# The Collocation of Measurement Points in Large Open Indoor Environment

Kaikai Sheng*, Zhicheng Gu*, Xueyu Mao*, Xiaohua Tian*, Weijie Wu*, Xiaoying Gan*, Xinbing Wang†
*Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China
†School of Electronic, Info. & Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
Email: shengkaikai@sjtu.edu.cn, {zchenggu, maoxueyu}@gmail.com, {xtian, weijiewu, ganxiaoying, xwang8}@sjtu.edu.cn

***Abstract*—With the pervasion of mobile devices, crowdsourcing based received signal strength (RSS) fingerprint collection method has drawn much attention to facilitate the indoor localization since it is effective and requires no pre-deployment. However, in large open indoor environment like museums and exhibition centres, RSS measurement points cannot be collocated densely, which degrades localization accuracy. This paper focuses on measurement point collocation in different cases and their effects on localization accuracy. We first study two simple preliminary cases under assumption that users are uniformly distributed: when measurement points are collocated regularly, we propose a collocation pattern which is most beneficial to localization accuracy; when measurement points are collocated randomly, we prove that localization accuracy is limited by a tight bound. Under the general case that users are distributed asymmetrically, we show the best allocation scheme of measurement points: measurement point density $\rho$ is proportional to $(c\mu)^{2/3}$ in every part of the region, where $\mu$ is user density and $c$ is a constant determined by the collocation pattern. We also give some guidelines on collocation choice and perform extensive simulations to validate our assumptions and results.**

## I. Introduction

Different from outdoor localization that is addressed by Global Positioning System (GPS), indoor localization still remains a big challenge. Effective localization techniques using infrared [1] or RF and ultrasound [2] can achieve high localization accuracy within centimeters. However, these approaches need extra deployments which increase system cost.

With the pervasion of smartphones and ubiquity of Wi-Fi Access Points (APs), the RSS fingerprint based method is the most popular one, which avoids the cost of specialized infrastructure deployment. Typically, a RSS fingerprint based method can be divided into two phases:

• *Offline Phase:* the location fingerprints are collected in each measurement point to build the radio map by performing a site-survey of RSS from multiple APs.

• *Online Phase:* estimate the user location by utilizing localization algorithm based on the user's RSS vector and the radio map.

Since the number of measurement points is large and RSS is time-variant (See [3][4]), numerous RSS fingerprints need to be recalibrated frequently, which indicates that site-survey is both laborious and time-consuming. This problem can be addressed by employing normal users to perform site-survey instead of expert surveyors like [4][5][6][7]. This crowdsourcing based RSS fingerprints collection method is proved to be effective and inexpensive.

Generally speaking, there are mainly two types of indoor environment, which are large open environment and room-level environment, respectively. In room-level environment like the office or school laboratory, the whole region is divided into multiple small rooms or cells by obstructions (walls). In these scenarios, the collocation of measurement points and accessible region for people is restricted by the dense indoor infrastructures like printers, desks or instruments. For example, people can only walk in almost straight line through the hallway and measurement points can usually be collocated at the centers of small rooms or cells. Previous localization systems like Zee in [7] and LiFS in [8], which leverage dense indoor constraints and reference points (like elevator, stairs), can achieve high localization accuracy within meters.

However, in large open indoor environment like museums and exhibition centres, there are no or almost no obstructions (walls) in the whole region. The large open indoor environment usually has two characters. The first one is that people density is usually high, which indicates extensive requests for localization. The second is that the large open indoor environment usually has a large open area with sparse indoor infrastructures, which indicates that most of region area is accessible to people and measurement points can be collocated in a more proper way with almost no restrictions. But localization systems like Zee and LiFS may not perform well due to lack of indoor constraints and reference points. There is an imperative calling for accurate localization in these scenarios. However, it introduces the following challenges:

• *Fingerprint Similarity:* Bahl *et al.* in [9] introduced the Wall Attenuation Factor (WAF) model which describes the RSS attenuation through walls. Since there is lack of obstructions like walls, the signal pass loss is low, which results in similarity of RSS fingerprints in neighboring measurement points. So the RSS fingerprints based localization algorithm is hard to determine the user location.

• *Computation Complexity:* To get specific radio map or site-survey in large open indoor environment, numerous measurement points need to be collected, which compose a large RSS fingerprint database. Due to the low pass loss, each measurement points contains more RSS of Aps than that in room-level environment. Moreover, extensive localization requests also contribute to the computation complexity.

• *Budget Constraint:* Although crowdsourcing is an inexpensive way to collect RSS fingerprints in measurement points, the cost of site-survey is still high considering the scalability in this scenario.

Because of the above challenges, the total number of measurement points is limited and the collocation of measurement points cannot be as dense as that in room-level environment. This indicates that proper collocation of measurement points is important. In this paper, the metric used to evaluate its performance is the expected quantization location error (EQLE), which is the expected (average) distance error from the user actual location to the nearest measurement point. In online phase, algorithms such as Discrete Space Estimator in [10] and the nearest neighbor algorithm in [8], which estimate the location whose RSS fingerprint is the most similar to user's RSS vector, is a widely-used localization algorithm. Actually, if employing above algorithm, EQLE can be regarded as the lower bound of location error since it is the expected distance error even if the estimation algorithm always chooses the nearest measurement point correctly. Moreover, since the collocation of measurement points is not dense in large open indoor environment, the location error will be dominantly determined by EQLE.

In this paper, we investigate on different cases of measurement point collocation and their effects on EQLE when the number of measurement points is fixed. Specifically, we mainly have three-fold contributions as follows.

• Assuming that users are uniformly distributed in a certain region, if the measurement points are collocated regularly, we prove that EQLE can be minimized when measurement points are at the intersecting locations of a mesh of equilateral triangles. Besides, this collocation pattern can minimize the maximum quantization location error (MQLE) and maximize the distance of nearest neighbouring measurement points (DNN) which can increase performance of localization algorithm.

• Assuming that users are uniformly distributed in a certain region, if the measurement points are collocated randomly, we prove that EQLE has a lower bound and this lower bound becomes tight when number of measurement points is large.

• Under the general case that users are distributed asymmetrically, we prove that EQLE can be minimized when the density of measurement points $\rho$ is proportional to $(c\mu)^{2/3}$ in every part of the region, where $\mu$ is user density and $c$ is a constant determined by the collocation pattern in this region.

The rest of this paper is organized as follows. We present theoretical analysis for two simple preliminary cases, regular collocation and random collocation in section II and III, respectively. In section IV, we show the best allocation scheme of measurement points when users are distributed asymmetrically. In section V, we validate our assumptions and results by extensive simulations. Finally section VI offers the conclusion.

## II. Regular Collocation

In crowdsourcing based RSS fingerprints collection method, crowdworkers devote various efforts in different incentive mechanisms. With high rewards, crowdworkers can be encouraged to update RSS information at certain measurement points which have been collocated regularly in a proper way. In this section, assuming that users are uniformly distributed, we prove that EQLE and MQLE can be minimized and DNN can be maximized when measurement points are at the intersecting locations of a mesh of equilateral triangles as shown in Figure 1(a). We also show specific values of EQLE, MQLE and DNN in a mesh of equilateral triangles and a mesh of grids and compare their performance.

### A. Background and Modeling

Chen *et al.* in [11] developed a simple iterative algorithm that finds an optimized indoor landmark deployment and showed deployment pattern with up to eight landmarks in the indoor environment. However, multiple indoor landmarks increase the computation complexity of this algorithm. Actually, as is shown in Figure 1(b), the measurement points are usually collocated at the intersecting locations of a mesh of grids like [8]. Kaemarungsi *et al.* in [12] showed specific analysis on how grid spacing influences the location error. A mesh of grids can be regarded as a mesh network that two groups of parallel lines with the identical spacing intersect vertically.

In this section, we study a more general case, the regular collocation, where measurement points are at the intersecting locations of a mesh network that two groups of parallel lines with the various spacing intersect at a certain angle as shown in Figure 1(c). Users are uniformly distributed in a certain region with area[1] $S$ and $N$ measurement points. Before theoretical analysis, we give following definitions first.

