國立臺灣大學資訊管理研究所碩士論文

指導教授： 林永松 博士

室內空間定位系統中實測點選擇演算法

Measure Point selection Algorithms for Wireless Indoor Positioning Systems

研究生： 陳大鈞 撰

中華民國九十三年七月

謝詞

正如兩年前進來時恩師林永松老師所說的，研究所兩年的時間說長不長，說短也不短。兩年的時光感覺過了好久好久，但迎新座談會時的記憶卻猶如昨日景象般清晰。兩年的過程有苦也有樂，而許多艱困的時期都仰賴許多人的幫忙才得以度過難關的。這裡我首先想感謝的就是我的指導老師，林永松老師。有了您一再不厭其煩的指導，屢次為我指點迷津，我才得以完成我的碩士論文，並且通過嚴酷的論文口試。

再來我想要感謝的是我的口試委員們：孫雅麗老師、李瑞庭老師、以及陳炳宇老師。由於你們寶貴的指導與建議，使得本論文能夠更加的完備。另外要感謝的是跟我一起打拼兩年的研究室伙伴們：瑩珍、坤威、翔騰、育先、耿宏、俊達、明立、閔元、榮耀、柏鈞。這兩年的時光有你們真好，可以彼此聊天打氣，在畢業旅行以及其他時刻一同吃喝玩樂，遊山玩水;在遭遇問題與困難時有你們分憂解勞，而論文撰寫與口試這一關也是靠著與你們之間彼此的砥礪才得以完成。這份情誼我一生也不會忘懷的。同時也感謝研究室學弟妹們：建宏、書平、孝澤、琳智、明源，感謝你們畢業典禮時的祝福，以其在口試期間給予許多的幫忙協助，讓口試的過程能更加順利的進行。

接著我想感謝我的爸媽，沒有你們一直以來的栽培與關懷，便造就不了今天的我;妹妹從旁關心與適時的協助，讓我能夠更加專注於課業。我也想感謝我最要好的朋友宇誠，一直以來你一直扮演著亦師亦友的角色，能夠在最適當的時機給予我關心與打氣，並分享我的困擾，也給予我在課業與論文上莫大的幫助。

最後我要感謝的是我的女友卲芳，能夠在研究所階段遇見妳實在是太好了，有了妳我不再只是一個人，有了妳也使我有更大的動力經營我的人生，並為我們的未來奮鬥。也非常感謝妳在我論文口試之前的包容與諒解，有妳真好。

僅以此論文獻給所有曾經幫助與祝福我的人，並向你們致上最高的敬意，

陳大鈞 謹識

2004.7.27

論文摘要

近年來，室內空間中使用者定位方面的服務逐漸成為一項熱門的議題，同時 IEEE802.11 無線網路技術之成熟，也使其成為室內空間定位上的首選。而由於 802.11 之 RF 訊號易受到室內空間中的障礙物以及人體的影響而衰減，因此傳統的室外空間定位演算法，如三角定位法是不適用的。

為了能在室內空間中精確地定位，許多研究提出為室內空間事先建立一個場地訊號測量資訊的資料庫是有必要的。在實際定位過程中，藉由比對訊號資料庫與行動端點收到各 AP 的訊號強度值，可得出該行動端點最有可能的定位點。然而實際測量全部各點的訊號值卻是非常耗費人力的。因此，本研究的目標即在於，透過精心選擇適量的實測點並收集其訊號實測值，輔以良好的訊號推估演算法，可藉精確地推估出其餘各點的訊號強度值，並減少人力的浪費。

本研究分為兩階段，第一個階段提出透過實測值以推測出推估值之最佳訊號強度推估演算法;第二階段則根據第一階段的訊號推估演算法，以及給定的實測密度，使用求解組合性最佳化的模擬退火法以求解最佳的實測點位置組合。.

關鍵詞：室內定位、室內定位演算法、IEEE802.11、訊號強度、最佳化、模擬退火法、組合性最佳化

THESIS ABSTRACT

GRADUATE INSTITUTE OF INFORMATION MANAGEMENT

NATIONAL TAIWAN UNIVERSITY

NAME: TA-CHUN CHEN **MONTH/YEAR: JULY, 2004**

ADVISER: YEONG-SUNG LIN

MEASURE POINT SELECTION ALGORITHMS FOR WIRELESS INDOOR POSITIONING SYSTEMS

Recently, the service of indoor positioning system has gradually become a hot issue; and with the maturation of IEEE 802.11 wireless technology, it has been the first choice for indoor positioning system. Owing to the sensitivity of RF signal of 802.11 which may attenuated by obstacles and human body, traditional outdoor positioning algorithm, such as triangle positioning algorithm, is not suitable to use for indoor positioning.

In order to accurately position in indoor space, many researches have pointed out that a previously built RSSI (Received Signal Strength Indicator) database is necessary. By comparing the RSS vector received at mobile nodes with RSSI database, we can precisely position the location of mobile users. However, collecting RSS for all grids of indoor space costs lots of human resource. Hence, the purpose of this thesis is to propose a method, which selects measure points elaborately, and collocates with a nice RSS inference algorithm, and then we can build up well RSSI database with relatively lower cost.

In this research we proposed a method that selects suitable quantity of measure points at elaborately selected locations, and infers the signal strength of the other points based on

these selected measure points to reduce signal strength collecting cost.

Keywords: Indoor position, Indoor Positioning Algorithm, IEEE 802.11, Signal Strength, Optimization, Simulated Annealing, Combinatorial Optimization

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background

With the emerging wireless technology, such as IEEE802.11, many issues and applications have been proposed, and the indoor positioning service is one of those issues. The main goal of indoor positioning service is to make users interacting with their surrounding environment. Consider the following example, while the user moving around in the Palace Museum with a handheld device such as PDA, he or she may want to know the history of the China in front of them, or they would like to be directed to the nearest service counter; furthermore, the positioning service can deliver the location-based content to the user's handheld device as soon as they need immediately. Figure1-1 shows typical indoor positioning graph, and we may position mobile users with various positioning method.

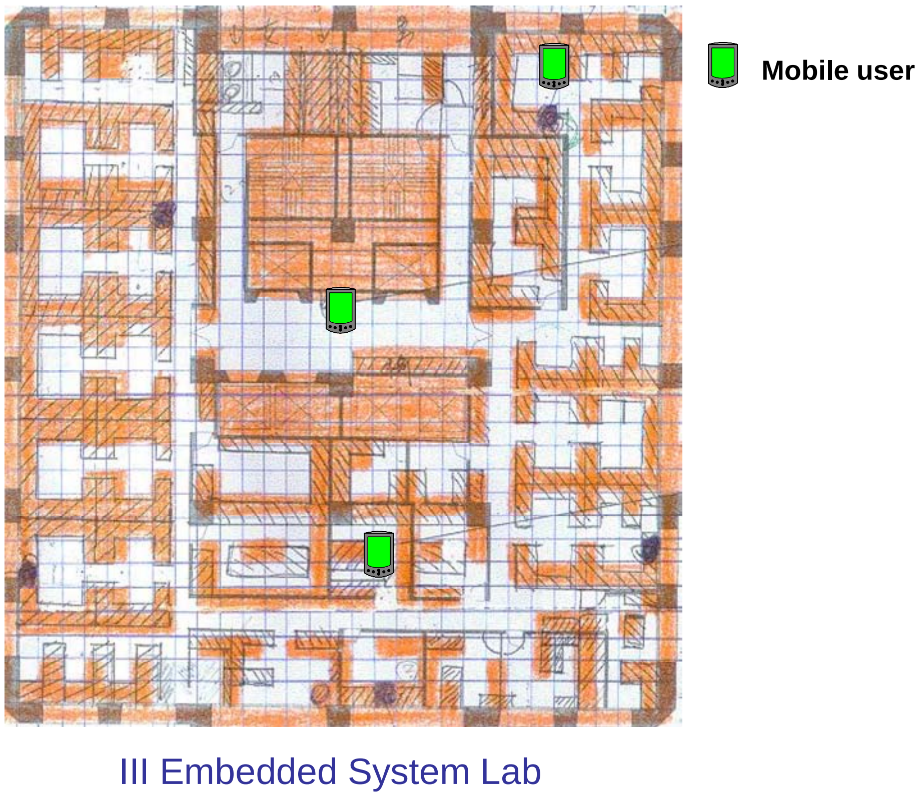

Figure 1–1 Typical Indoor Positioning Field

The granulation of location information needed varies from one application to another. One may acquire the exact location information, such as the distance how far they are from some proximity, while the others may just want to know which region they probably are. In general, the accuracy of positioning in the Indoor space depends on the purpose of the application, and the most important, the computation cost and complexity of the indoor positioning systems varies with the positioning accuracy. General wireless positioning system diagram proposed in [7] is as Figure1-2:

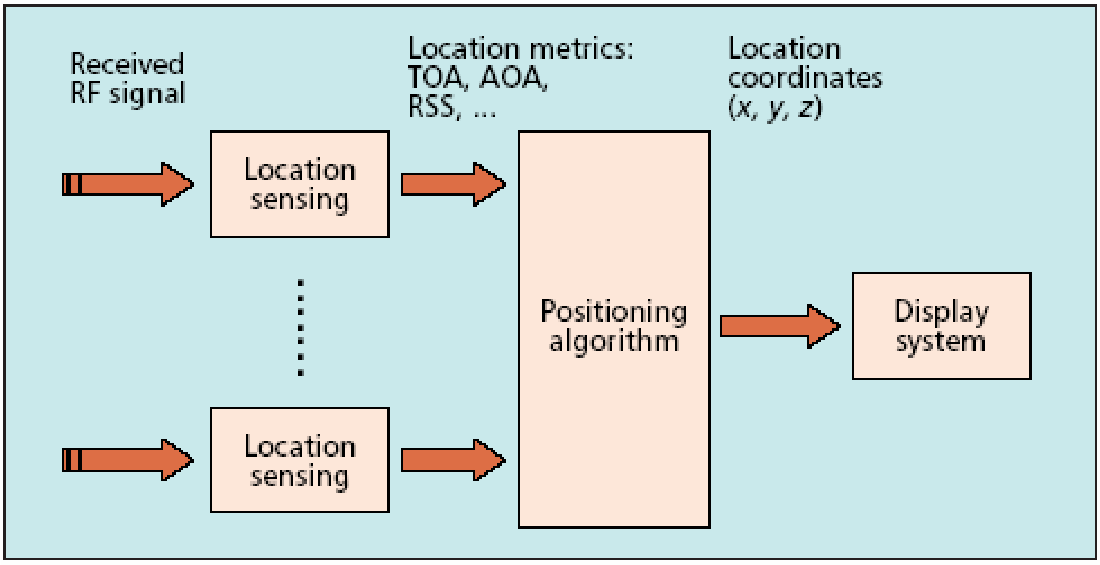

Figure 1–2 General Wireless Positioning System Diagram

The main elements of the system are a number of location sensing devices that measure metrics related to the relative position of a mobile device with respect to a known reference point, a positioning algorithm that process metrics reported by location sensing elements to estimate the location coordinates of mobile device, and a display system that illustrates the location of the mobile device to users.

To our knowledge, there are few indoor positioning systems use pure 802.11 architecture without any other additional entities. Most application with high positioning accuracy relies on additional hardware or use more than two types of signal simultaneously [5]. However, these applications suffer from the limitation of scalability and introduced more cost, because additional hardware or protocol modifications are required, and thus lacks portability.

[1][2][3] developed indoor positioning systems relied on pure IEEE802.11b-based system architecture without any protocol modification or additional hardware. These

applications utilize off-the-shelf technology, which is relatively inexpensive and also reduce the limitation of hardware and protocol modification constraints.

