_0, \gamma, \alpha]\) 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 11 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 11 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 12 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 signal strength measurements 15, 17, and 23. These measurements were made, as shown in Figure 9, 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.

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.
This section evaluates the localization performance of grid resolution and also evaluates proposed clustering-based search techniques. The impact of the Squared Error, the multiple nearest neighbors, and the proposed clustering-based search technique on a best match. This section evaluates the Least Minimum Squared Error (LMSE) for a best match. This section evaluates the Localization Algorithm (ARIADNE) radio propagation model constructs an imprecise signal strength map SS-MAP on a grid of locations. To locate a mobile "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 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. 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. Clustering-based search method works similarly to multiple nearest neighbors technique, however, it is more flexible in that it does not restrict the number of neighbors. Instead, clustering-based search algorithm selects a set of reference 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 points are 0.75 meter and 1.5 meter, respectively (see Figure 9). Six different SS-MAPs for the 6 days $Day_{6}$ were constructed to evaluate the localization performance of the three strategies.

![Figure 13: Signal strength shaded surface at sniffers A, B, and C](image)

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 x 1.5 meter apart, and for the floor plan in Figure 9, 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.

#### 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): $0.75 \times 1.5$, $1.5 \times 1.5$, and $3.0 \times 3.0$.

![Table 4: Grid resolution on the performance of Localization](image)

Lower estimation error at finer grid resolution is due to more candidate points in the vicinity of the true position. Figure 14 illustrates the impact of grid resolution. In Figure 14, the true position of the mobile user is denoted by $\circ$ 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 ~ 4 with sign $\circ$ in the figure) may be selected as the final cluster group of estimates. The center of this cluster is given at position 8. Alternatively, if finer grid resolution is to be

![Figure 14: Coverage map for each sniffer (A, B, and C)](image)
used (reference positions at cross points of both solid and dashed lines), then all points in the figure (1 ~ 16 with sign *) may be included in the final cluster. And the center will be in position ⊕, 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 [39] 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 ADNE 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 minutes. So, a dynamic realtime signal strength map is possible as long as structure conditions in a building remain stable.

5. DISCUSSION
To improve the localization performance, the following two problems must be solved. (a) 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; (b) 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 13).

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. Assume a mobile user is moving along a corridor in Figure 9. The distance limit of the mobile user between two continuous locations within a sampling period is no more 5.0 meters. Experiments with mobile users were conducted in this work. Table 5 provides the the localization performance for stationary and mobile users.

The path information in Table 5 is corresponding to the data collection positions along the corridor in Figure 9. For example, Path 1-7 denotes the scenario of a mobile user moving from position 1 to position 7 along the corridor. The numeric value in the table is the average estimation error in meters for all data collection positions along the path. The

| Table 3: Localization performance of six experimental measurements |
|----------------------------------|---------------------|---------------------|---------------------|---------------------|
| Clustering | LMSE  | 2-N  | 3-N  |
| err  | std   | err  | std   | err  | std   |
| Day1  | 2.8372 | 2.4304 | 2.7442 | 2.6349 | 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 | 2.0800 |
| 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 | 2.1866 |
| Avg   | 2.8487 | 2.2725 | 3.2759 | 2.2531 | 3.8780 | 2.7445 | 3.6516 | 2.7173 |

| 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 |

Figure 14: Grid resolution on the performance of clustering localization.

Figure 13: Grid resolution on the performance of clustering localization.
bottom row is the overall localization performance for both cases.

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 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