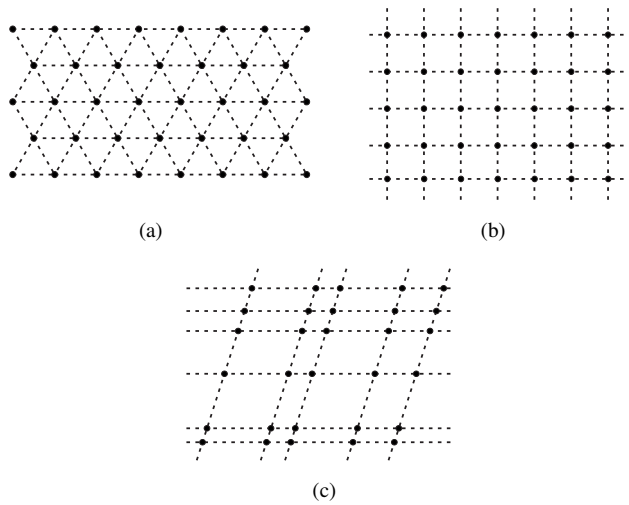



Fig. 1: (a) a mesh of equilateral triangles, (b) a mesh of grids, (c) the regular collocation.

*Definition 1:* **(neighboring region)** The neighboring region for a measurement point refers to the region which it is the nearest measurement point to any user located in.

As shown in Figure 2, the neighboring region for measurement point $M$ is surrounded by dash line. The nearest neighbor algorithm will return the estimated location $M$ for any user located in its neighboring region. The whole region can be partitioned into neighboring regions in only one way.

*Definition 2:* **(neighboring triangle)** The neighboring triangle refers to the triangle combined with three measurement points with no other measurement points in.

[1]Throughout this paper, with abuse of notation, we shall use $S$ to represent both a certain region and its area.

As shown in Figure 2, $\triangle M_1M_2M_3$ is a neighboring triangle combined with $M_1$, $M_2$ and $M_3$. Different from the neighboring region, the whole region can be partitioned into neighboring triangles in multiple ways.

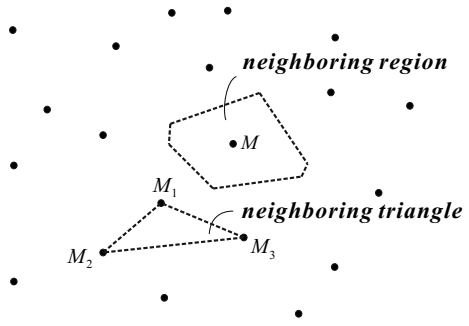



Fig. 2: Neighboring region & neighboring triangle.

*Definition 3:* **(measurement point density)** If there are $N$ measurement points in a certain region with the area $S$, the measurement point density $\mu = \frac{N}{S}$.

Intuitively, the reciprocal of measurement point density $\frac{S}{N}$ can be regarded as the average area of neighboring region and $\sqrt{\frac{S}{N}}$ can be regarded as the size scale of the neighboring region.

For mathematical tractability, we introduce the following assumption and approximation. Any indoor region has indoor infrastructures more or less, which indicates that parts of region are inaccessible to people and measurement points. Since the large open indoor environment usually has a large open area with sparse indoor infrastructures, in our theoretical analysis, we assume that there is no infrastructure and the whole region is accessible to people and measurement points. In section V, the simulation value in a large open area with sparse infrastructures has a small deviation from the theoretical value, which indicates that the influence of sparse indoor infrastructures is limited. In large open indoor environment, we also have the following proposition:

*Proposition 1:* If measurement points are collocated regularly, the ratio of the number of measurement points at the region boundary can be scaled as $\Theta\left(\frac{1}{\sqrt{N}}\right)$.

*Proof:* In the regular collocation, the perimeter can be scaled as $\Theta\left(\sqrt{S}\right)$. Since the spacing of measurement points can be scales as $\Theta\left(\sqrt{\frac{S}{N}}\right)$, the number of measurement points at the region boundary can be scaled as $\Theta\left(\sqrt{N}\right)$. So the ratio of number of these points to $N$ can be scaled as $\Theta\left(\frac{1}{\sqrt{N}}\right)$. This finishes the proof of Proposition 1. ■

In the large open indoor environment, both $S$ and $N$ are large, which indicates that the measurement points at the region boundary account for a small fraction of all measurement points. As an approximation, we will ignore the effect of measurement points at the region boundary. For example, the sum of the area of neighboring triangles is considered to be equal to that of neighboring regions. We will also verify this approximation by extensive simulations in section V.

### B. Theoretical Analysis

*Lemma 1:* The number of neighboring triangles is double of that of measurement points.

*Proof:* Since the sum of three interior angles for a single neighboring triangle is $\pi$ and a certain region can be partitioned into neighboring triangles, the number of neighboring triangles is

$$\frac{N \times 2\pi}{\pi} = 2N.$$

This finishes the proof of Lemma 1. ■

In the above proof, based on Proposition 1, we ignore the effect of measurement points at the region boundary.

*Lemma 2:* If measurement points are regularly collocated, the whole region can be partitioned into acute or right neighboring triangles.

*Proof:* The proof is straightforward. Since the regular collocation is the mesh of parallelograms and any parallelogram can be partitioned into two acute or right triangles, we can draw the above conclusion. This finishes the proof of Lemma 2. ■

*Lemma 3:* The boundaries of neighboring regions in an acute or right triangle are three midperpendiculars of triangle edges.

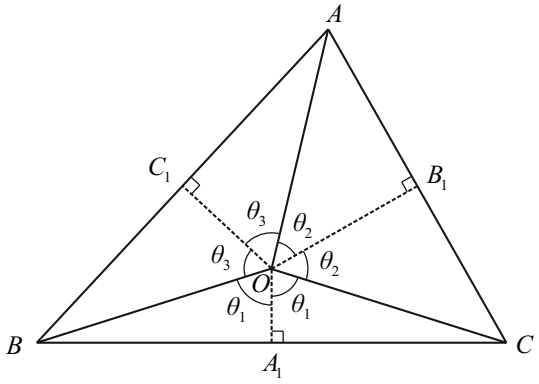



Fig. 3: Neighboring regions in the triangle.

*Proof:* In Figure 3, $OA_1$, $OB_1$ and $OC_1$ are three midperpendiculars on the $BC$, $AC$, $AB$. Using the property of the midperpendicular, for any node in region $\square BC_1OA_1$, $B$ is the nearest measurement point to it in $A$, $B$ and $C$. Hence, $\square BC_1OA_1$ is the neighboring region for $B$ in $\triangle ABC$. Similarly, we can know that $\square AC_1OB_1$ and $\square CB_1OA_1$ are the neighboring regions for $A$ and $C$, respectively. So three midperpendiculars $OA_1$, $OB_1$ and $OC_1$ are the boundaries of neighboring regions in $\triangle ABC$. This finishes the proof of Lemma 3. ■

From Lemma 3, it is apparent that the crossing point of three midperpendiculars $O$ is the circumcenter of $\triangle ABC$. Moreover, the MQLE of an acute or right triangle is the radius of its circumcircle, where the corresponding position is the circumcenter $O$.

*Lemma 4:* If a region can be partitioned into acute or right neighboring triangles, the nearest measurement point to any

node inside a neighboring triangle must be one of three triangle vertexes.

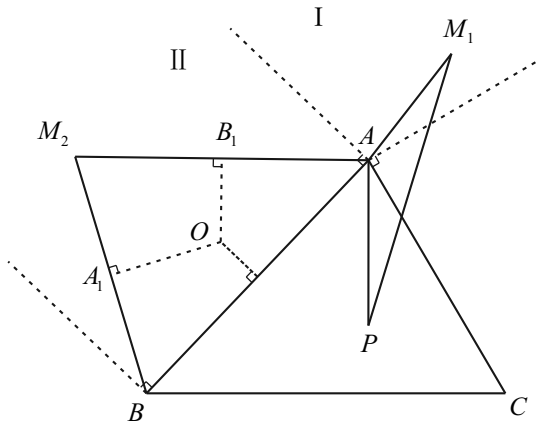



Fig. 4: Two possible positions of outside measurement points.

*Proof:* We consider two cases according to positions of the measurement point outside the neighboring triangle as follows.

**Case I:** As is shown in Figure 4, if the measurement point $M_1$ is in region I, for any node $P$ in $\triangle ABC$, $\triangle APM_1$ is an obtuse triangle ($\triangle APM_1$ may reduce to a line segment when node $P$ is in special positions). Hence, $PA < PM_1$, which indicates that compared to any other measurement point in region I, $A$ is the nearest one to any node $P$ inside $\triangle ABC$.

**Case II:** As is shown in Figure 4, if the measurement point $M_2$ is in region II, $\triangle ABM_2$ is an acute or right triangle. Using the property of the midperpendicular, $\square B_1OA_1M_2$ is the neighboring region for $M_2$ in region $\square ACBM_2$. Since $\triangle ABM_2$ is an acute or right triangle, the crossing point of three midperpendiculars (circumcenter) is inside $\triangle ABM_2$ or on edge $AB$, which indicates that $\square B_1OA_1M_2$ is totally inside $\triangle ABM_2$. Hence, for any node $P$ inside $\triangle ABC$, $PM_2$ is always longer than $PA$, $PB$ or $PC$, which means the nearest measurement point to $P$ is between $A$, $B$ or $C$.