Positioning methods of these applications are achieved by the previous construction of the radio map of the information of the indoor space. Owing to the sensitivity of the RF signal, which may be interfered by the obstacles of indoor spaces, or it may be attenuated by human bodies, it is difficult to infer the location of users by time differencing method such as AOA (Angle of Arrival), TOA (Time of Arrival), or TDOA (Time Differentiate of Arrival). Such method requires time synchronization of mobile devices with base stations, and is not suitable for indoor environment [6]. Thus, these applications require construction of radio map of indoor space previously. Radio map of an indoor building includes the RSS (Received Signal Strength) vectors at every free space in the map. General radio map is a two dimensional discrete grids. In [1], they define each grid as either following four types: AP, Obstacle, Measure point, and Infer point. Obstacle grid stores the type of this obstacle and its attenuation efficient; Measure point and Infer point stores the RSS received from each AP. The system architecture of [1] is in Figure 1-3:

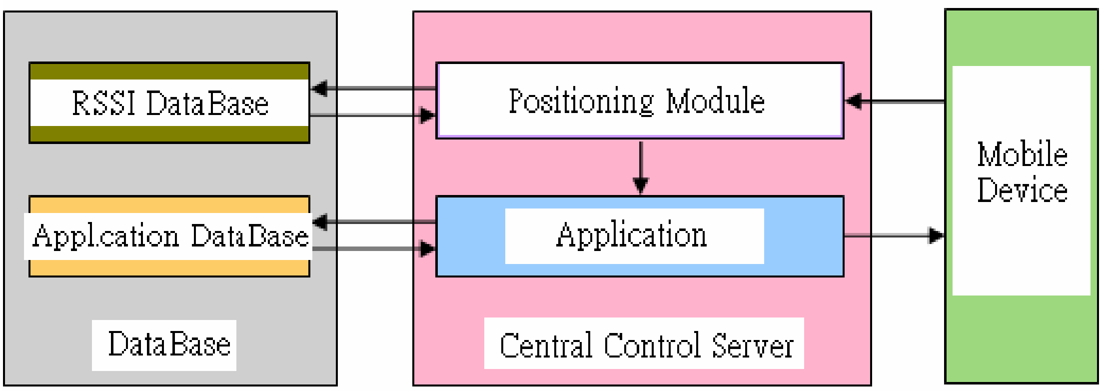

Figure 1–3 System Architecture of III Wireless Indoor Positioning System

To realize this positioning system, it is divided into two processes: Offline-process and On-line process. In offline process, we collect RSS vectors of each point from each AP. The on-line mode is based upon the RSS information collected in the off-line mode; in on-line process, the mobile device periodically transmit the RSS vector received from each AP to the Central Control Server, and the Positioning Module in the Central Control Server would compare the RSS vector with the RSSI (Received Signal Strength Indicator) Database in the Database to find out where the mobile device is. This kind of systems achieves high positioning performance, but relies highly on the accuracy of RSSI Database build in off-line process, which is also called the radio map mentioned previously.

1.2 Motivation

In the previous section, we can see that positioning accuracy relies highly on the accuracy of the RSSI Database. To build a good RSSI Database, we have to collect RSS for all free space points, which introduces lots of human-resource. In fact, if we can collect RSS for only some specifically selected measure points, and if we also have an excellent algorithm to infer RSS for those infer points based on known measure point, then we can reduce lots of RSS collecting cost and acquire a considerable well RSSI Database. Our research has two phases. In phase1, we proposed several algorithms to infer RSS of each infer points based on those collected RSS measure points, then we compare which one is the best based on two RSSI Databases. In phase two, we designed an algorithm to select the optimal measure points based on the best algorithm proposed in phase one. The optimization algorithm in phase2 is based on simulated annealing method, which can be implemented with large-scale combinatorial optimization problem.

Chapter 2 Literature Survey

2.1 Location and Tracking System

Related work of location and tracking systems can be separated into four categories: (1) IR-based systems (2)indoor RF-based systems (3)wide-area cellular-based systems, and (4)The others, such as ultrasound-based systems. The relative work of these location and tracking systems are introduced and discussed below.

The most representative work in IR-based location systems is the Active badge system [4]. User carries a handheld device which emits IR signal periodically. IR receivers are placed at known positions in indoor space. Upon receiving IR from handheld device, the IR receiver would relay this signal to location manager software. This system provides high accuracy, yet suffers from some drawbacks such as the poor signal range due to the IR, and it incurs significant hardware setup cost.

The RF-based positioning systems are RADAR [1] and Duress Alarm Location System (DALS) [8]. These two systems uses RF signal strengths (RSS) to determine user location, but DALS differs from our system and RADAR that it requires special hardware and infrastructure deployment over and above a wireless network, and it does not consider the factor that the human orientation may have extremely effect on

RF signals and the factor of RF propagation model. RADAR system is developed by Microsoft Research Center which is based on RF signal without additional hardware. Access Points (AP) are located in some way that provide overlapping coverage in the area where is going to be used for positioning. User carries handheld device such as PDA or mobile phone equipped with a wireless LAN card which responsible for communicating with AP. RADAR uses the RSS received at each handheld device from each AP to infer the user's location. It build a Radio Map which is a database of locations in the indoor space that stores the signal strength of the beacons emanating from AP observed or estimated at those locations. RADAR proposed two types of method to create the Radio MAP, the first one is empirical method, which requires a mobile user walking around the building with a handheld device to record the RSS from each AP at each locations. The second method is mathematical method, which involved computing the RSS at each location with a mathematical model of indoor RF signal propagation. This mathematical model considers both free space path loss and attenuation due to obstructions like walls or obstacles between AP and the mobile user.

In the wide-area cellular area, several location determination systems have been proposed [9]. The popular way of measuring mobile phones base upon the signal attenuation such as angle of arrival (AOA), time of arrival (TOA) and time differentiate of arrival (TDOA). These methods have well performance in the outdoor environment, while they are not suitable for indoor environment because the RF signal suffered from multiple reflections, for example, multi-path, shadow fading, and non-light-of-sight problem. It is also no off-the-shelf technology to provide the time synchronization supplies the TOA or TDOA method and requires additional hardware (may be relatively expensive). Furthermore, Global Positioning System (GPS) is a powerful outdoor positioning system, but it is not effective for indoor positioning system because the GPS transmission would be blocked by buildings.

2.2 Simulated Annealing

Optimal measure point selection problem is an NP-hard problem. For each point, we have to decide whether it is a measure point or an infer point, in other words, we have to partite these points into two set, one is measure points, and the other is infer points. Partitioning problem is a kind of combinatorial problem. A general approximation algorithm that runs in polynomial-time needs to be used in order to solve this kind of problems. It is difficult to find such algorithm to obtain near optimal solution. Simulated annealing (SA) [12] is considered an approximation algorithm where it is applicable to various problems in general. The SA algorithm can be considered as a version of an "iterative improvement algorithm" which considers only specific transitions and terminates in the first local minima found. Unlike those algorithms, simulated annealing allows various types of transitions in which some of them may be opposite towards achieving the goal. For instance, cost increasing transitions are also accepted along with cost decreasing transitions whereas iterative improvement algorithm would allow only cost-decreasing ones to pass. However, it is proven that eventually simulated annealing produces more optimal solution than the original iterative improvement algorithm.

Metropolis algorithm was the original idea behind the optimization technique of SA. Kirkpatrick et. al, [11] has used Metropolis algorithm as a global optimizer. Thus, simulated annealing is also known as global optimizer. This algorithm is then applied to the physical design of computers. The advantage of using simulated annealing is its ability to scale for large scale optimization problems and its robustness towards achieving local optima convergence.

SA starts with an initial solution, *s*. A neighbor to this solution *s'*, is then generated as the next solution and the change in cost. The aim of a generation mechanism is to generate a new solution *s'* out of the solution *s* by means of a random perturbation in one of the variables of *s*. We adopted the generation mechanism of Spatial Simulated Annealing (SSA) [13], which is a modification of SA. In SSA, *s'* is done by moving one randomly chosen measure point $\overline{x_i}$ over a vector $\vec{h}$ with the direction of $\vec{h}$ drawn randomly, and $|\vec{h}|$ taking a random value between 0 and $\vec{h}_{max}$. One of the modifications of SSA as compared to ordinary SA is that $\vec{h}_{max}$ initially is equal to half the length of the sampling region, and decreases with time. This increase the efficiency of the demanding recalculations after each modification in the sampling scheme, because it can be expected that with optimization of sampling schemes, successful modifications consist of increasingly smaller values of $|\vec{h}|$ as the SSA process advances. This is because the process deals with many similar variables. Therefore, moving measure points randomly over large distance will not contribute much to finding the minimum towards the end of the optimization process. At the end the final value of the control parameter $\vec{h}_{max}$ will be almost equal to zero.

After *s'* has been generated, $\Delta F(s, s')$ is evaluated. If a reduction in cost is found, the current solution is replaced by the generated neighbor, otherwise we decide with a certain probability whether *s* remains or *s'* becomes the current solution. The probability of accepting a transition that causes an increase, ΔF, in the cost is usually called the acceptance function and is set to $\exp(-\Delta F/T)$ where *T* is the control parameter that corresponds to temperature in the analogy with the physical annealing process. In SA, the algorithm is started with a relatively high value of *T*, to have a better chance to avoid being prematurely trapped in a local minimum. The control parameter is lowered in steps until it approaches to zero. After termination, the final configuration is taken as the solution of the problem at hand. That is, simulated annealing is a generalization of the local search algorithm. As the Figure 2-1 and Figure 2-2 shows below, SA has the chance to climb out of local optima and eventually find the global optima. The general pseudo code of SA is also showed below.

```
SimulatedAnnealing()
  S = Initial solution to the optimization problem;
  T = Initial Temperature
  While T > μ
    While (InnerLoop Condition)
       M = A Randomly Chosen Modification to S;
       If func(Gain(M), T) Apply M to S;
    End While
    T = Update(T);
  End While
```

Figure 2–1 Pseudo code of general simulated annealing procedure

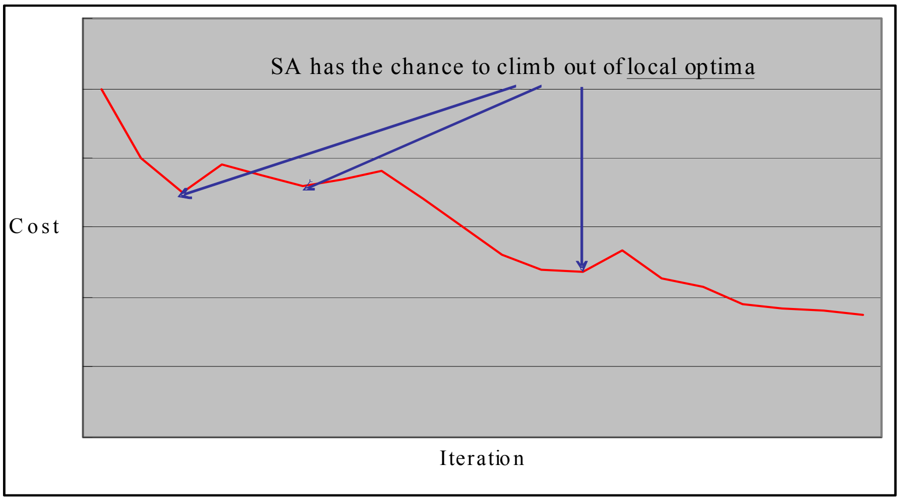

Figure 2–2 the relationship of Cost and Iteration in Simulated Annealing

Chapter 3 Solution Approach

We divide our problem into two phases. Phase one is to compare several RSS inference algorithms which infer RSS of an infer point by their surrounding measure points, and find the best one using a previously built RSSI Database of floor 5 of NTU Management Building. In phase two, we proposed an optimization algorithm using simulated annealing (SA) to select optimal measure point based on the inferring algorithm in phase1.

3.1 Phase1: RSS Inference Algorithms

We proposed several RSS inference algorithms below.

3.1.1 Notation

Notation	Description
I	Infer point.
S_i	The ith measure point of I.
AP	Access Point.

D_i	The distance from *I* to S_i.
RSS(p)	The RSS of a measure or infer point *p*.
r	Radius of current searching area circle.
V_1	(X_1, Y_1), vector determined from *AP* to S_1.
V_2	(X_2, Y_2), vector determined from *AP* to S_2.
V_3	(X_3, Y_3), vector determined from *AP* to S_3.
V_4	(X_4, Y_4), vector determined from *AP* to S_4.
V	(X, Y), vector determined from *AP* to I.
α	The linear combination coefficient of V_1.
β	The linear combination coefficient of V_2.
γ	The linear combination coefficient of V_3.

Table 3-1 Notations of phase1

3.1.2 Algorithm1

This algorithm uses two measure point S_1 and S_2 to infer RSS vector of an infer point *I*. See Figure 3-1.

