Achieving Range-free Localization Beyond Connectivity

Ziguo Zhong
Department of Computer Science
University of Minnesota
zhong@cs.umn.edu

Tian He
Department of Computer Science
University of Minnesota
tianhe@cs.umn.edu

Abstract

Wireless sensor networks have been proposed for many location-dependent applications. In such applications, the requirement of low system cost prohibits many range-based methods for sensor node localization; on the other hand, range-free localization depending only on connectivity may *underutilize* the proximity information embedded in neighborhood sensing. In response to the above limitations, this paper presents a range-free approach to capturing a *relative distance* between 1-hop neighboring nodes from their *neighborhood orderings* that serve as unique high-dimensional location signatures for nodes in the network. With little overhead, the proposed design can be conveniently applied as a transparent supporting layer for many state-of-the-art connectivity-based localization solutions to achieve better positioning accuracy. We implemented our design with three well-known localization algorithms and tested it in two types of outdoor test-bed experiments: an 850-foot-long linear network with 54 MICAz motes, and a regular 2D network covering an area of 10000 square feet with 49 motes. Results show that our design helps eliminate estimation ambiguity with sub-hop resolution, and reduces localization errors by as much as 35%. In addition, extensive simulations reveal an interesting feature of robustness for our design under unevenly distributed radio propagation path loss, and confirm its effectiveness for large-scale networks.

Categories and Subject Descriptors

C.2.4 [**Computer-Communication Networks**]: Distributed Systems

General Terms

Algorithms, Design, Performance, Experimentation

Keywords

Wireless Sensor Networks, Range-free Localization

1 Introduction

Wireless sensor networks (WSN) have been considered as a promising tool for many location-dependent applications [1, 2], e.g., battlefield surveillance [3], environment data collection [4], event or human localization [5, 6]. In addition, some of the routing protocols [7, 8] and network management mechanisms proposed for such networks are built on the assumption that geographic parameters of each sensor node are available. Although sensor node localization plays an important role in all those systems, it is itself a challenging problem due to extremely limited resources available at each low-cost and tiny sensor node.

Many ideas have been proposed for node localization in WSN. Based on whether accurate ranging is required, there are basically two types of methods: (i) range-based localization and (ii) range-free localization. Range-based localization could achieve good accuracy but costly for requiring either per-node ranging hardware [10, 12, 14, 16, 22] or careful system calibration and environment profiling [9, 11, 40], and thus is not appropriate for large-scale outdoor sensor networks. Range-free approaches localize nodes based on simple sensing, such as wireless connectivity [26, 27, 29, 32, 33], anchor proximity [25, 28, 30], or localization events detection [36, 37]. Among these, connectivity-based solutions feature a low overall system cost, however, by sacrificing localization accuracy.

Our work is motivated by the finding that localization by means of mere connectivity may *underutilize* the proximity information available from neighborhood sensing. Although radio signal strength (RSS) is not considered a good choice for physical distance estimation in many scenarios because of unknown radio path loss factors, multi-path effects, hardware discrepancies, antenna orientation, and so forth [40, 41, 42], it does provide some useful distance-related information beyond indicating connectivity among neighboring nodes. Our experimental study confirms that in outdoor open-air scenarios, the radio signal strength weakens approximately monotonically with the physical distance, especially from the viewpoint of a single node, where RSS might provide heuristic information about which neighboring node is closer and which is further.

Based on our empirical study, this paper introduces a novel range-free approach to extracting relative distance information from *neighborhood orderings* which serve as unique high-dimensional location signatures for sensor

nodes in the network. Instead of offering yet another new localization method, the design described in this paper can be conveniently applied as a transparent supporting layer for many state-of-the-art connectivity-based localization algorithms, providing a low-cost but effective solution for system accuracy.

We augmented three range-free localization algorithms, i.e., MDS-MAP [26], DV-Hop [27], RPA [33], with our design, and evaluated the effectiveness of the proposed design in two types of outdoor test-bed systems: an 850-foot-long linear network with 54 MICAz motes, and a regular 2-dimensional network covering an area of 10000 square feet with 49 motes. System evaluation showed noticeable performance gains including eliminating estimation ambiguity and reducing localization errors by as much as 35%. In addition, extensive simulation demonstrated the effectiveness of our design for large-scale networks and revealed an interesting feature of robustness to the unknown and spatially unevenly distributed radio path loss.