Considering case I and II, the nearest measurement point to any node inside a neighboring triangle must be one of three triangle vertexes. This finishes the proof of Lemma 4. ■

If measurement points are regularly collocated in a certain region, using Lemma 2 and Lemma 4, the EQLE and MQLE of any neighboring triangle in this region is determined by its shape and area without being influenced by other measurement points outside the triangle.

*1) EQLE:* We introduce the Lemma 5 and Lemma 6 to simplify the proof.

*Lemma 5:* If the area of an acute or right neighboring triangle is fixed, under the assumption that users are uniformly distributed in this triangle, the EQLE can be minimized when this neighboring triangle is a equilateral triangle.

*Proof:* See Appendix A. ■

Using Jensen's Inequality, we can derive the Lemma 6.

*Lemma 6:* For $n$ nonnegative variables $x_1, x_2, \ldots, x_n$,

$$\frac{\sum_{i=1}^{n} x_i^{3/2}}{n} \geq \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^{3/2}, \tag{1}$$

where the equality holds if and only if $x_1 = x_2 = \ldots = x_n$.

*Proof:* See Appendix B. ■

*Theorem 1:* EQLE can be minimized when measurement points are at the intersecting locations of a mesh of equilateral triangles.

*Proof:* Using Lemma 2, the region in regular collocation can be partitioned into acute or right neighboring triangles. Using Lemma 1, the number of neighboring triangles is $2N$. So we can use $S_1, S_2, \ldots, S_{2N}$ to denote the area of these triangles. Equilateral triangles whose area is equal to $S_1, S_2, \ldots, S_{2N}$ is denoted by $S_{\triangle 1}, S_{\triangle 2}, \ldots, S_{\triangle 2N}$, respectively. Their edge length is denoted by $a_1, a_2, \ldots, a_{2N}$, respectively. Applying Lemma 5, we know that [2]

$$\mathbf{E_e}(S_i) \geq \mathbf{E_e}(S_{\triangle i}) = c_e a_i \ \ (i = 1, 2, \ldots, 2N), \tag{2}$$

where $c_e$ is a constant denoting the ratio of the EQLE to the edge length in equilateral triangles. Together with Lemma 2, Lemma 4 and noting that

$$S_i = S_{\triangle i} = \frac{\sqrt{3}}{4} a_i^2 \ \ (i = 1, 2, \ldots, 2N), \tag{3}$$

we have

$$\begin{aligned} \mathbf{E_e}(S) &= \left(\sum_{i=1}^{2N} S_i \mathbf{E_e}(S_i)\right)/S \\ &\geq \left(\sum_{i=1}^{2N} S_{\triangle i} \mathbf{E_e}(S_{\triangle i})\right)/S \\ &= \left(\sum_{i=1}^{2N} \frac{\sqrt{3}}{4} c_e a_i^3\right)/S. \end{aligned} \tag{4}$$

Since the whole area $S$ equals to the sum of area of neighboring triangles, noting equation (3),

$$S = \sum_{i=1}^{2N} S_i = \sum_{i=1}^{2N} S_{\triangle i} = \frac{\sqrt{3}}{4} \sum_{i=1}^{2N} a_i^2. \tag{5}$$

Using Lemma 6 and noting equation (5),

$$\begin{aligned} \left(\sum_{i=1}^{2N} \frac{\sqrt{3}}{4} c_e a_i^3\right)/S &= \frac{\sqrt{3} c_e N}{2S} \left(\frac{\sum_{i=1}^{2N} \left(a_i^2\right)^{3/2}}{2N}\right) \\ &\geq \frac{\sqrt{3} c_e N}{2S} \left(\frac{\sum_{i=1}^{2N} a_i^2}{2N}\right)^{3/2} \\ &= \frac{\sqrt{2} c_e}{\sqrt[4]{3}} \sqrt{\frac{S}{N}}. \end{aligned} \tag{6}$$

Together with equation (4),

$$\mathbf{E_e}(S) \geq \frac{\sqrt{2} c_e}{\sqrt[4]{3}} \sqrt{\frac{S}{N}}. \tag{7}$$

[2] $\mathbf{E_e}(\cdot)$ is a notation for calculating the EQLE in a region.

The equality in (7) holds when all localization triangles are equilateral triangles with the same edge length, which indicates that the measurement points are at the intersecting locations of a mesh of equilateral triangles. This finishes the proof of Theorem 1. ■

In the proof of Theorem 1, we use Lemma 5 and Lemma 6 to reduce the optimal collocation of all measurement points to the optimal choice of neighboring triangles.

*2) MQLE:* In reality, some application scenarios are sensitive to the MQLE. We will show that a mesh of equilateral triangles can also minimize MQLE. Lemma 7 is introduced to avoid formula redundancy in the proof.

*Lemma 7:* If $2\theta_1 + 2\theta_2 + 2\theta_3 = 2\pi$, then

$$\sin 2\theta_1 + \sin 2\theta_2 + \sin 2\theta_3 \leq \frac{3\sqrt{3}}{2}, \tag{8}$$

where the equality holds when $\theta_1 = \theta_2 = \theta_3 = \pi/3$.

*Proof:* See Appendix C. ■

*Theorem 2:* MQLE can be minimized when measurement points are at the intersecting locations of a mesh of equilateral triangles.

*Proof:* From Lemma 3, MQLE of an acute or right neighboring triangle is the radius of its circumcircle, where the corresponding position is the circumcenter. Together with Lemma 2 and Lemma 4, the MQLE of this region is the largest circumcircle radius of a certain neighboring triangle.

Using $S_m$ and $R_m$ to denote the largest triangle area and its corresponding radius of the circumcircle, it is obvious that[3]

$$\mathbf{E_m}(S) \geq R_m. \tag{9}$$

As is shown in Figure 3,

$$S_m = \frac{1}{2} R_m^2 \left(\sin 2\theta_1 + \sin 2\theta_2 + \sin 2\theta_3\right). \tag{10}$$

Since $S_m$ is largest neighboring triangle area, using Lemma 1,

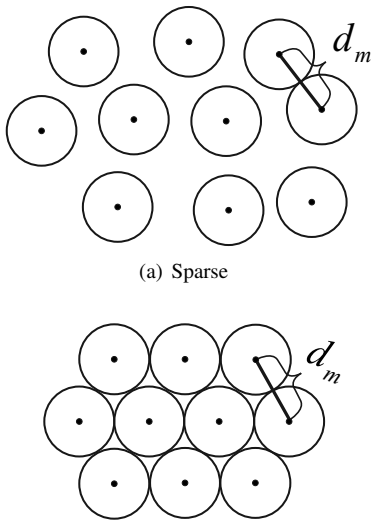
$$S_m \geq \frac{S}{2N}. \tag{11}$$

Together with (9) and (10) and using Lemma 7,

$$\begin{aligned} \mathbf{E_m}(S) &\geq R_m \\ &= \sqrt{\frac{2S_m}{\sin 2\theta_1 + \sin 2\theta_2 + \sin 2\theta_3}} \\ &\geq \sqrt{\frac{S}{N}} \cdot \sqrt{\frac{1}{\sin 2\theta_1 + \sin 2\theta_2 + \sin 2\theta_3}} \\ &\geq \frac{\sqrt{2\sqrt{3}}}{3} \sqrt{\frac{S}{N}}, \end{aligned} \tag{12}$$

where the equality holds when all neighboring triangle are equilateral triangles with identical area. This finishes the proof of Theorem 2. ■

[3] $\mathbf{E_m}(\cdot)$ is a notation for calculating the MQLE in a region.




(a) Sparse

(b) Dense

Fig. 5: Sparse packing & dense packing.

*3) DNN:* According to our discussion in section I, The low pass loss of signal in large open indoor environment results in the similarity of RSS fingerprints between neighboring measurement points, which degrades the performance of RSS fingerprint based localization algorithm. Since large DNN can make RSS more spatially varying, it is beneficial to increase performance of localization algorithm. Using the Thue's Theorem (The proof is shown in [13]), we will prove that a mesh of equilateral triangles can also maximize DNN.

*Theorem 3:* DNN can be maximized when measurement points are at the intersecting locations of a mesh of equilateral triangles.

*Proof:* This theorem is equivalent to the one that given the DNN $d_m$, the region area can be minimized when the measurement points are at the intersecting locations of a mesh of equilateral triangles.