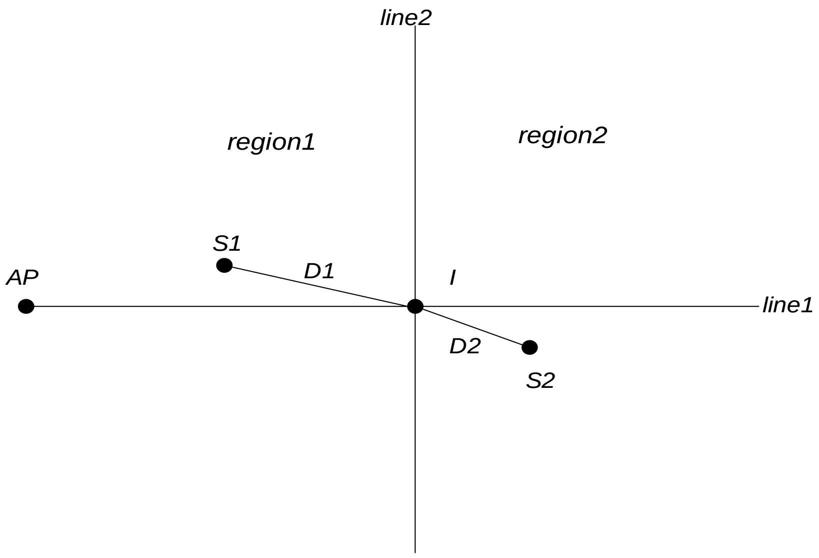

Figure 3–1 Diagram of Algorithm1

Line1 is the line determined by *AP* to *I*, *line2* is the line determined by a line which is perpendicular to *line1*. *Line2* divides searching area into two regions: *region1* and *region2*. In both regions, we search for a measure point *S* which is the nearest point to *I*. Now S_1 and S_2 become the two nearest measure points in each region and D_1, D_2 is the distance from each measure point to *I*. Then we evaluate *RSS(I)* by RSS of S_1 and S_2 with the inverse of D_1 and D_2 to be their weight. The equation of evaluating *RSS(I)* is as follows:

$$RSS(I) = RSS(S_1) \cdot (\frac{D_2}{D_1 + D_2}) + RSS(S_2) \cdot (\frac{D_1}{D_1 + D_2})$$

The main idea of this algorithm is to find two nearest measure point from *I*, and one may has larger RSS than *I*, and the other one may has smaller RSS than *I*. If we calculate the average of these two points with the weight of their inverse of distance from *I*, then we may get a rough RSS of *I*. The mean error of algorithm1 is 3.9dbm.

3.1.3 Algorirhm2

Algorithm2 uses three measure point S_1, S_2, and S_3 to infer RSS of *I*. We first find three measure points S_1, S_2, and S_3 which forms a convex hull that surrounds *I*. Figure 3-2 shows the concept of Algorithm2.

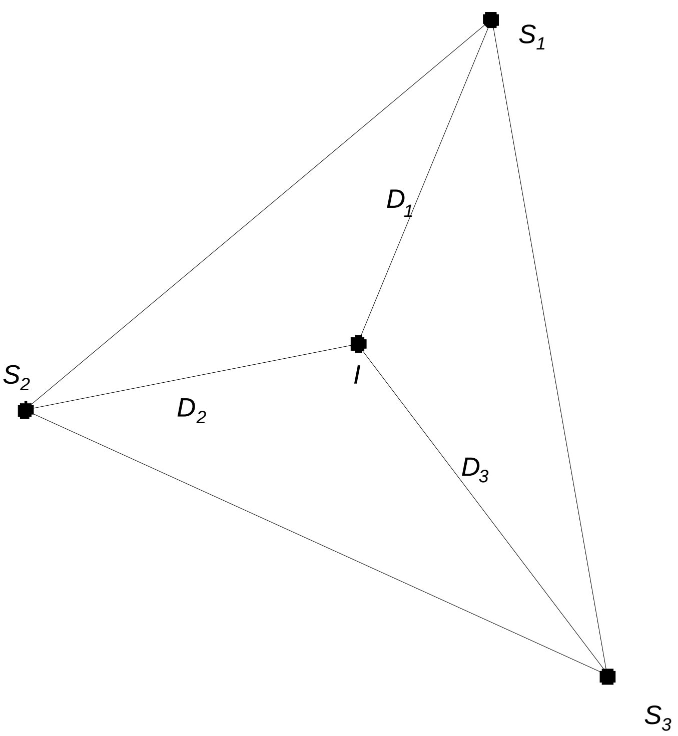

Figure 3–2 Diagram of Algorithm2

Then we can calculate *RSS(I)* by these three points, the following equation shows how to evaluate *RSS(I)*:

$$RSS(I)=\frac{\frac{1}{d1}}{\frac{1}{d1}+\frac{1}{d2}+\frac{1}{d3}}\cdot RSS(S_1)+\frac{\frac{1}{d2}}{\frac{1}{d1}+\frac{1}{d2}+\frac{1}{d3}}\cdot RSS(S_2)+\frac{\frac{1}{d3}}{\frac{1}{d1}+\frac{1}{d2}+\frac{1}{d3}}\cdot RSS(S_3)$$

As we can see, Algorithm3 adopt the fashion used in Algorithm1 that giving the weight of each S_i the inverse of $D_{i.}$ The main idea is to find a triangle which has the minimum area, and also surrounds *I*. The vertex concludes this triangle may have similar RSS with *I*, and may never larger or smaller than *RSS(I)* simultaneously. The mean error of algorithm2 is 2.47dbm.

3.1.4 Algorithm3

Algorithm3 have I to be the origin of a 2-dimensional coordinates. The x-axis and y-axis divides this 2-d coordinates as 4 regions. For each region, we search for a measure point nearest to *I*. Now we can use S_1, S_2, S_3 , *and S4* to evaluate *RSS(I)*. Figure 3-3 shows the concept of Algorithm3.

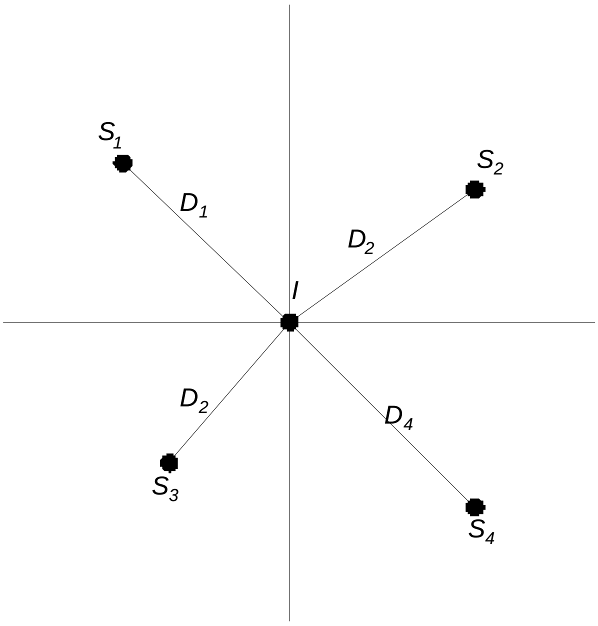

Figure 3–3 Diagram of Algorithm3

The equation calculates RSS of *I* are showed below:

$$RSS(I)=\frac{\frac{1}{D_1}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_1)+\frac{\frac{1}{D_2}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_2)+$$
$$\frac{\frac{1}{D_3}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_3)+\frac{\frac{1}{D_4}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_4)$$

The main idea is simple. We simply find four measure point around *I* and evaluate *RSS(I)* by the weight of inverse of D_i. The mean error of Algorthm3 is 3.16dbm.

3.1.5 Algorithm4

Algorithm4 is a little similar to Algorithm3, but has the lowest mean error so far. It works as follows: *Line1* is determined by *AP* and *I*, *Line2* is a line perpendicular to *Line1* and passes through *I*. *Line1* and *Line2* divides searching area as four region: *region1, region2, region3, and region4*. Furthermore, a circle with radius *r* limits the searching area into this circle. We start searching from $r=1$, and find a measure point which is nearest to *I* for each region. If there are more than two measure points which have the same minimum distance to *I*, then we choose the one S_i with the minimum angle determined by vector *AP-to-Si* and vector *AP-to-I*. If we can find four measure points in the circle with current *r*, then we doubled *r* in the next iteration, until all four measure points have been found. *Line3* is a line perpendicular with *Line1* and passes through *AP*. Region divides by *Line3* which is opposite to the region where *I* locates is the restrict region. Even if circle determined by *r* has covered this region, we can not find measure point located at this region. Figure 3-4 shows the concept of Algorithm4.

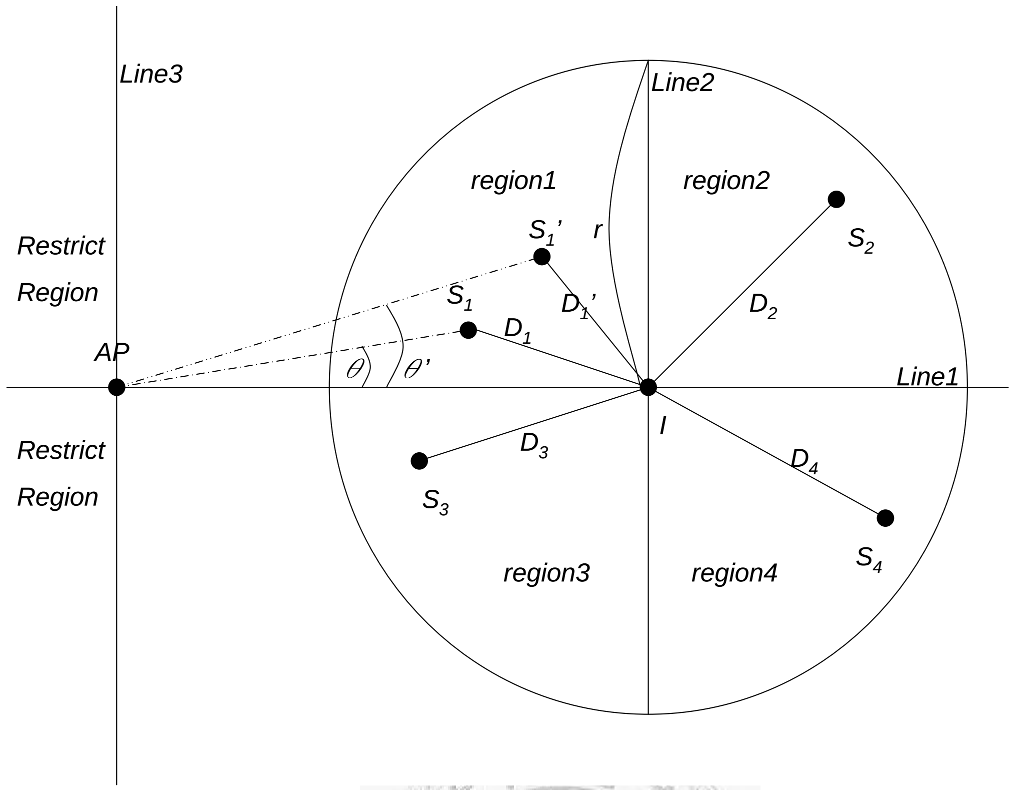

Figure 3–4 Diagram of Algorithm4

In Figure 3-4, we can see that there are two measure points S_1 and S_1*'* with distance D_1 and D_1*'* to *I* respectively. If D_1 equals D_1*'*, then we have to choose the one with the smaller angle. θ is smaller than θ', so we choose S_1 as the measure point in *region1*. The following equation evaluates *RSS(I)*, which is the same with Algorithm3:

$$RSS(I)=\frac{\frac{1}{D_1}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_1)+\frac{\frac{1}{D_2}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_2)+$$
$$\frac{\frac{1}{D_3}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_3)+\frac{\frac{1}{D_4}}{\frac{1}{D_1}+\frac{1}{D_2}+\frac{1}{D_3}+\frac{1}{D_4}}\cdot RSS(S_4)$$

The mean error of Algorithm4 is 2.20dbm.