The rest of the paper is organized as follows: Section 2 surveys related work. Section 3 explains the motivation behind the paper with empirical data. The main design is introduced in Section 4. Section 5 briefs three range-free protocols on which we evaluated our work. Section 6 reports outdoor test-bed experiments. Section 7 discusses results from simulation. Finally, Section 8 concludes the paper.

2   Related Work

Based on whether ranging is conducted at the resource-constrained sensor nodes, most of the previous work about node localization can be categorized into one of following two classes: (i) range-based [9, 10, 11, 14, 12, 15, 16, 17, 18, 19, 20, 21, 22, 24], and (ii) range-free localization [25, 26, 27, 28, 30, 29, 31, 32, 34, 35, 36, 37, 38].

Range-based solutions try to estimate absolute distances or angles among randomly deployed sensor nodes and then apply triangulation or multilateration for location calculation. Many range-based methods use techniques such as Time of Arrival (ToA) [13], Time Difference of Arrival (TDoA) (e.g., Cricket [10], AHLos [11], TPS [12]) and Angle of Arrival (AOA) (e.g., APS [22], SpinLoc [18]) to measure distance or angles among nodes and anchors (also called beacons or reference nodes with pre-known location information). Those methods can be accurate but costly by adding per-node additional hardware [10, 12, 14], requiring intensive tuning [17] or not suitable for large-scale systems due to their limited effective range [10]. Although some research has tried to utilize RSS (Receive Signal Strength) with noise filtering for distance estimation or for wireless finger-print matching (e.g., Radar [9], wMDS [43], SpotOn [45], Indoor GPS [46], Sequence [48], Ranking [49]), empirical studies [39, 40, 41] have concluded that unless careful calibration and environment profiling can be accomplished, RSS is not a good choice for accurate ranging.

Range-free methods have applied many smart ideas for pursuing a low-cost design. Early range-free solutions made use of the proximity information to anchor nodes. Typical examples are Centroid [25], APIT [28], Concave [35] and Self [30], in which the high cost of anchors is supposed to

be amortized with a large number of low-cost ordinary senor nodes. To achieve a good accuracy, however, a high anchor density is required, which is impractical for large-scale systems. Concurrently, wireless connectivity-based protocols such as DV-Hop [27], MDS-MAP [26], RPA [33], Amorphous [32] and so on, proposed using local neighborhood sensing to build hop-based virtual distances for large-scale sensor network localization. In those systems, only a small number of anchors are necessary for constructing the global coordinates, which significantly reduces the system cost. Recent work helps solve problems of "holes" [31, 34] and "complex shapes" [29], contributing to connectivity-based solutions in practical irregular node deployment with obstacles. However, we found that localization by means of connectivity alone does not make full use of information available from local neighborhood sensing.

Realizing the limitations of previous works, this paper presents the idea of *regulated signature distance* (RSD), a metric of the proximity among 1-hop neighboring nodes. Acting as a transparent supporting layer, our design can effectively improve the system accuracy of state-of-the-art connectivity-based localization with little extra cost.

3   Empirical Data as Motivation

This paper is motivated by our experimental data showing that in the outdoor environments,

- Network-wide monotonic relationship between radio signal strength and physical distance does not hold, but

- Per-node monotonic RSS-Distance relationship holds well, i.e., any single node's RSS sensing results for its neighboring nodes can be used as an indicator for the relative "near-far" relationship among neighbors.

In the following, we first explain results from a preliminary test, and then provide data obtained from large-scale outdoor experiments for verification.