As is shown in Figure 5(a), if we draw circles centering on each measurement point with radius $d_m/2$, there is no overlap in these circles. Recall the Thue's Theorem that regular hexagonal packing is the densest circle packing in the plane and the density of this circle configuration is $\pi/\sqrt{12}$ as shown in Figure 5(b). This indicates that

$$\frac{S_c}{S} = \frac{\pi N (d_m/2)^2}{S} \leq \frac{\pi}{\sqrt{12}}, \tag{13}$$

where $S_c$ is the sum of the area of all circles. Hence,

$$S \geq \frac{\sqrt{3}}{2} N d_m^2. \tag{14}$$

The equality in (14) holds when these circles are in a regular hexagonal packing, which is equivalent to that the measurement points are at the intersecting locations of a mesh of equilateral triangles. This finishes the proof of Theorem 3. ■

Since the proof of Theorem 3 does not base on the regular collocation condition, the Theorem 3 is a general theorem,

which indicates that DNN in a mesh of equilateral triangles is the maximum in all collocation patterns.

### C. Comparisons of Collocation Patterns

The specific values of EQLE, MQLE and DNN in a mesh of grids or equilateral triangles are shown in Table I. From this table, EQLE and MQLE in a mesh of grids are about 1.4% and 14.1% larger than those in a mesh of equilateral triangles. DNN in a mesh of grids is about 7.5% lower than that in a mesh of equilateral triangles. This indicates that a mesh of equilateral triangles is more beneficial to localization accuracy and algorithm performance than a mesh of grids.

However, the difficulty of measurement point deployment varies in different collocation patterns. Since the grid floor tiles are widely used in indoor environment, it is easy to locate measurement points in a mesh of grids. Compared to this, locating measurement points in a mesh of equilateral triangles needs extra assistance such as markers on the floor.

## III. Random Collocation

With low rewards, crowdworkers will not bother to update RSS information at certain measurement points. Actually, some indoor localization systems like Zee (See [7]) can run in the background on a smartphone without requiring any explicit crowdworker participation. In this section, we still assume that users are uniformly distributed in a region. Since users are also crowdworkers, the measurement points where they update RSS information will be randomly and dynamically collocated in the region. We will show that EQLE is lower bounded by $\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}$ and this lower bound becomes tight when number of measurement points is large.

### A. Background and Modeling

In a region with infinite area and infinite measurement points, we use $r$ to denote the distance of a user located at a certain node to its closest measurement point. For a homogeneous two-dimensional Poisson point process where measurement points are uniformly randomly collocated, the probability density function (p.d.f.) of the distance to nearest measurement point is (See [14], Chapter 8)

$$f(r) = 2\pi\mu r \cdot e^{-\mu\pi r^2}, \tag{15}$$

where $\mu$ is the measurement point density. Using the conclusion in [15], the expected distance of the user to its nearest measurement point is

$$\mathbf{E}(r) = \int_0^\infty r f(r)\, dr = \frac{1}{2\sqrt{\mu}}. \tag{16}$$

However, in the indoor environment, the region has a certain shape with finite area $S$ and finite measurement points $N$, in which case the above conclusion cannot apply to.

For convenience of analysis, we give some definitions. In a region with an arbitrary shape, the area of a certain part of the region which the distance of the user to any node located in is less than $R$ can be expressed as

$$S(R) = \int_0^R r\theta(r) dr, \tag{17}$$

where the function $\theta(r)$ denotes the sum of angle covered by the region at the distance $r$ to the user node and $\theta(r) \leq 2\pi$ as shown in Figure 6. In this region, the largest distance to the user node is denoted by $R_m$ and the total area $S$ satisfies

$$S = S(R_m) = \int_0^{R_m} r\theta(r) dr. \tag{18}$$

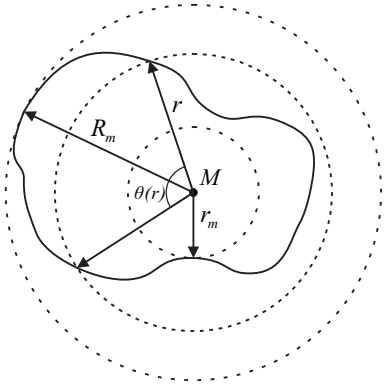



Fig. 6: $R_m$, $\theta(r)$ and $r_m$ in a region with an arbitrary shape.

### B. Theoretical Analysis

*Lemma 8:* In a region with an arbitrary shape, when the user is located at a certain node of this region, the EQLE of the region can be expressed as

$$\mathbf{E_e}(S) = \int_0^{R_m} \left(1 - \frac{S(r)}{S}\right)^N dr. \tag{19}$$

*Proof:* See Appendix D. ■

*Lemma 9:* When the user is located at a center of a circular region, the EQLE is $\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}$.

*Proof:* See Appendix E. ■

*Theorem 4:* If $N$ measurement points are uniformly randomly collocated in the region, the EQLE is lower bounded by $\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}$ and this bound becomes tight when $N$ is large.

*Proof:* We will first show that when the user is located at a certain node of a region with an arbitrary shape, the EQLE is no smaller than that when the user is located at the centre of a circular region which has identical area and identical number of measurement points. We use $S_o(R)$ and $R_{mo}$ to denote the area within the distance $R$ to the user node and the radius in the circular region.

Since the area of two regions is identical,

$$\begin{aligned} S &= S(R_m) \\ &= \int_0^{R_m} r\theta(r) dr \end{aligned} \tag{20}$$

and

$$\begin{aligned} S &= S_o(R_{mo}) \\ &= \int_0^{R_{mo}} r \cdot 2\pi dr. \end{aligned} \tag{21}$$

TABLE I: EQLE, MQLE and DNN in a mesh of grids or equilateral triangles.

| Collocation pattern | EQLE | MQLE | DNN |
|---|---|---|---|
| Equilateral triangles | $\left(\frac{\sqrt{2\sqrt{3}}}{9}+\frac{\sqrt{2\sqrt{3}}\ln 3}{12}\right)\sqrt{\frac{S}{N}}\approx 0.377\sqrt{\frac{S}{N}}$ | $\frac{\sqrt{2\sqrt{3}}}{3}\sqrt{\frac{S}{N}}\approx 0.620\sqrt{\frac{S}{N}}$ | $\frac{\sqrt{6\sqrt{3}}}{3}\sqrt{\frac{S}{N}}\approx 1.075\sqrt{\frac{S}{N}}$ |
| Grids | $\left(\frac{\sqrt{2}}{6}+\frac{1}{12}\ln\left(\frac{2+\sqrt{2}}{2-\sqrt{2}}\right)\right)\sqrt{\frac{S}{N}}\approx 0.383\sqrt{\frac{S}{N}}$ | $\frac{\sqrt{2}}{2}\sqrt{\frac{S}{N}}\approx 0.707\sqrt{\frac{S}{N}}$ | $\sqrt{\frac{S}{N}}$ |

Recalling that $\theta(r)\leq 2\pi$ and noting (20) and (21), we know that $R_m \geq R_{mo}$ and $S(r)\leq S_o(r)$ when $r\leq R_{mo}$.

Using Lemma 8 and noting that $S(r)\leq S_o(r)$ and $R_m \geq R_{mo}$, the EQLE in a circular region satisfies

$$\begin{aligned}\mathbf{E_e}\left(S_o\left(R_{mo}\right)\right) &= \int_0^{R_{mo}}\left(1-\frac{S_o(r)}{S}\right)^N dr \\ &\leq \int_0^{R_{mo}}\left(1-\frac{S(r)}{S}\right)^N dr \\ &\leq \int_0^{R_m}\left(1-\frac{S(r)}{S}\right)^N dr \\ &= \mathbf{E_e}\left(S\left(R_m\right)\right).\end{aligned} \tag{22}$$

Using Lemma 9, we know

$$\mathbf{E_e}\left(S_o\left(R_{mo}\right)\right)=\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}. \tag{23}$$

Substituting (23) into (22),

$$\mathbf{E_e}\left(S\left(R_m\right)\right)\geq\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}. \tag{24}$$

In the following, we will show that this lower bound becomes tighter with $N$ increasing. We use $r_m$ to denote the shortest distance of a user node to the region boundary as shown in Figure 6. Noting that

$$\int_0^{r_m}\left(1-\frac{S(r)}{S}\right)^N dr=\int_0^{r_m}\left(1-\frac{S_o(r)}{S}\right)^N dr, \tag{25}$$

the difference of EQLE in two regions satisfies

$$\begin{aligned}&\mathbf{E_e}\left(S\left(R_m\right)\right)-\mathbf{E_e}\left(S\left(R_{mo}\right)\right)\\ &=\int_0^{R_m}\left(1-\frac{S(r)}{S}\right)^N dr-\int_0^{R_{mo}}\left(1-\frac{S_o(r)}{S}\right)^N dr\\ &=\int_{r_m}^{R_m}\left(1-\frac{S(r)}{S}\right)^N dr-\int_{r_m}^{R_{mo}}\left(1-\frac{S_o(r)}{S}\right)^N dr\\ &\leq\int_{r_m}^{R_m}\left(1-\frac{S(r)}{S}\right)^N dr\\ &\leq\int_{r_m}^{R_m}\left(1-\frac{S(r_m)}{S}\right)^N dr\\ &=\int_{r_m}^{R_m}\left(1-\frac{\pi r_m^2}{S}\right)^N dr.\end{aligned} \tag{26}$$

From the inequality (26), when $N$ increases, the difference of EQLE in two regions will decrease exponentially. So far we have proved that if the user is located at any node of a region with an arbitrary shape, the EQLE is all lower bounded by $\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}$ and this bound becomes tighter as $N$ increases.