3.1.6 Algorithm5

Algorithm5 is modified from Algorithm4. So far, we evaluate RSS based on the distance which is from *I* to each measure points. If we give the weight of each RSS of measure point based on the coefficient of linear combination, we may get more precise inferring value of RSS. Figure 3-5 shows the concept of Algorithm5:

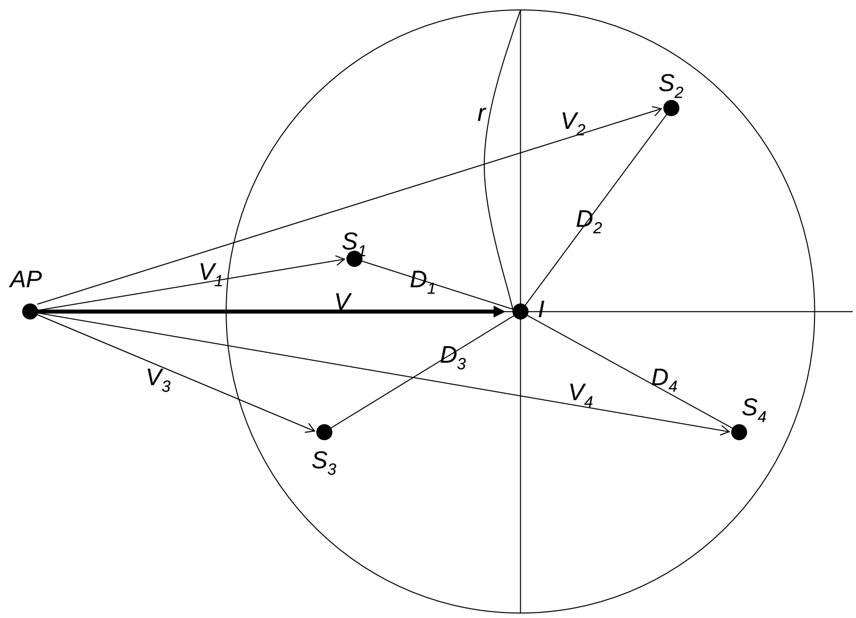

Figure 3–5 Diagram of Algorithm5

In Figure 3-5, vector *V* can be represented by various linear combinations of the four vectors V_1, V_2, V_3, and V_4, in other words, the multipliers of the four vectors have any kind of combinations. Our objective is to find a combination that satisfies the following objective functions and the constraints:

$$\min \ (\alpha D_1 + \beta D_2 + \gamma D_3 + (1-\alpha-\beta-\gamma) D_4)$$

subject to:

$$\alpha X_1 + \beta X_2 + \gamma X_3 + (1-\alpha-\beta-\gamma) X_4 = X \qquad (1)$$

$$\alpha Y_1 + \beta Y_2 + \gamma Y_3 + (1-\alpha-\beta-\gamma) Y_4 = Y \qquad (2)$$

$$0 < \alpha < 1 \qquad (3)$$

$$0 < \beta < 1 \qquad (4)$$

$$0 < \gamma < 1 \qquad (5)$$

$$0 < (1-\alpha-\beta-\gamma) < 1 \qquad (6)$$

The equation that evaluates *RSS(I)* is:

$$RSS(I) = \alpha \cdot RSS(S_1) + \beta \cdot RSS(S_2) + \gamma \cdot RSS(S_3) + (1-\alpha-\beta-\gamma) \cdot RSS(S_4)$$

As we can see from the objective function, coefficient of each vector relates to the distance from each measure point to *I*. Constraint (3), (4), (5), (6) force each measure point to contribute at least a little, and would not dominate to *RSS(I)*. The RSS evaluating equation shows that measure point with larger distance has smaller contribution to the inferring *RSS(I)*.

3.2 Phase2: Optimal Measure point Selection Algorithm

In phase1, we proposed several inference algorithms and also found the best one. In phase2, we proposed an optimal algorithm to select measure points. This optimal algorithm would select measure points based on algorithm5 proposed in phase1, and solved by simulated annealing (SA). We also proposed a contrast algorithm to compare with the performance with Algorithm5. We defined an objective function for algorithm5 and the contrast algorithm, and would be used by SA to evaluate the cost of each combination of measure points.

3.2.1 Notation

Notation	Description
M	Positioning field map, a 2-dimensional m*n coordinates.
D	Randomly generated number, $0 \le D \le 1$.
$\vec{h}$	The shift vector for a candidate measure point t.
$\lvert\vec{h}\rvert$	Length of $\vec{h}$, $1 \le \lvert\vec{h}\rvert \le h_{\max}$.
$h_{\max}$	The maximum distance in *M.*
$L_{\max}$	The maximum try iteration for current temperature.
S	A combination of selected measure points.
S^i	Combination of selected measure points in stage *i.*
S^{i+1}	Combination of selected measure points in stage $i+1$.

t	A measure point randomly picked in *S*.
T	Current temperature.
T_f	Final (lowest) temperature.
α_T	Gradient of lower temperature.
α_h	Gradient of shorten $\lvert\vec{h}\rvert$.
E^i	Energy of state *i*.
ΔE	$E^i - E^{i+1}$

Table 3-2 Notations in phase2

3.2.2 Objective Function of Algorithm5

The most important part of SA is the objective function. SA would use the objective function to evaluate the cost (or the energy state) of current solutions. Thus, the objective function should be as close the requirement of the inference algorithm as possible. In phase1, the inference algorithm select four measure points which are the closest points to the infer point, and also have the minimum angle formed by the AP-to-measure-point and the AP-to-infer-point two vectors. In Algorithm5, we can see that both distance and angle should be considered jointly, so we evaluate both angle and distance for each measure point simultaneously. Here is the pseudo code of our objective function:

```
Compute_Objective_Function(S)
begin
  E = 0;
  for(each inferring point I){
    for(each AP A){
      find four sampling point S1,S2,S3,S4 of I based on Algorithm5;
      E = E + angle(AS1, AI) + angle(AS2, AI) +
          angle(AS3, AI) + angle(AS4, AI) +
          D1 + D2 + D3 + D4;
    }
  }
  return E;
end
```

E represents the energy state of current state, which represents the cost of current solution. In other words, lower energy has lower cost. This algorithm calculates angles and distance of each measure points and summarized in E because lager angle and distance means higher cost.

3.2.3 Objective Function of Contrast Algorithm

The contrast algorithm evaluates and summarizes the distance from infer point to its nearest measure point, and would be compared with algorithm5. Here we show the pseudo code of objective function of this contrast algorithm:

```
Compute_Objective_Function_Contrast_Algorithm()
begin
E = 0;
  for (each inferring point I){
    find the nearest sampling point S of I;
    E=E+distance(S,I);
  }
  return E;
end
```

The objective of this algorithm is to minimize the distance from each infer point to its nearest measure point, thus makes the measure points to be evenly distributed.

3.2.4 Annealing Process

Here is the pseudo code of SA for optimal measure point selection:

SA_Sampling_Point_Selection()

begin

  Randomly generate an initial combination of sampling point S^i;

  $E^i = Compute\_Objective\_Function(S^i)$;

  $L = 0$;

  while $(T > T_f)\{$

    while$(L<L_{\max})\{$

      Randomly select a sampling point t from S^i;

      do$\{$

        Randomly generate $\vec{h}$ where $1 \le |\vec{h}| \le h_{\max}$;

        $t' = t$ moves along with $\vec{h}$;

      $\}$

      while (t' is out of boundary of M or t' is not in free space);

      $S^j = S^i$ where t in S^i replaced with t';

      $E^j = Compute\_Objective\_Function(S^j)$;

      if$(E^i - E^j \ge 0)$

        Accept S^j, $S^i = S^j$;

      else$\{$

        Randomly generate D where $0 \le D \le 1$;

        if$(\exp((E^i - E^j)/T) < D)$

          Accept S^j, $S^i = S^j$;

      $\}$

      $L = L + 1$;

    $\}$

    $T = T * \alpha_T$, $h_{\max} = h_{\max} * \alpha_h$;

    $L = 0$;

  $\}$

  Output S^i;

end

The whole SA process begins with a randomly generated initial combination of measure points, and also evaluates the energy state (the cost) of the initial solution. The following is the temperature lowering process. For each round of temperature lowering process, there are L_{max} times of try; each try we randomly generate a neighbor solution S^j from current solution S^i, and evaluate its cost. The randomly selected point t in S^i will be shifted along with randomly generated vector $\vec{h}$. Initial length of $\vec{h}$ is the maximum distance in M, and would be iteratively shortened while temperature is lowered, and become one unit of moving distance finally. If the cost of this neighbor solution is lower than current solution, then the current solution would be replaced by the neighbor solution. Otherwise, we evaluate if $\exp((E^i - E^j)/T) < D$, and then the neighbor solution should be accepted. $\exp((E^i - E^j)/T)$ is the acceptance function with value between zero and one. With higher temperature, the acceptance function would have lower value, that is, the neighbor solution has higher probability to be accepted even its cost is larger than current solution. This mechanism avoids SA to be trapped in local optima and has the ability to achieve global optima. If the try has reached L_{max}, temperature would be lowered and $|\vec{h}|$ would be shortened. The parameter of α_T and α_h is 0.99, T is 2.25, and T_f is 0.015, L_{max}=200.

Below is the flow chart of the SA for optimal measure point selection process:

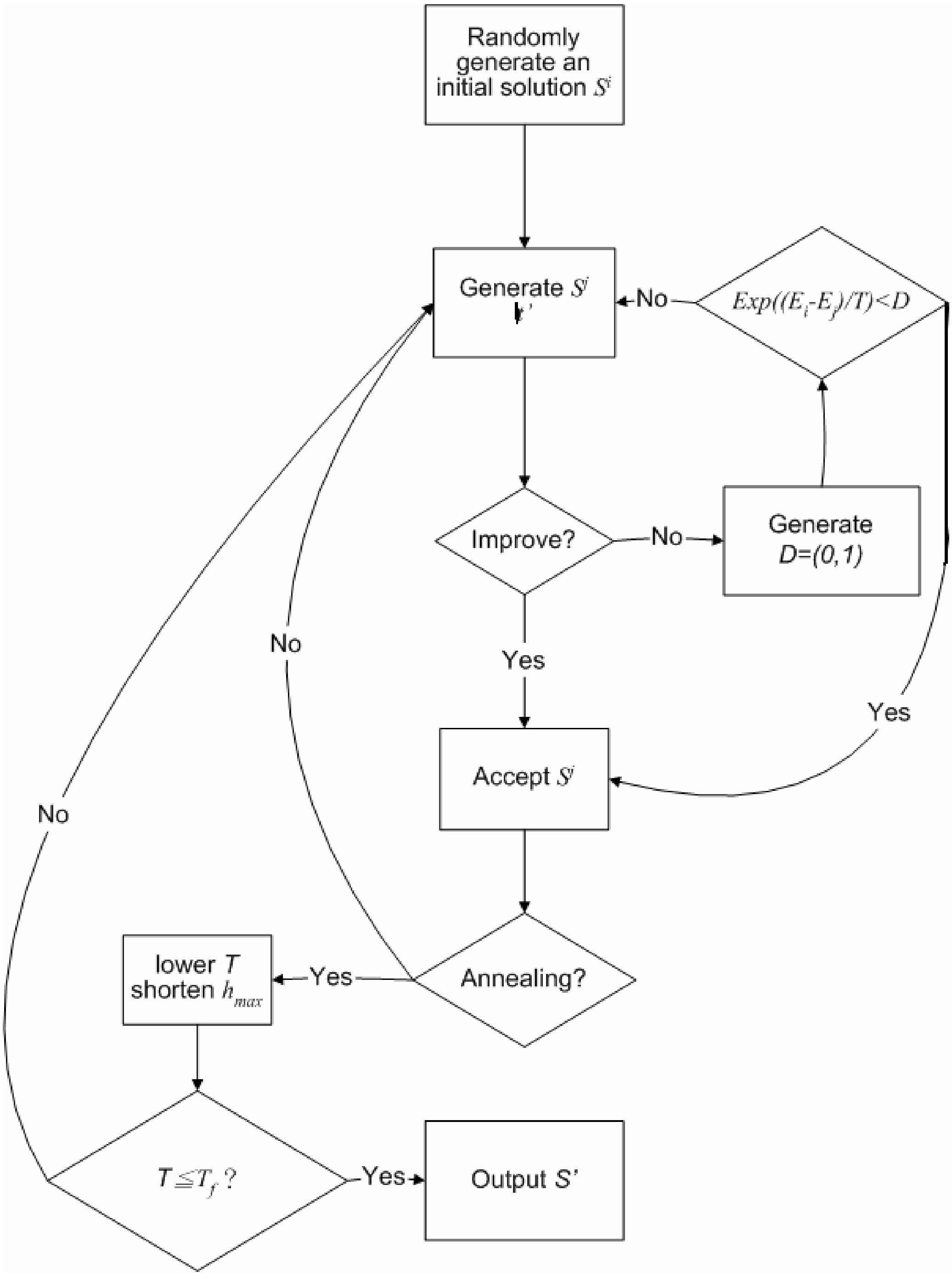

Figure 3–6 flow chart of Measure point Selection with SA

Chapter 4 Experimental Result

To evaluate the performance of our algorithms, we build RSSI Databases for two test beds, one at floor five of National Taiwan University Management Building, and one at the Institute for Information Industry (III) Embedded System Laboratory, to evaluate the solutions computed by SA process. We have experiments for various densities of measure points computed by two algorithms, one for SA, and one for the contrast algorithms. We show the experiment result below, including the location of measure points and the mean error of various densities in two test beds.