3.1   Preliminary Experiments

Figure 1 shows RSS sensing results from MICAz nodes in several outdoor experiments conducted in two types of environment: grass land and parking lot. In the test, we placed 9 sender nodes at different distances from a receiver node. Each sender node broadcast 100 packets with 0dBm sending power, and the receiver node recorded the RSS upon receiving the packet. In the grass-land scenario, we performed the test twice with two different receiver nodes placed at the same location and without moving or switching sender nodes (Grass Land Test1 and Grass Land Test 2, respectively). In
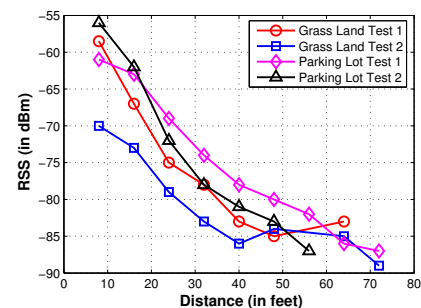

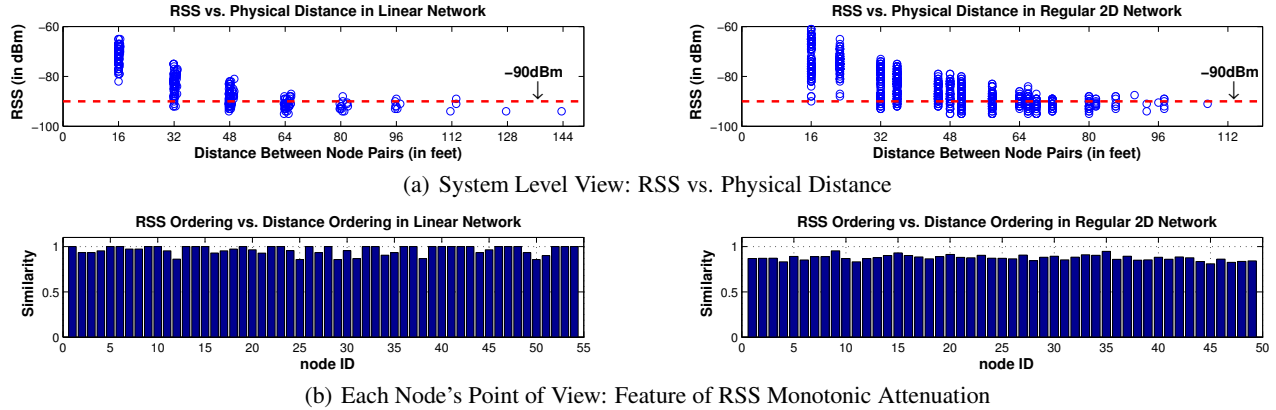**Figure 1. Experimental Results: RSS vs. Distance**

(a) System Level View: RSS vs. Physical Distance

(b) Each Node's Point of View: Feature of RSS Monotonic Attenuation

Figure 2. Empirical Date from Large Scale Experiments

the parking lot scenario, identical sets of nodes were tested during day-time (Parking Lot Test 1) and at night (Parking Lot Test 2). Tests were conducted multiple times, and results did not show significant changes in the overall shapes of the curves shown in Figure 1.

Figure 1 tells that at the system level, using absolute values of RSS for distance estimation is not reasonable since identical RSS values may correspond to different distances. However, for each individual curve (i.e., from the viewpoint of a single node), RSS values mostly decreased monotonically with increasing distance, conveying information about relative "near-far" relationships among 1-hop neighbors.

3.2 Large-scale Experiments

We then conducted large-scale outdoor experiments with two types of networks to verify the phenomena found in the preliminary test. The first experiment was a linear network containing 54 MICAz nodes with a 16-foot intermediate distance between adjacent nodes covering a 850-foot length along a road. In the second experiment, we constructed a regular 2D network with a 7×7 grid-shaped layout including 49 nodes occupying an open-air parking lot area of 10000 square feet. The setup of the large-scale experiments will be further detailed in Section 6.

Figure 2 reports the empirical data obtained from the two test-beds. Figure 2(a) plots the sensed RSS value for each pair of nodes against the distance between them, which verifies that monotonic RSS-distance relationship does not hold for the whole network. In both the linear network and the regular 2D network, on one hand, RSS may vary dramatically for identical distance. For example, as shown in the right sub-figure of Figure 2(a), RSS ranges from -60dBm to -90dBm for a 16-foot distance in the 2D network. On the other hand, a single RSS value may correspond to a wide range of distances. For instance, as shown in the left sub-figure, -90dBm could range from 32 feet to 112 feet in the linear network; even worse, -90dBm RSS covers almost all of the distance spectrum, i.e., from 16 feet to 112 feet, in the 2D network showing in the right sub-figure of Figure 2(a).