Hence, if users are uniformly distributed in this region, the EQLE is also lower bounded by $\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}$. This bound becomes tighter as $N$ increases, which indicates that it is a tight bound when $N$ is large. This finishes the proof of Theorem 4. ■

*Proposition 2:*

$$\lim_{N\to\infty}\frac{(2N)!!}{(2N+1)!!}\cdot 2\sqrt{\frac{N}{\pi}}=1. \tag{27}$$

*Proof:* See Appendix F. ■

From the Proposition 2, we can conclude that when $N$ is large, the lower bound approximates

$$\frac{(2N)!!}{(2N+1)!!}\cdot\sqrt{\frac{S}{\pi}}\approx\frac{1}{2}\sqrt{\frac{\pi}{N}}\cdot\sqrt{\frac{S}{\pi}}=\frac{1}{2}\sqrt{\frac{S}{N}}.$$

Together with Theorem 4, $\frac{1}{2}\sqrt{\frac{S}{N}}$ can be regarded as the approximate value for the EQLE in a region with arbitrary shape when number of measurement points is large. We find that the approximate value $\frac{1}{2}\sqrt{\frac{S}{N}}$ is exactly equal to that in equation (16). Intuitively, when $N$ is large, the nearest measurement point will extremely approach the user node, where the effect of the region shape and boundary can be ignored and the problem is reduced to the two-dimensional homogeneous Poisson point process. Compared with EQLE in regular collocation (See Table I), EQLE in random collocation is much larger than that in regular collocation. However, random collocation do not require explicit deployment of measurement points and the budget is low.

## IV. Asymmetrical Distribution

In section II and III, we hold the strong assumption that users are uniformly distributed in the region. However, in reality, user density varies in different parts of the region, which results in an asymmetrical distribution. In this section, we show the best allocation scheme of measurement points when the number of measurement points is fixed.

### A. Modeling

If users are asymmetrically distributed, the problem can be modeled as follows. The whole region $S$ can be divided into different parts denoted by $S_1, S_2, \ldots, S_l$, where $l$ is the number of parts. In each part, users are uniformly distributed, where the probability density is denoted by $\rho_1, \rho_2, \ldots, \rho_l$. Noting that the term $\rho_i S_i$ indicates the probability of users locating in the $i_{th}$ region, the area and probability density of each part satisfy

$$\sum_{i=1}^{l}\rho_i S_i=1. \tag{28}$$

Without generality, the collocation pattern of measurement points in each part can be different. Since the EQLE is proportional to the size scale of the neighboring region, the EQLE of the $i_{th}$ region is $c_i\sqrt{\frac{S_i}{N_i}}$, where $N_i$ denotes the number of measurement points allocated to the $i_{th}$ region and $c_i$ is a collocation pattern constant. As is shown in Table I, the collocation pattern constant in a mesh of equilateral triangles and a mesh of grids is approximately 0.377 and 0.383, respectively. The collocation pattern constant in a random collocation is approximately 0.5 if number of measurement points is large. Given the above definitions, we investigate on how to minimize EQLE of the whole region when the total number of measurement points $N$ is fixed.

### B. Theoretical Analysis

Using Hölder's Inequality, we can derive Lemma 10.

*Lemma 10:* For $n$ nonnegative variables $x_1, x_2, \ldots, x_n$ and $n$ nonnegative variables $y_1, y_2, \ldots, y_n$,

$$\sum_{i=1}^{n}\frac{y_i}{\sqrt{x_i}} \geq \frac{\left(\sum\limits_{i=1}^{n} y_i^{2/3}\right)^{3/2}}{\left(\sum\limits_{i=1}^{n} x_i\right)^{1/2}}, \tag{29}$$

where the equality holds if and only if

$$\frac{y_1^{2/3}}{x_1} = \frac{y_2^{2/3}}{x_2} = \ldots = \frac{y_n^{2/3}}{x_n}.$$

*Proof:* See Appendix G. ∎

*Theorem 5:* EQLE of the whole region is minimized when measurement point density $\mu_i \propto (c_i\rho_i)^{2/3}$, in which case the minimum EQLE is $\left(\sum\limits_{i=1}^{l} S_i(c_i\rho_i)^{2/3}\right)^{3/2}/N^{1/2}$.

*Proof:* EQLE of the whole region can be expressed as

$$\begin{aligned}\mathbf{E_e}(S) &= \sum_{i=1}^{l}\rho_i S_i \mathbf{E_e}(S_i)\\ &= \sum_{i=1}^{l}\rho_i S_i c_i\sqrt{\frac{S_i}{N_i}}.\end{aligned} \tag{30}$$

Using Lemma 10, the expression (30) satisfies

$$\begin{aligned}\mathbf{E_e}(S) &= \sum_{i=1}^{l}\frac{\rho_i c_i S_i^{3/2}}{\sqrt{N_i}}\\ &\geq \frac{\left(\sum\limits_{i=1}^{l}\left(\rho_i c_i S_i^{3/2}\right)^{2/3}\right)^{3/2}}{\left(\sum\limits_{i=1}^{l} N_i\right)^{1/2}}\\ &= \frac{\left(\sum\limits_{i=1}^{l} S_i(c_i\rho_i)^{2/3}\right)^{3/2}}{N^{1/2}}.\end{aligned} \tag{31}$$

The equality in (31) holds if and only if

$$\frac{S_i(c_i\rho_i)^{2/3}}{N_i} = k \;\; (i = 1, 2, \ldots, l),$$

where $k$ is a constant. This equality condition is equivalent to

$$\frac{(c_i\rho_i)^{2/3}}{\mu_i} = k \;\; (i = 1, 2, \ldots, l)$$

or $\mu_i \propto (c_i\rho_i)^{2/3}$. This finishes the proof of Theorem 5. ∎

Using the ratio constant $k$, the minimum EQLE can be simplified as

$$\frac{\left(\sum\limits_{i=1}^{l} S_i(c_i\rho_i)^{2/3}\right)^{3/2}}{N^{1/2}} = \frac{\left(\sum\limits_{i=1}^{l} S_i k u_i\right)^{3/2}}{N^{1/2}} = Nk^{3/2}.$$

As a special case in which the collocation pattern in each part is identical, we can draw a simpler conclusion that EQLE is minimized when $\mu_i \propto \rho_i^{2/3}$.

## V. Simulation

In this section, we will perform extensive simulations to validate our assumptions and results in above analysis.

### A. Regular Collocation

We create an $1\times1$ square region with 1024 (1020) measurement points in grid (equilateral triangle) collocation pattern. Figure 7(a) and Figure 7(b) show collocation patterns without indoor infrastructures while Figure 7(c) and Figure 7(d) show collocation patterns with indoor infrastructures accounting for 16% of total area. In each case, we compare the EQLE of a mesh of grids and that of a mesh of equilateral triangles. In each experiment, 100 users are distributed uniformly in the accessible region. Repeating each experiment for 1000 times, we obtain EQLE by calculating the mean value of distance of the nearest measurement points to users. Experiment results are illustrated in Table II.

As is shown in the table, when there is no indoor infrastructure, EQLE obtained by simulation is almost equal to theoretical EQLE. This validates the accuracy of approximation that the effect of measurement points at region boundary can be ignored.