Symbol	Meaning
■	Obstacle
□	Infer point
*	Measure point
AP	Access Point

Table 4-1 Symbols for Test bed diagram

4.1 Test bed1: F5 of NTU Management Building

entity	Quantity
AP	5
Obstacles	691
Boundary nodes	90
Free space	396

Table 4-2 Parameters of test bed1

Fig 4-1 shows the radio map of test bed1.

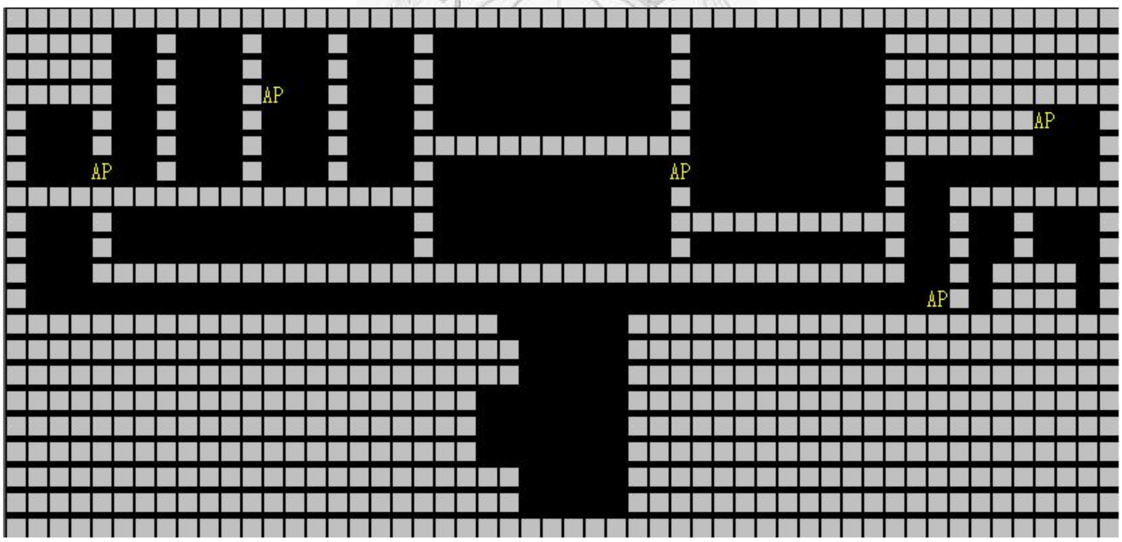

Figure 4–1 Radio map of Floor 5 of NTU Management Building

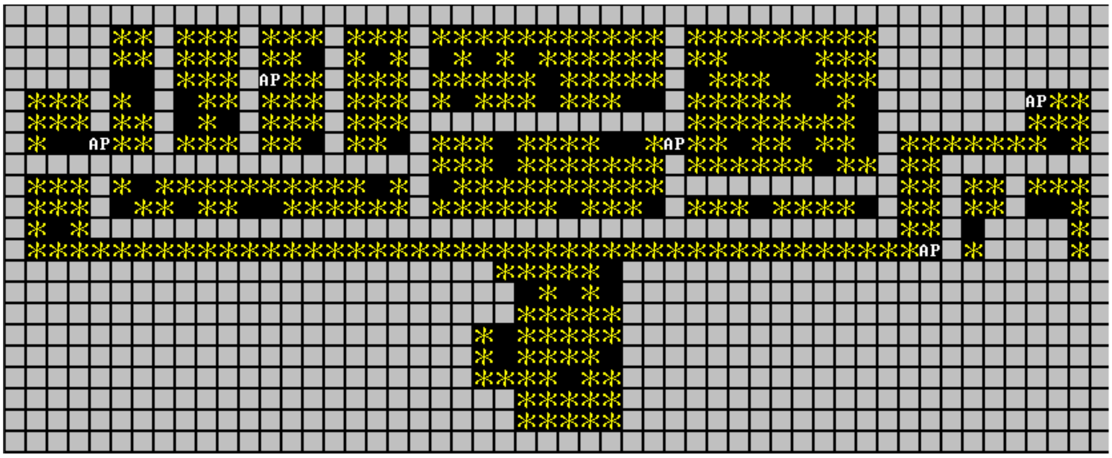

Figure 4–2 measure points with Alg5[1] with density: 0.8 mean error: 1.99

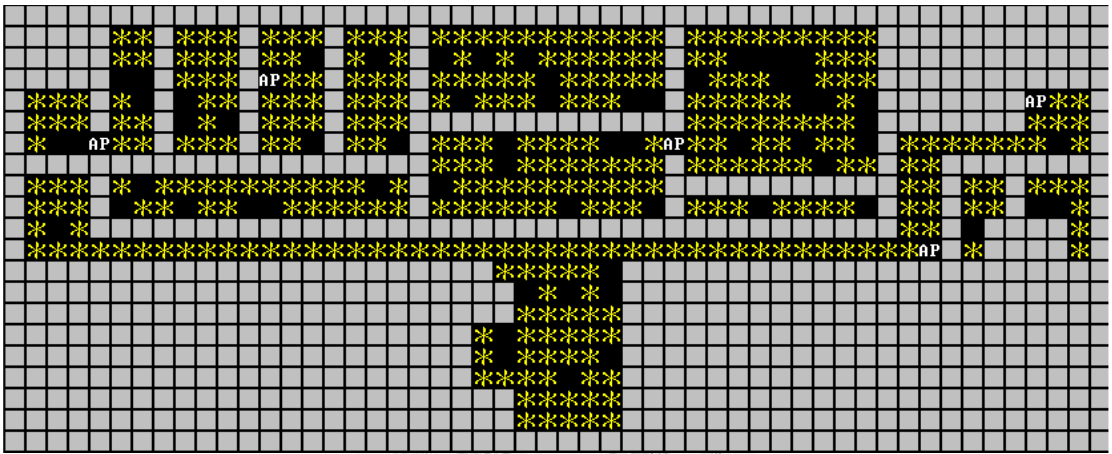

Figure 4–3 measure points with CA[2] with density: 0.8 mean error: 2.15

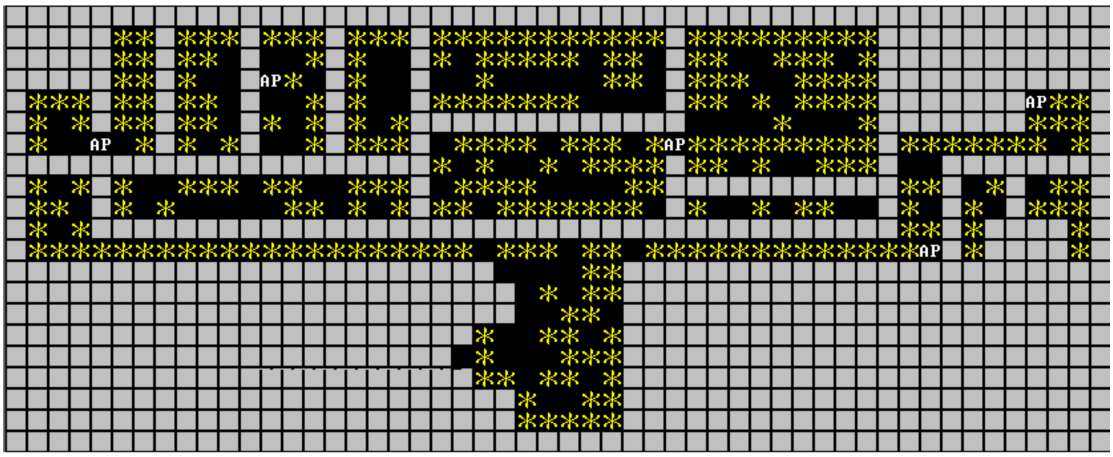

Figure 4–4 measure points with Alg5 with density: 0.6 mean error: 2.10

[1] Alg5 means objective function of Algorithm5.
[2] CA means objective function of Contrast Algorithm.