However, examining the data from the viewpoint of a single node tells a different story. For any node, say u_i, we can obtain an ordered node list, say A, by listing u_i's 1-hop neighbors according to their RSS values sensed at u_i in decreasing order; and another node list, say B, by ordering u_i's

1-hop neighbors with increasing physical distance. Ideally, if the sensed RSS decreases monotonically with increasing distance, A and B should be identical. We define the *similarity* between two lists A and B as the percentage of *accordant node pairs* between them. For example, let $A = (u_1, u_2, u_3)$ and $B = (u_1, u_3, u_2)$, then $\{u_1, u_2\}$ is an accordant node pair between A and B since node u_1 is ordered ahead of u_2 in both A and B; while $\{u_2, u_3\}$ is not since their ordering gets reversed from A to B. We can see that if A and B are consistent with their similarity close to 1, the monotonic feature holds.

Figure 2(b) shows the similarity results for all nodes in two test-beds. We can see from the left sub-figure that in the linear network, most of the nodes have a similarity close to 1 (the minimum, mean and maximum similarities are 0.86, 0.96 and 1, respectively). It means that in the linear network, from single node's point of view, RSS values for 1-hop neighbors are approximately monotonic with the distance. This finding still holds for the 2D regular network as shown in the right of Figure 2(b), where the minimum, mean and maximum similarities are 0.81, 0.88 and 0.96, respectively.

Above experiments confirm that the monotonic RSS-distance relationship does not hold at the system level, but approximately holds from the viewpoint of a single node.

3.3 Analysis and Discussion

In addition to the physical distance between two nodes, there are many factors that affect RSS sensing results. Table 1 lists some major aspects. We marked an aspect with a "$\sqrt{}$" if pre-deployment engineering efforts could possibly be applied to reduce its impact, or a "$\times$" if it would be hard or costly to address.

At the sender side, besides the sending power, the carrier frequency, modulation, baud rate and etc. determine the band-width, center frequency and spectrum shape [47], which all affect the RSS at the receiver side. Most of those parameters can be configured with small offset errors and maintained relatively stable during the runtime. Antenna is-

Table 1. Major Factors Affecting RSS Sensing

Types of Factors	P
RF Transmit Parameters: Sending Power, Carrier Frequency, Modulation …	$\sqrt{}$
Antenna Issues: TX/RX Gain, Radiation Pattern, Orientation, Height …	$\sqrt{}$
Random Noise: Interference, Mobile Effects, Electronic Pulse …	$\sqrt{}$
Propagation Path Loss: Terrain, Vegetation, Obstacle, Magnetic Field …	$\times$
Node-level Sensing Discrepancy: LNA, ADC Ref. Voltage, Ground Noise …	$\times$

sues such as isotropic gain, orientation and etc. can also be carefully engineered in the design phase. For transient random noise, traditional filtering methods are able to help reduce its impact. All of the above are considered addressable without significant in-field calibration.

On the contrary, unpredictable environmental factors are much harder to handle. For example, radio path loss is unknown and costly to profile in most cases since it is temporally dynamic and spatially unevenly distributed. Another difficult issue is the sensing hardware discrepancy among different nodes. For example, a tiny bias of the ADC reference voltage or small variance of LNA (low noise amplifier) gain caused by different ground noise levels, may lead to different RSS values at two nodes, even when their received signal strengths are equivalent. Runtime sensing discrepancy among nodes is caused by many dynamic reasons and per-node in-field calibration could be very costly.

The above two "×" factors are hard to address in a deployed system, however, from the viewpoint of a single node, their impacts could be less severe. Firstly, a 1-hop neighborhood area is much smaller than the whole region covered by the network, so one node's local sensing alleviates the problem of spatially unevenly distributed path loss. In addition, RSS from a single node's sensing avoids the issue of node-level receiver side hardware discrepancy. Therefore, as confirmed by our empirical data, the monotonic RSS-distance relationship holds much better in the case of one node.