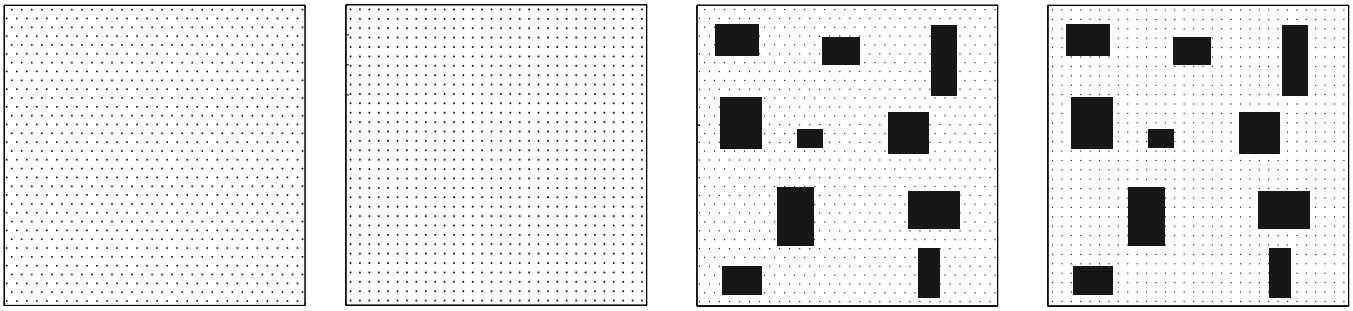
When there are indoor infrastructures which accounts for 16% of total area, EQLE obtained by simulation is about 1.7% larger than theoretical EQLE, which indicates that the influence of sparse indoor infrastructures is limited. This is because the region covered by infrastructures is inaccessible for both people and measurement points, which indicates that the measurement point density of accessible region is the same as that when there are no infrastructures. However, indoor infrastructures increase the length of inner boundary greatly, which enhances the effect of measurement points at the boundary. Since the large open indoor environment usually has a large region area with sparse indoor infrastructures, the deviation from theoretical EQLE is still limited.

### B. Random Collocation

In the simulation of random collocation, we evaluate the relationship between the EQLE and the number of measurement points in a $0.2 \times 5$ rectangular region and $1 \times 1$ square region respectively, where the simulation results are shown in Figure 8. Noting that curves in this figure is logarithmic scale for the

TABLE II: EQLE of collocation patterns in various environments.

| Collocation Pattern | Theoretical | No infrastructures | Infrastructures |
|---|---|---|---|
| Equilateral triangles | 0.011810 | 0.011810 | 0.011997 |
| Grids | 0.011956 | 0.011955 | 0.012185 |



(a) A mesh of equilateral triangles without infrastructures.

(b) A mesh of grids without infrastructures.

(c) A mesh of equilateral triangles with infrastructures.

(d) A mesh of grids with infrastructures.

Fig. 7: Different collocation patterns in various environments.

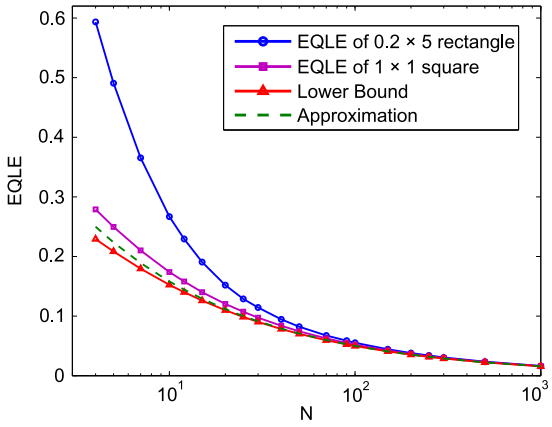



Fig. 8: EQLE versus the number of measurement points.

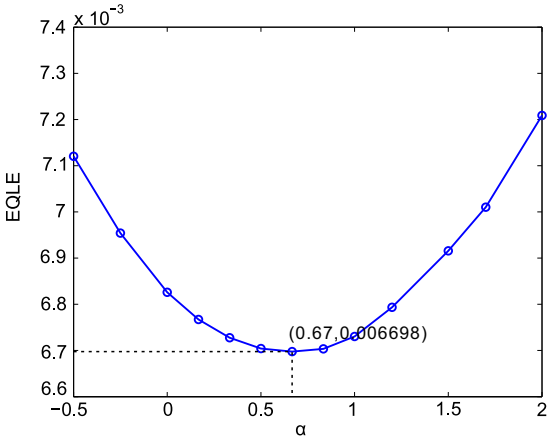



Fig. 9: EQLE of the region when measurement points are allocated following $\mu_i \propto \rho_i^{\alpha}$.

x-axis and linear scale for the y-axis, the EQLE in two region approaches the lower bound fast with $N$ increasing. EQLE of $1 \times 1$ square is always lower than that of $0.2 \times 5$ rectangular since users will have less probability to locate near the region boundary. When $N$ is large, both the EQLE and the lower bound will be approximately equal to $\frac{1}{2}\sqrt{\frac{S}{N}}$, which validates our results.

### C. Asymmetrical Distribution

In this simulation, we create an $1 \times 2$ rectangular region, where 1000 users locate in left half part with probability 0.9 and locate in the other half part with with probability 0.1. We use the same collocation pattern for both parts, which indicates that collocation pattern constants in two parts are identical. We allocate measurement points following

$$\mu_i \propto \rho_i^{\alpha},$$

where $\alpha$ is the independent variable. Simulation results are shown in Figure 9, where each data point in the simulation is derived from 1000 experiments. In this figure, the EQLE is minimized when $\alpha = 0.67$, which validates our theoretical analysis in section IV.

## VI. Conclusion and Future Work

In this paper, we mainly investigate on the collocation of measurement points in large open indoor environment by employing EQLE as the metric. Assuming that users are uniformly distributed, we first study two simple preliminary cases: (1) When the measurement points are collocated regularly, EQLE can be minimized when measurement points are at the intersecting locations of a mesh of equilateral triangles. Besides, this collocation pattern can minimize MQLE and maximize DNN which is beneficial to increase performance of localization algorithm. (2) When the measurement points are collocated randomly, EQLE has a lower bound and this lower bound becomes tight when the number of measurement points is large.

We develop the collocation scheme under the general case that users are distributed asymmetrically. We draw a simple conclusion that EQLE is minimized when $\rho$ is proportional to the $(c\mu)^{2/3}$ in each part, where $\mu$ is the user density and $c$ is the collocation pattern constant. And in each part, we can choose proper collocation pattern considering the difficulty of deployment, target localization accuracy and budget.

In the future work, we will study if a mesh of equilateral triangles can minimize both EQLE and MQLE in any collocation patterns, not only in regular collocation patterns. Furthermore, we derive our results based on the localization algorithm which chooses "closest point or nearest neighbor in signal space (see [16]). However, k-nearest neighbor (often referred as "KNN") localization algorithm (see [17]) has been also widely employed in indoor localization systems like [18]. The KNN algorithm first choose $k$ nearest measurement points, and then calculate their center. Enhanced KNN algorithms are like EWKNN (see [19]) which change the number of neighbors to improve localization accuracy or CFK (see [20]) which uses clustering technique to divide neighbors into different clusters and chooses one cluster to calculate the estimated position. Other algorithm which is referred as "smallest polygon" (see [21], [22]) will choose some nearest neighbors which can form different polygons and recognize the centre of the smallest polygon as the estimated position. So it is important to derive the optimal collocation patterns of measurement points when utilizing these localization algorithms.

## VII. Acknowledgement

This work is supported by NSF China (No. 61325012, 61271219, 61221001, 61202373, 61102052, 61102051, 61428205); China Ministry of Education Doctor Program (No. 20130073110025); SEU SKL project (No. 2012D13, No. 2014D07); Shanghai Basic Research Key Project (No. 13510711300, 12JC1405200, 11JC1405100); Shanghai International Cooperation Project: (No. 13510711300).

## References

[1] E. Aitenbichler and M. Muhlhauser, "An IR local positioning system for smart items and devices," in *Proc. IEEE ICDCS*, 2003, pp. 334–339.

[2] N. B. Priyantha, "The cricket indoor location system," Ph.D. dissertation, MIT, 2005.

[3] K. Kaemarungsi and P. Krishnamurthy, "Properties of indoor received signal strength for wlan location fingerprinting," in *Proc. ACM MOBIQUTIOUS*, 2004, pp. 14–23.

[4] S. Yang, P. Dessai, M. Verma, and M. Gerla, "FreeLoc: Calibration-free crowdsourced indoor localization," in *Proc. IEEE INFOCOM*, 2013, pp. 2481–2489.

[5] P. Bolliger, "Redpin-adaptive, zero-configuration indoor localization through user collaboration," in *Proc. ACM MELT*, 2008, pp. 55–60.

[6] S. Teller, J. Battat, B. Charrow, D. Curtis, R. Ryan, J. Ledlie, and J. Hicks, "Organic indoor location discovery," *Technical Report CSAIL TR-2008-075, MIT*, vol. 75, p. 16, 2008.

[7] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: zero-effort crowdsourcing for indoor localization," in *Proc. ACM Mobicom*, 2012, pp. 293–304.