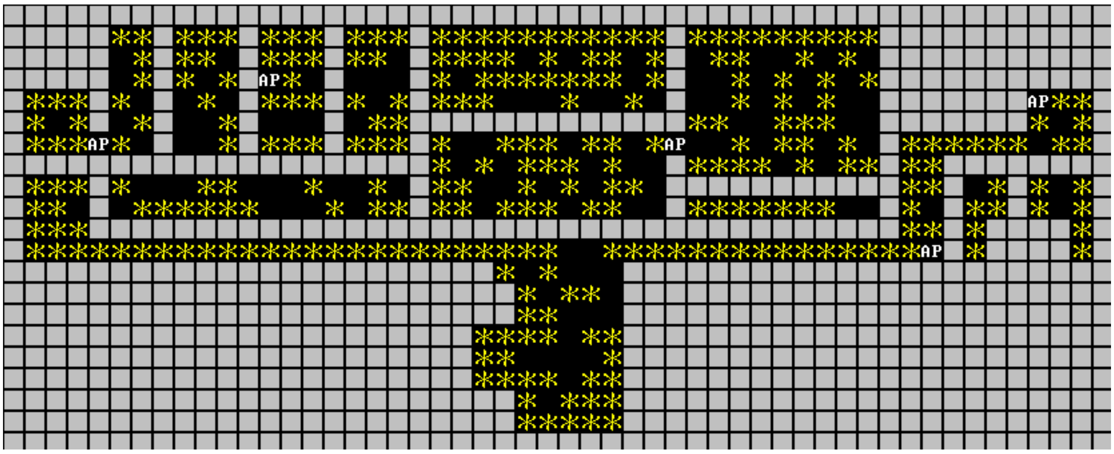

Figure 4–5 measure points with CA with density: 0.6 mean error: 2.19

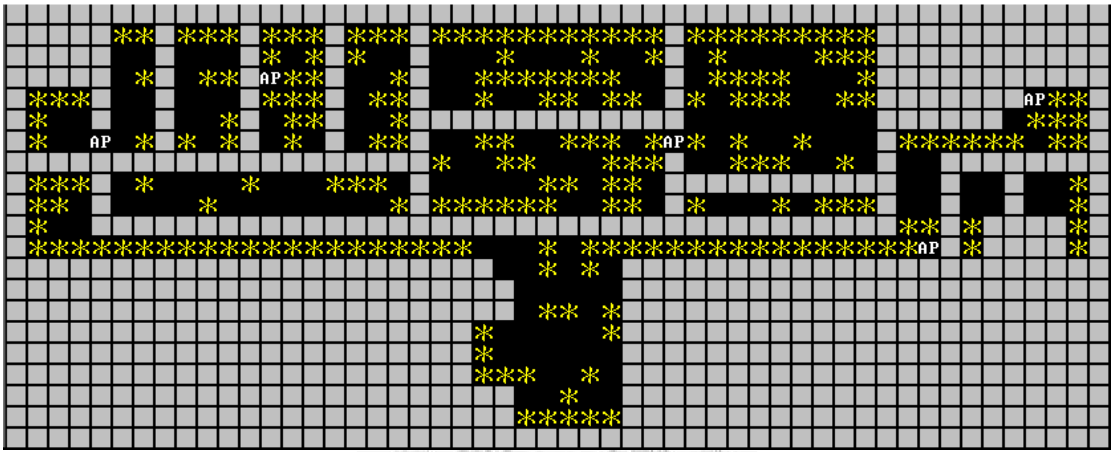

Figure 4–6 measure points with Alg5 with density: 0.4 mean error: 2.25

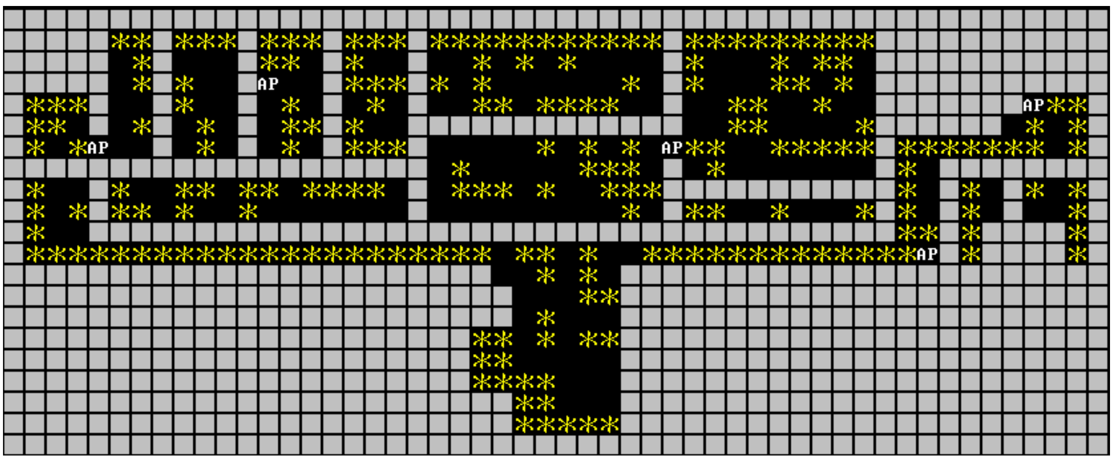

Figure 4–7 measure points with CA with density: 0.4 mean error: 2.26

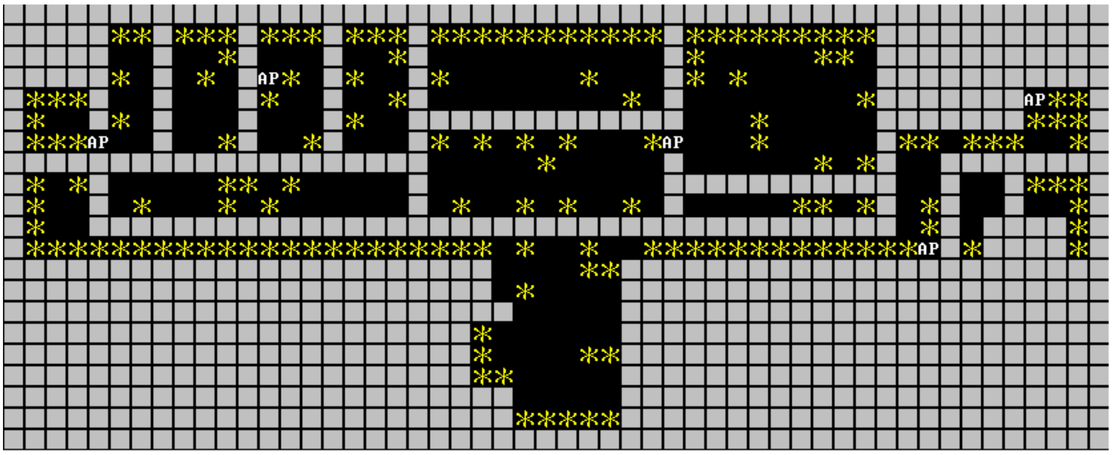

Figure 4–8 measure points with Alg5 with density: 0.2 mean error: 2.34

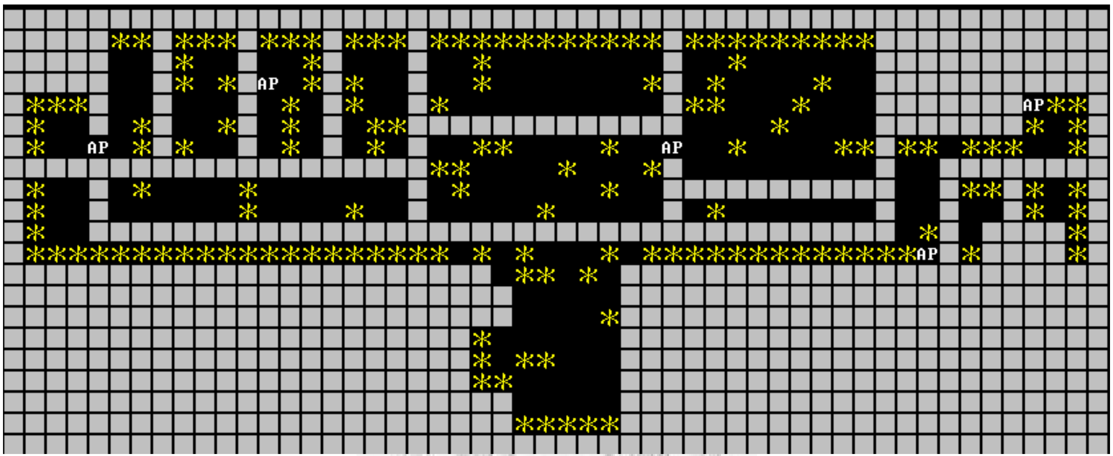

Figure 4–9 measure points with CA with density: 0.2 mean error: 2.39

	Mean error (dbm)	
density	Algorithm5	uniform
0.2	2.34	2.39
0.4	2.25	2.26
0.6	2.10	2.19
0.8	1.99	2.15
1	1.95	1.95

Table 4-3 mean error of two objective functions in test bed1

	Maximum error (dbm)	
density	Algorithm5	uniform
0.2	15	22
0.4	22	20
0.6	17	16
0.8	17	14
1	21	21

Table 4-4 maximum error of two objective functions in test bed1

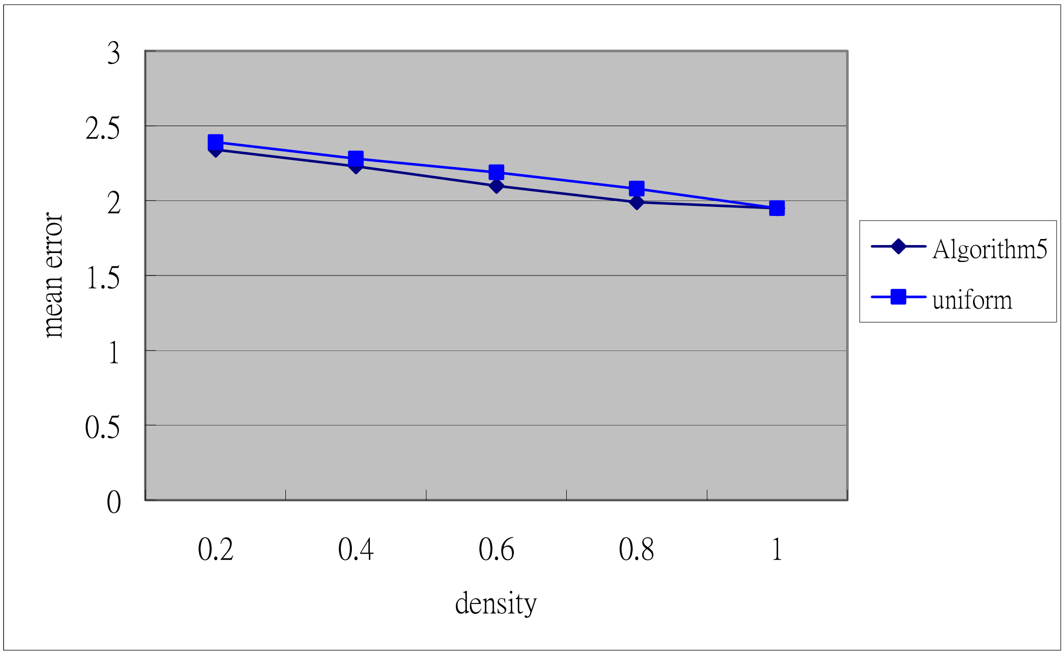

Figure 4–10 Line graph of the performance of two objective functions in test bed1