[8] Z. Yang, C. Wu, and Y. Liu, "Locating in fingerprint space: wireless indoor localization with little human intervention," in *Proc. ACM Mobicom*, 2012, pp. 269–280.

[9] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM*, vol. 2, 2000, pp. 775–784.

[10] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proc. ACM MobiSys*, 2005, pp. 205–218.

[11] Y. Chen, J.-A. Francisco, W. Trappe, and R. P. Martin, "A practical approach to landmark deployment for indoor localization," in *Proc. IEEE SECON*, vol. 1, 2006, pp. 365–373.

[12] K. Kaemarungsi and P. Krishnamurthy, "Modeling of indoor positioning systems based on location fingerprinting," in *Proc. IEEE INFOCOM*, vol. 2, 2004, pp. 1012–1022.

[13] H.-C. Chang and L.-C. Wang, "A simple proof of Thue's Theorem on circle packing," *arXiv preprint arXiv:1009.4322*, 2010.

[14] N. Cressie, "Statistics for spatial data," *Terra Nova*, vol. 4, no. 5, pp. 613–617, 1992.

[15] C. Bettstetter, "On the minimum node degree and connectivity of a wireless multihop network," in *Proc. ACM MobiHoc*, 2002, pp. 80–91.

[16] A. G. Dempster, B. Li, and I. Quader, "Errors in determinstic wireless fingerprinting systems for localisation," in *Proc. IEEE ISWPC*, 2008, pp. 111–115.

[17] A. Hossain, H. N. Van, Y. Jin, and W.-S. Soh, "Indoor localization using multiple wireless technologies," in *Proc. IEEE MASS*, 2007, pp. 1–8.

[18] C.-N. Huang and C.-T. Chan, "ZigBee-based indoor location system by k-nearest neighbor algorithm with weighted RSSI," *Procedia Computer Science*, vol. 5, pp. 58–65, 2011.

[19] B. Shin, J. H. Lee, T. Lee, and H. S. Kim, "Enhanced weighted K-nearest neighbor algorithm for indoor Wi-Fi positioning systems," in *Proc. IEEE ICCM*, vol. 2, 2012, pp. 574–577.

[20] J. Ma, X. Li, X. Tao, and J. Lu, "Cluster filtered KNN: A WLAN-based indoor positioning scheme," in *Proc. IEEE WoWMoM*, 2008, pp. 1–8.

[21] E. Martin, O. Vinyals, G. Friedland, and R. Bajcsy, "Precise indoor localization using smart phones," in *Proc. ACM MM*, 2010, pp. 787–790.

[22] A. Roxin, J. Gaber, M. Wack, and A. Nait-Sidi-Moh, "Survey of wireless geolocation techniques," in *Proc. IEEE Globecom Workshops*, 2007, pp. 1–9.

## Appendix A
## Proof of Lemma 5

Before proof of Lemma 5, we first calculate the EQLE of an arc with radius $r$ as shown in Figure 10.

The EQLE in the region $\widehat{AB}$ is

$$\begin{aligned}\mathbf{E_e}\left(\widehat{AB}\right) &= \frac{\int_0^\theta \int_0^r x^2 dxdy}{S_{\widehat{AB}}} \\ &= \frac{1}{3}r^3\theta / \frac{1}{2}r^2\theta \\ &= \frac{2}{3}r,\end{aligned} \tag{32}$$

where $\theta$ denotes the radian of the arc. From the above analysis, we know that no matter what the radian of the arc is, the EQLE of an arc is only determined by its radius $r$.

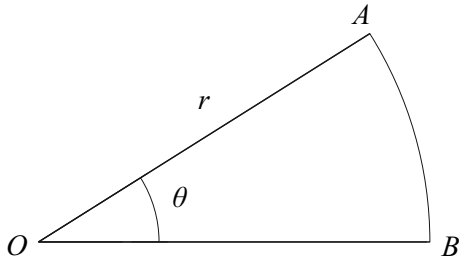



Fig. 10: An arc with radius $r$.

Equipped with the above conclusion, we can show the proof of Lemma 5.

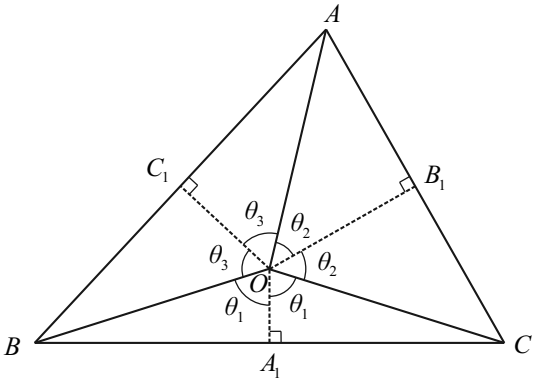



Fig. 11: A neighboring triangle.

From Lemma 3, the boundaries of neighboring region in an acute or right triangle is the three midperpendiculars of the triangle edges. As is shown in Figure 11, the crossing point of three midperpendiculars (circumcenter) is $O$ and $\angle BOA_1 = \angle COA_1 = \theta_1$, $\angle COB_1 = \angle AOB_1 = \theta_2$, $\angle AOC_1 = \angle BOC_1 = \theta_3$, $|OA| = |OB| = |OC| = R$.

In the $\triangle BOA_1$, based on the conclusion of above analysis, using $\mathbf{E_{e|B}}(BOA_1)$ to denote the expected distance deviated from $B$ in $\triangle BOA_1$,

$$\mathbf{E_{e|B}}(\triangle BOA_1) = \lim_{N\to\infty}\left(\sum_{i=1}^{N}\frac{2}{3}\sqrt{|BA_1|^2+(i\cdot\Delta h)^2}\right)/N, \tag{33}$$

where $N\Delta h = |OA_1|$. Transforming (33) to the integral form,

$$\begin{aligned}
&\mathbf{E_{e|B}}(\triangle BOA_1) = \frac{2}{3|OA_1|}\int_0^{|OA_1|}\sqrt{|BA_1|^2+h^2}dh\\
&=\frac{2}{3|OA_1|}\left(\frac{|OA_1||OB|}{2}+\frac{|BA_1|^2}{2}\ln\left(\frac{|OA_1|+|OB|}{|BA_1|}\right)\right)\\
&=\frac{|OB|}{3}+\frac{|BA_1|^2}{3|OA_1|}\ln\left(\frac{|OA_1|+|OB|}{|BA_1|}\right)\\
&=\frac{R}{3}+\frac{Rsin^2\theta_1}{3\cos\theta_1}\ln\left(\cot\frac{\theta_1}{2}\right).
\end{aligned} \tag{34}$$

Noting that

$$\begin{cases}
S_{\triangle BOA_1} = S_{\triangle COA_1} = R^2\sin 2\theta_1/4\\
S_{\triangle COB_1} = S_{\triangle AOB_1} = R^2\sin 2\theta_2/4\\
S_{\triangle AOC_1} = S_{\triangle BOC_1} = R^2\sin 2\theta_3/4\\
S_{\triangle ABC} = R^2(\sin 2\theta_1+\sin 2\theta_2+\sin 2\theta_3)/2
\end{cases},$$

the EQLE of $\triangle ABC$ is

$$\begin{aligned}
&\mathbf{E_e}(\triangle ABC) = \frac{2R}{3(\sin 2\theta_1+\sin 2\theta_2+\sin 2\theta_3)}\cdot\\
&\left[sin^3\theta_1\ln\left(\cot\frac{\theta_1}{2}\right)+sin^3\theta_2\ln\left(\cot\frac{\theta_2}{2}\right)+\right.\\
&\left.sin^3\theta_3\ln\left(\cot\frac{\theta_3}{2}\right)\right]+\frac{R}{3},
\end{aligned} \tag{35}$$

where $\theta_1+\theta_2+\theta_3 = \pi$. Note that

$$S_{\triangle ABC} = R^2(\sin 2\theta_1+\sin 2\theta_2+\sin 2\theta_3)/2,$$

and together with (35),

$$\begin{aligned}
&\mathbf{E_e}(ABC) = \frac{1}{3}\sqrt{\frac{2S_{\triangle ABC}}{\sin 2\theta_1+\sin 2\theta_2+\sin 2\theta_3}}\\
&+\frac{2}{3}\sqrt{\frac{2S_{\triangle ABC}}{(\sin 2\theta_1+\sin 2\theta_2+\sin 2\theta_3)^3}}\cdot\\
&\left[sin^3\theta_1\ln\left(\cot\frac{\theta_1}{2}\right)+sin^3\theta_2\ln\left(\cot\frac{\theta_2}{2}\right)\right.\\
&\left.+sin^3\theta_3\ln\left(\cot\frac{\theta_3}{2}\right)\right].
\end{aligned} \tag{36}$$

If the area of $\triangle ABC$ is fixed, the EQLE can be minimized when $\theta_1 = \theta_2 = \theta_3 = \pi/3$, which indicates that the neighboring triangle is an equilateral triangle.

## Appendix B
## Proof of Lemma 6

We let $f(x) = x^{3/2}$ $(x\geq 0)$. The second-order derivative of $f(x)$ is

$$f''(x) = \frac{3}{4}x^{-1/2}. \tag{37}$$

Since $f''(x) > 0$, $f(x)$ is a concave function. Using Jensen Inequality,

$$\begin{aligned}
\frac{\sum_{i=1}^{n}x_i^{3/2}}{n} &= \frac{\sum_{i=1}^{n}f(x_i)}{n}\\
&\geq f\left(\frac{\sum_{i=1}^{n}x_i}{n}\right)\\
&= \left(\frac{\sum_{i=1}^{n}x_i}{n}\right)^{3/2},
\end{aligned} \tag{38}$$

where the equality holds if and only if $x_1 = x_2 = \ldots = x_n$.

## Appendix C
## Proof of Lemma 7

Noting that $2\theta_1+2\theta_2+2\theta_3 = 2\pi$ and the AM-GM inequality,

$$\begin{aligned}
&\sin 2\theta_1+\sin 2\theta_2+\sin 2\theta_3\\
=&\sin 2\theta_1+\sin 2\theta_2-\sin(2\theta_1+2\theta_2)\\
=&2\sin(\theta_1+\theta_2)(\cos(\theta_1-\theta_2)-\cos(\theta_1+\theta_2))\\
\leq&2\sin(\theta_1+\theta_2)(1-\cos(\theta_1+\theta_2))\\
=&2\sqrt{1-\cos^2(\theta_1+\theta_2)}(1-\cos(\theta_1+\theta_2))\\
=&2\sqrt{27\left(\frac{1}{3}-\frac{1}{3}\cos(\theta_1+\theta_2)\right)^3(1+\cos(\theta_1+\theta_2))}\\
\leq&2\sqrt{\frac{27}{16}}\\
=&\frac{3\sqrt{3}}{2}.
\end{aligned} \tag{39}$$

The equality in (39) holds when

$$\begin{cases} \cos(\theta_1-\theta_2)=1 \\ \frac{1}{3}-\frac{1}{3}\cos(\theta_1+\theta_2)=1+\cos(\theta_1+\theta_2) \end{cases},$$

where $\theta_1=\theta_2=\theta_3=\pi/3$.

## Appendix D
## Proof of Lemma 8

Since $N$ measurement points are uniformly randomly collocated in the region, the probability that the distance of a measurement point to the user node $d$ is no larger than $r$ can be expressed as

$$P(d\le r)=\frac{S(r)}{S}. \tag{40}$$

So the cumulative distribution function (c.d.f.) for the distance of the nearest measurement point to the user node is

$$F(r)=1-\left(1-\frac{S(r)}{S}\right)^N. \tag{41}$$

Hence, the p.d.f. for the distance of the nearest measurement point to the user node is

$$\begin{aligned} f(r)&=F'(r)\\ &=\frac{NS'(r)}{S}\left(1-\frac{S(r)}{S}\right)^{N-1}. \end{aligned} \tag{42}$$

Noting that $f(R_m)=0$, the EQLE is

$$\begin{aligned} \mathbf{E_e}(S)&=\int_0^{R_m} rf(r)\,dr\\ &=\int_0^{R_m}\frac{rNS'(r)}{S}\left(1-\frac{S(r)}{S}\right)^{N-1}dr\\ &=\int_0^{R_m}(-r)\,d\left(1-\frac{S(r)}{S}\right)^N\\ &=(-r)\left(1-\frac{S(r)}{S}\right)^N\Big|_0^{R_m}+\\ &\quad\int_0^{R_m}\left(1-\frac{S(r)}{S}\right)^N dr\\ &=\int_0^{R_m}\left(1-\frac{S(r)}{S}\right)^N dr. \end{aligned} \tag{43}$$

This finishes the proof of Lemma 8.

## Appendix E
## Proof of Lemma 9

If the user is located at the centre of the circular region,

$$\begin{aligned} S(r)&=\int_0^r x\theta(x)dx\\ &=\int_0^r x\cdot 2\pi dx\\ &=\pi r^2 \end{aligned} \tag{44}$$

and

$$R_m=\sqrt{\frac{S}{\pi}}. \tag{45}$$

Substituting (44) and (45) into the (19) of Lemma 8,

$$\begin{aligned} \mathbf{E_e}(S)&=\int_0^{R_m}\left(1-\frac{S(r)}{S}\right)^N dr\\ &=\int_0^{\sqrt{S/\pi}}\left(1-\frac{\pi r^2}{S}\right)^N dr. \end{aligned} \tag{46}$$

Using $r=\sin\alpha\sqrt{S/\pi}$ to replace $r$ in (46),

$$\begin{aligned} \mathbf{E_e}(S)&=\int_0^{\pi/2}\left(1-\sin^2\alpha\right)^N d\left(\sin\alpha\sqrt{S/\pi}\right)\\ &=\sqrt{\frac{S}{\pi}}\int_0^{\pi/2}\left(\cos^2\alpha\right)^N d(\sin\alpha)\\ &=\sqrt{\frac{S}{\pi}}\int_0^{\pi/2}(\cos\alpha)^{2N+1}d\alpha\\ &=\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}. \end{aligned} \tag{47}$$

According to the Wallis' Formula,

$$\int_0^{\pi/2}(\cos\alpha)^n d\alpha=\begin{cases}\frac{\pi}{2}\cdot(n-1)!!/n!!, & n\,is\,even\\ (n-1)!!/n!!, & n\,is\,odd\end{cases}$$

Hence we can calculate that

$$\mathbf{E_e}(S)=\frac{(2N)!!}{(2N+1)!!}\sqrt{\frac{S}{\pi}}. \tag{48}$$

## Appendix F
## Proof of Proposition 2

From Stirling's approximation, we know that

$$\lim_{N\to\infty}\frac{N!}{(N/e)^N\sqrt{2\pi N}}=1. \tag{49}$$

Noting the (49), we have

$$\begin{aligned} &\lim_{N\to\infty}\frac{(2N)!!}{(2N+1)!!}\cdot 2\sqrt{\frac{N}{\pi}}\\ =&\lim_{N\to\infty}\frac{((2N)!!)^2}{(2N+1)!}\cdot 2\sqrt{\frac{N}{\pi}}\\ =&\lim_{N\to\infty}\frac{2^{2N}\cdot(N!)^2}{(2N+1)!}\cdot 2\sqrt{\frac{N}{\pi}}\\ =&\lim_{N\to\infty}\frac{2^{2N}\cdot(N/e)^{2N}\cdot 2\pi N}{((2N+1)/e)^{2N+1}\sqrt{2\pi(2N+1)}}\cdot 2\sqrt{\frac{N}{\pi}}\\ =&\lim_{N\to\infty}\frac{(2N/e)^{2N}\cdot 2\pi N}{((2N+1)/e)^{2N+1}\sqrt{2\pi(2N+1)}}\cdot 2\sqrt{\frac{N}{\pi}}\\ =&\lim_{N\to\infty}\left(\frac{2N/e}{(2N+1)/e}\right)^{2N}\cdot\lim_{N\to\infty}\frac{2\pi N}{(2N+1)/e}\\ &\cdot\lim_{N\to\infty}\frac{2\sqrt{N/\pi}}{\sqrt{2\pi(2N+1)}}\\ =&\frac{1}{e}\cdot\pi e\cdot\frac{1}{\pi}\\ =&1. \end{aligned} \tag{50}$$

## APPENDIX G
## PROOF OF LEMMA 10

Using Hölders Inequality,

$$\left(\sum_{i=1}^{n}\frac{y_i}{\sqrt{x_i}}\right)^{2/3}\left(\sum_{i=1}^{n}x_i\right)^{1/3}\geq\left(\sum_{i=1}^{n}y_i^{2/3}\right)^{3/2}, \tag{51}$$

where the equality holds if and only if

$$\frac{y_1^{2/3}}{x_1}=\frac{y_2^{2/3}}{x_2}=\ldots=\frac{y_n^{2/3}}{x_n}.$$

Derived from the (51), we have

$$\sum_{i=1}^{n}\frac{y_i}{\sqrt{x_i}}\geq\frac{\left(\sum\limits_{i=1}^{n}y_i^{2/3}\right)^{3/2}}{\left(\sum\limits_{i=1}^{n}x_i\right)^{1/2}}.$$