4.2 Test bed2: III Embedded System Lab

entity	Quantity
AP	5
Obstacles	573
Boundary nodes	60
Free space	352

Table 4-5 Parameters of test bed2

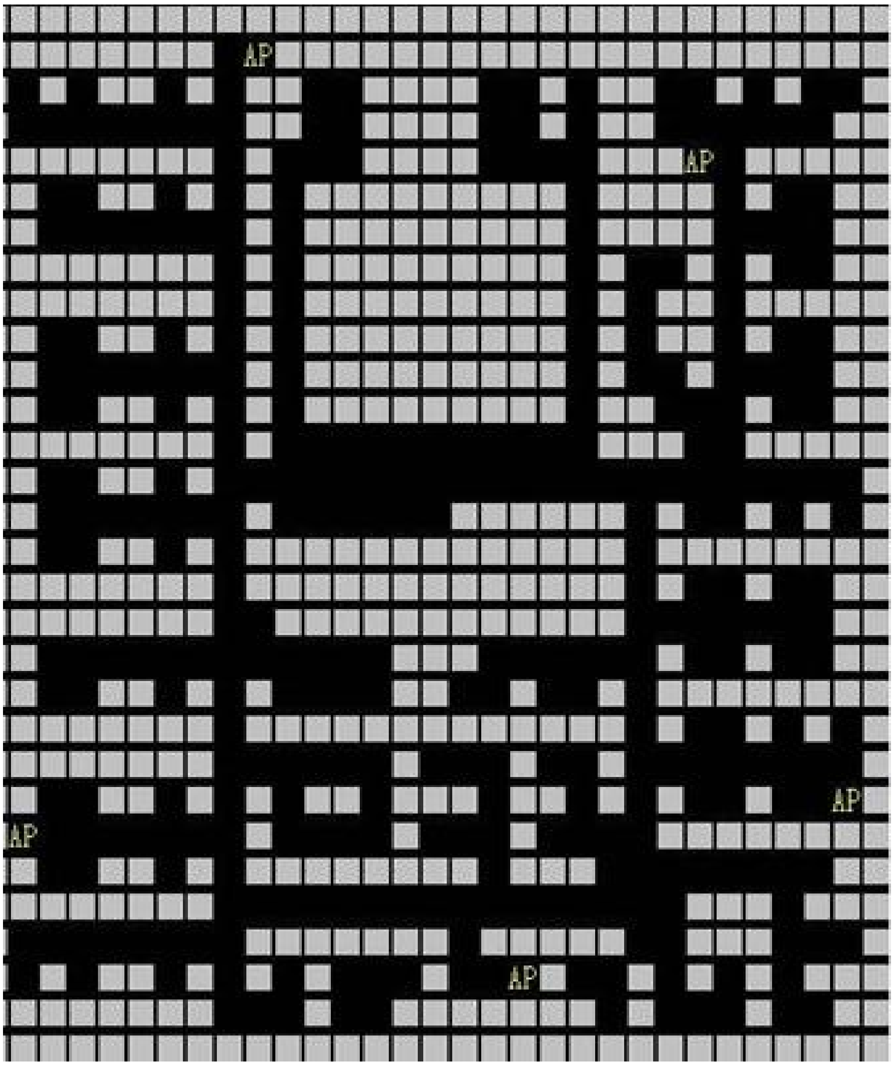

Figure 4–11 Radio map of III Embedded System Laboratory

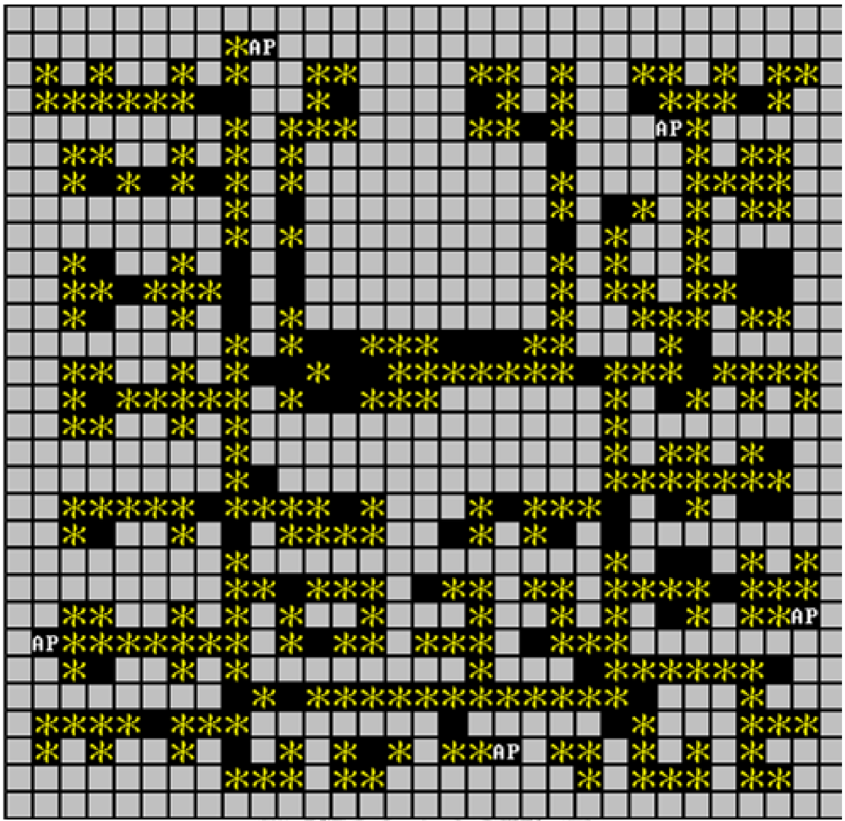

Figure 4–12 measure points with Alg5 with density: 0.8 mean error: 2.28

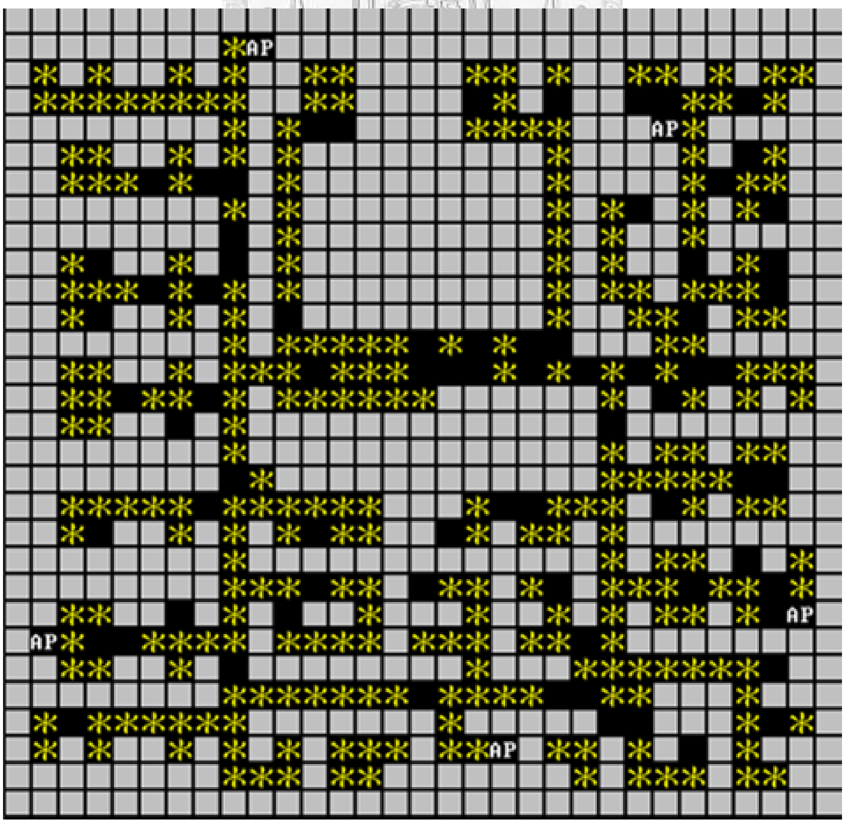

Figure 4–13 measure points with CA with density: 0.8 mean error: 2.38

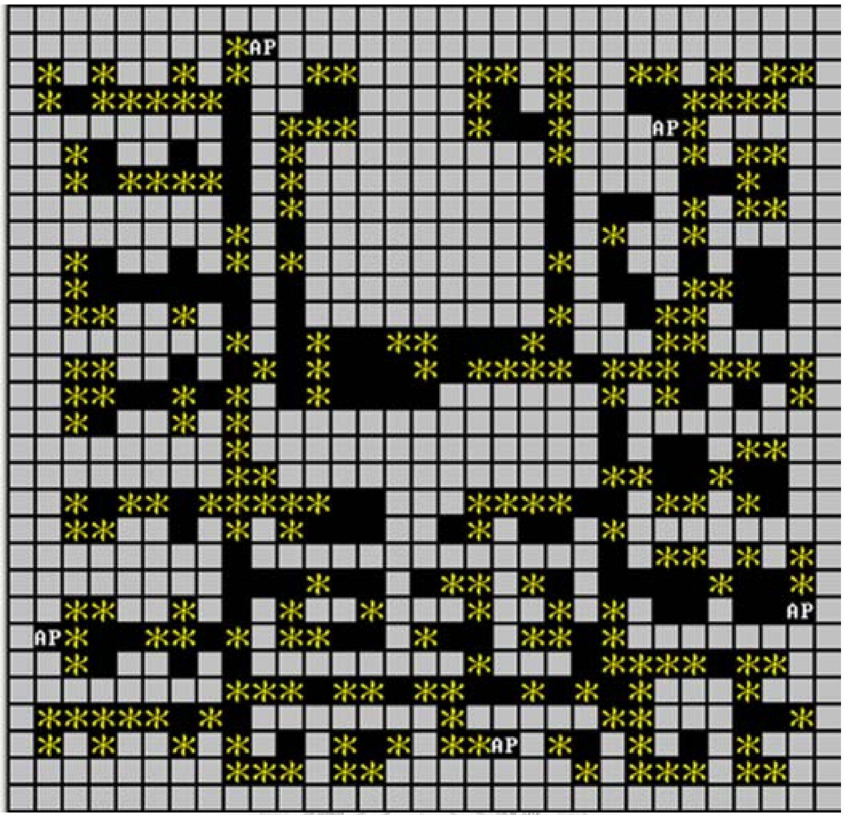

Figure 4–14 measure points with Alg5 with density: 0.6 mean error: 2.53

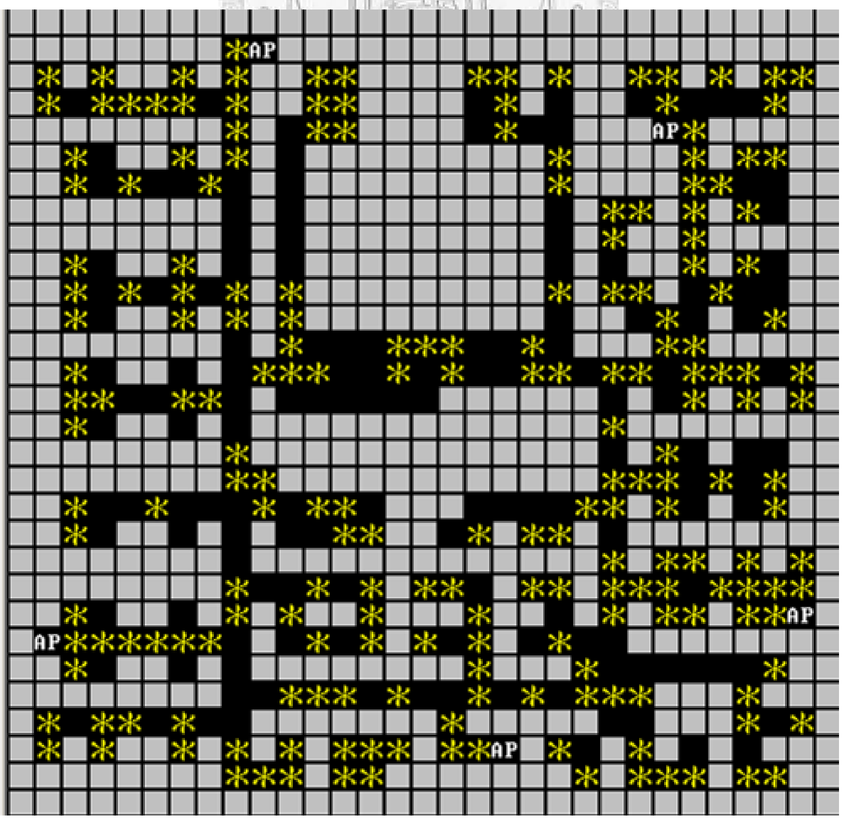

Figure 4–15 measure points with CA with density: 0.6 mean error: 2.60

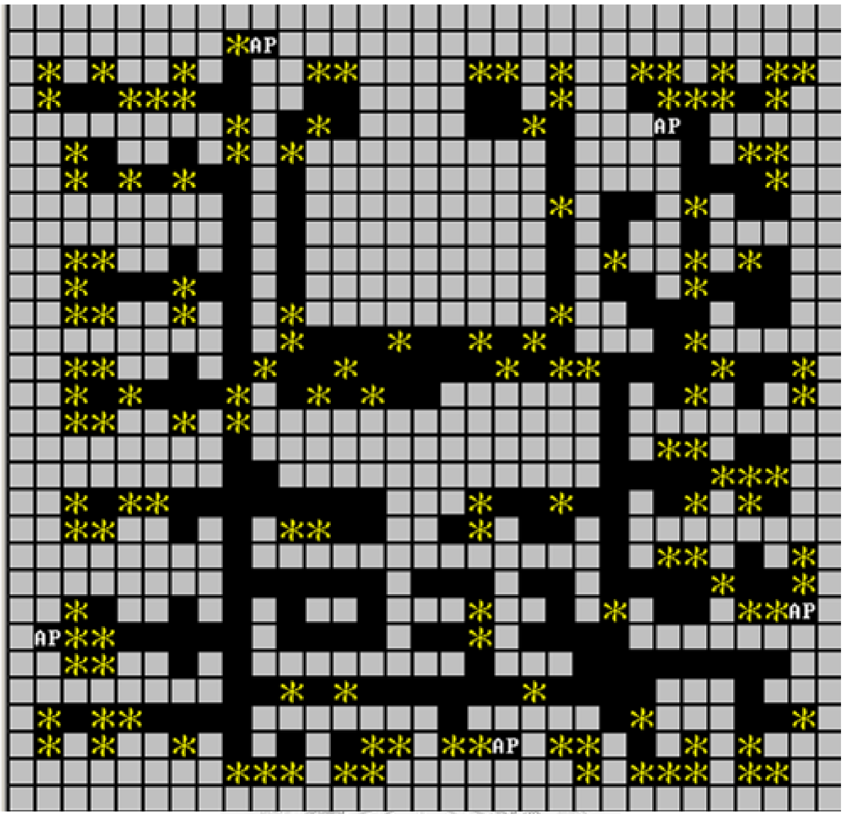

Figure 4–16 measure points with Alg5 with density: 0.4 mean error: 2.67

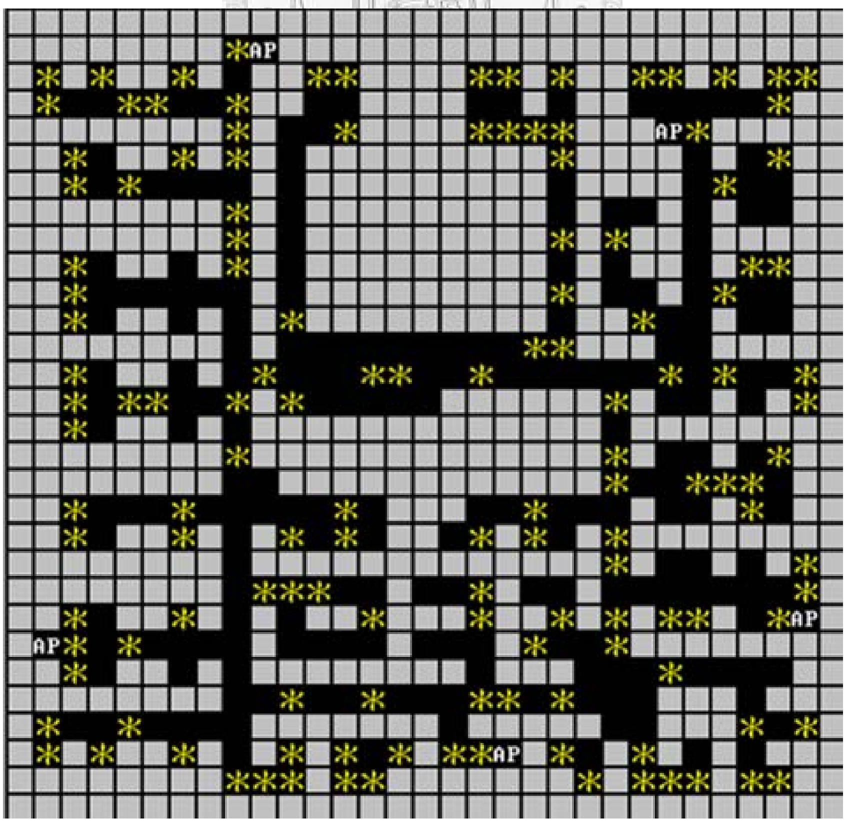

Figure 4–17 measure points with CA with density: 0.4 mean error: 2.82

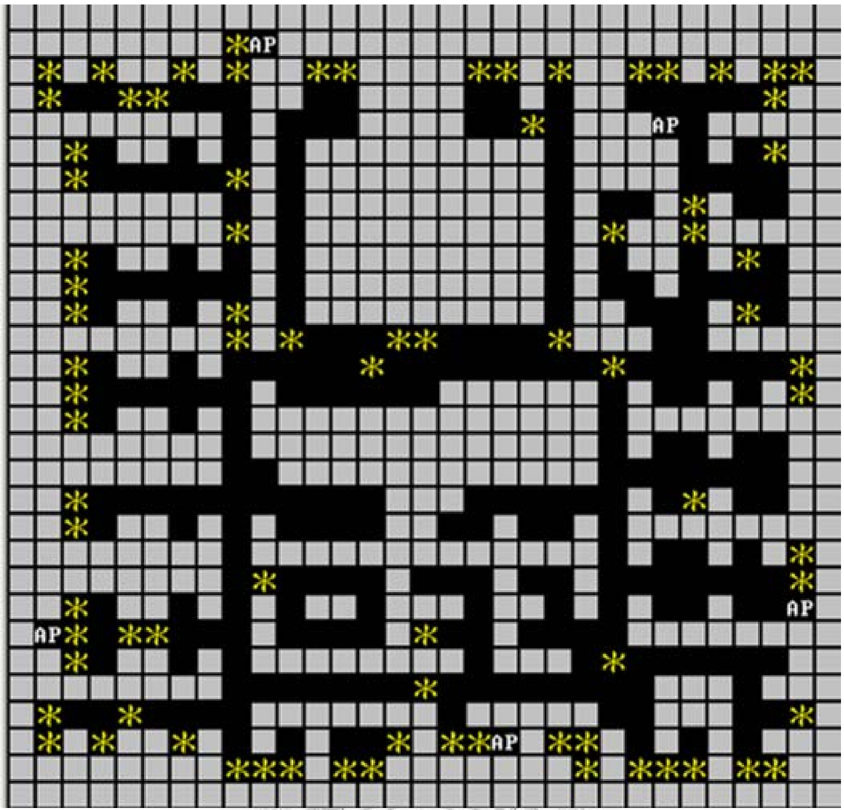

Figure 4–18 measure points with Alg5 with density: 0.2 mean error: 3.28

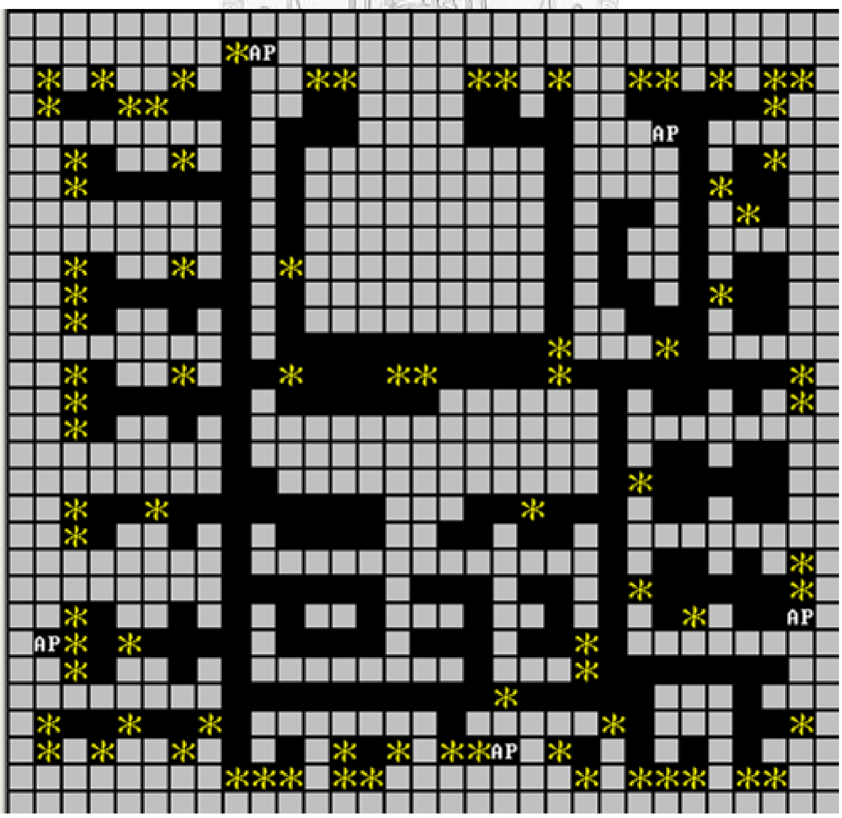

Figure 4–19 measure points with CA with density: 0.2 mean error: 3.44

	Mean error (dbm)	
density	Algorithm5	uniform
0.2	3.28	3.44
0.4	2.67	2.82
0.6	2.53	2.60
0.8	2.28	2.38
1	2.2	2.2

Table 4-6 mean error of two objective functions in test bed2

	Maximum error (dbm)	
density	Algorithm5	uniform
0.2	27	24
0.4	24	25
0.6	22	14
0.8	21	14
1	25	25

Table 4-7 maximum error of two objective functions in test bed2

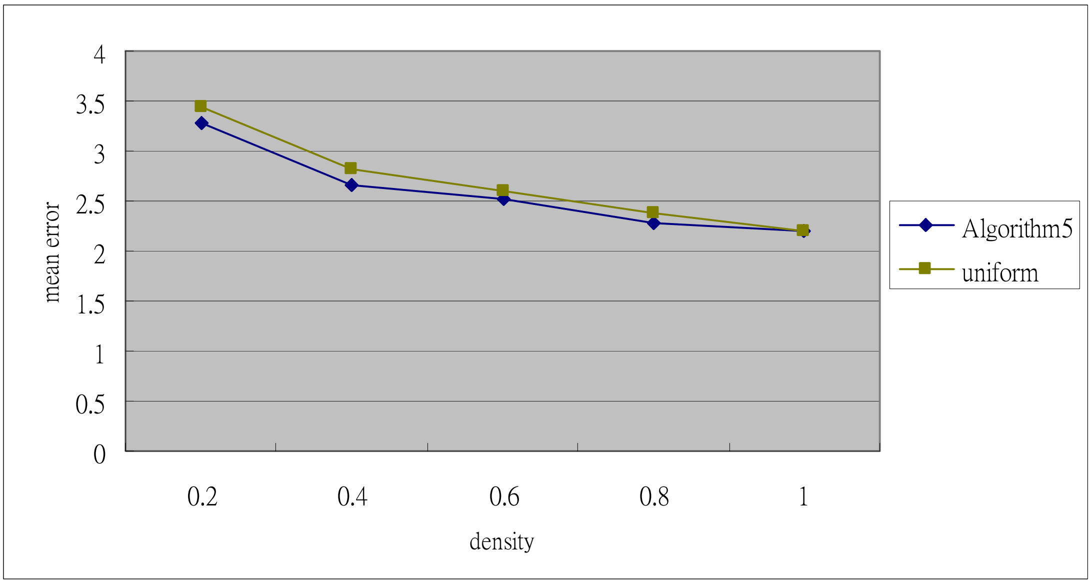

Figure 4–20 Line graph of the performance of two objective functions in test bed2

4.3 Result Discussion

According to the result of the two test bed, we can find that the results of objective function1 are all better than objective function2 in all density. The main reason why objective function1 is better is, objective function2 considers how to select measure points evenly, but do not consider the angles which determined by vector AP-to-measure point and vector AP-to-infer point. Objective function1 considers angles and distance jointly, thus has better performance than objective function2.

Chapter 5 Summary and Future Work

5.1 Summary

Indoor positioning systems can be realized in various ways. We can purely use off-the-shelf technologies, or add additional entities to improve the accuracy of positioning, or modifies wireless protocols to meet the system requirement, or even develop a new standard. There is a tradeoff between using off-the-shelf technologies and modifying them to meet system requirement more. Using more entities and modification also introduce more cost, but using purely off-the-shelf technologies requires more precise or algorithms and process to lower inaccuracy. We have developed an indoor positioning system using purely IEEE 802.11 technology without additional entities and protocol modification, and in this thesis, we proposed an idea to infer RSS based on RSS of known measure points, thus reduce additional cost of building up radio maps.

General optimization problems can be represented in mathematical forms and solved by mathematical techniques, such as Lagrangian relaxation method. But measure point selection problem in this thesis is difficult to be described in standard mathematical forms. Simulated Annealing (SA) is a substitute solution yet has very well performance solving hard combinatorial problems. We use SA to deal with the optimal measure point selection problem and get great performance. The result of SA can give us hints that in what density of measure points can we achieve satisfying positioning accuracy with acceptable RSS collecting cost. We can choose the density based on the requirement of positioning accuracy or the cost constraint to collect RSS.

5.2 Future Work

First, so far we considers every points in a field map of indoor space the same weight, which means every point in the indoor space have the same importance for mobile users. However, we can give different points the different weight, for example, the location of an antique being demonstrated is given weight 10, and the corridor toward toilet would be given 2. Hence, we can select more measure points for locations with higher weight, and fewer for lower weight.

Second, besides the location of measure points, the location of Access Points may also affect the accuracy to infer RSS for inferring points. We can consider both the location of measure points and the location of AP jointly, and find the best combination to achieve even higher accuracy. Traditional AP placement problem considers only the coverage rate, and if we consider RSS inferring accuracy jointly, then it would be a new issue for us to explore.

Reference

[1] 賴坤威，葉耿宏，陳大鈞, “IEEE802.11b 架構下之室內空間客戶定位系統之議題及解決方案,” 國立臺灣大學資訊管理學系網路通訊系統與應用實驗室, 2002 年十月。

[2] P. Bahl and V. N. Padmanabhan, "RADAR: An RF Based In-Building User Location and Tracking System," *Proc. IEEE INFOCOM, March 2000*, http://www.research.microsoft.com/sn.

[3] P. Bahl and V. N. Padmanabhan, “A Software System for Locating Mobile Users: Design, Evaluation, and Lessons,” Microsoft Research.

[4] Jeffrey Hightower and Gaetano Borriello, “Location Systems for Ubiquitous Computing,” University of Washington, Computer Science and Engineering, Box 352350, Seattle, WA 98195, August 24, 2001.

[5] Joshua A. Tauber, “Indoor Location Systems for Pervasive Computing,” 8 August 22.

[6] James J. Caffery, Jr. “A New Approach to the Geometry of TOA Location,” *IEEE Vehicular Technology Conference (VTC)*, September 2000, pp. 1943-1949..

[7] Kaveh Pahlavan and Xinrong Li, Worcester Polytechnic Institute, “Indoor Geolocation Science and Technology,” *IEEE Communication Magazine*, February 2002.

[8] T. W. Christ and P. A. Godwin, “A Prison Guard Duress Alarm Location System”, *Proc. IEEE International Carnahan Conference on Security Technology*, October 1993.

[9] S. Tekinay, “Wireless Geolocation Systems and Services,” *Special Issue of the*

IEEE Communication Magazine, April 1998.

[10] S. Y. Seidel and T. S. Rapport, “914 MHz path loss prediction Model for Indoor Wireless Communications in Multi-floored buildings,” *IEEE Trans. On Antennas & Propagation*, Feb. 1992.

[11] Kirkpatrick, S., C. D. Gelatt Jr., M. P. Vecchi, "Optimization by Simulated Annealing", *Science*, 220, 4598, 671-680, 1983.

[12] R.W. Eglese, “Simulated Annealing: A tool for Operational Research,” *European Journal of Operational Research*, vol. 46, 1990.

[13] Van Groenigen, J. W., Stein, A., 1998, “Constrained optimization of spatial sampling using continuous simulated annealing,” *J. Environ. Qual.*, Vol. 27, No. 5, pp. 1078-1086.