Unfortunately, this heuristic correlation is not utilized by previous localization methods based on mere connectivity, where only a binary "1" or "0" is evaluated for either connected or not, resulting in a degraded system accuracy.

4 Design: a Relative Distance

This section presents the main design of a range-free *relative distance* among 1-hop neighboring nodes.

4.1 Neighborhood Ordering as a Signature

Given the RSS sensing results for neighboring nodes, a node can obtain a *neighborhood ordering* with two steps:

- Sorting its 1-hop neighbors according to their signal strength by decreasing order, and

- Adding itself as the first element in the sorted node list.

A simple example is shown in Figure 3. In this figure, graph G on the left illustrates the connectivity of the network. On the right, each node generates a node list starting with itself and containing all its 1-hop neighbors which are ordered by decreasing signal strength, i.e., by increasing distance in an ideal case.
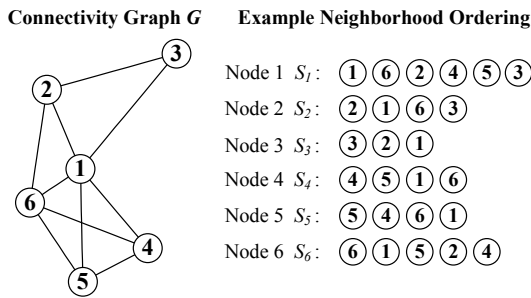

Figure 3. Neighborhood Ordering

For any node u_i, we consider its neighborhood ordering S_i as a *high-dimensional signature* of the node in the network. S_i has a vector format and contains all 1-hop neighbors of node u_i with three important features:

- S_i is unique for each node u_i.

- S_i is position-dependent. S_i embeds location-related information on both connectivity and proximity.

- S_i is obtained without ranging; instead, it is a range-free sensing observation beyond connectivity.

For the sake of clarity, in the following design part, we first use ideal neighborhood orderings for conveying ideas. Namely, S_i is consistent with the ordering according to physical distance. Later sections will verify the effectiveness and robustness of our design in practical noisy scenarios through both test-bed and simulation experimentation.

4.2 SD: Signature Distance

The high-dimensional signatures of sensor nodes can be obtained easily via local signal strength sensing. In this section, we explain the concept and rationale of *signature distance* (SD) which quantifies the difference between two high-dimensional signatures. SD is the first step toward a *relative distance* that effectively reflects the physical distance relationships among neighboring nodes in the network.

4.2.1 Formation, Definition and Calculation of SD

Say that a pair of nodes u_m and u_n get flipped between two signatures S_i and S_j, if the ordering of u_m and u_n in S_i gets reversed in S_j. For example, as shown in Figure 4, the ordered node pair $\{1,6\}$ in $S_2 = (2,1,6,3)$ gets reversed to $\{6,1\}$ in $S_5 = (5,4,6,1)$.
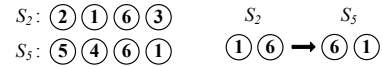

Figure 4. 1 Explicit Node-Pair Flip

There are three types of potential node-pair flips between two signatures S_i and S_j : (i) explicit flip, (ii) implicit flip, and (iii) possible flip. If node u_m and u_n appear in both S_i and S_j, then we can easily tell whether this node pair gets flipped or not, as the example shows for node pair $\{1,6\}$ in Figure 4. This type of flip is called explicit flip. Implicit flips and possible flips are somewhat tricky, as explained in the following with examples.

As shown in the left part of Figure 5, S_2 and S_5 have different sets of node elements. For example, $S_2 = (2,1,6,3)$ contains node 2 while $S_5 = (5,4,6,1)$ does not. In this case, many node pairs in S_2 do not have corresponding counterparts in S_5. For instance, $\{2,1\},\{2,6\}$ in S_2 have no related node pairs in S_5 since node 2 is absent in S_5. We solve this problem by attaching "wildcards" to S_2 and S_5, as depicted
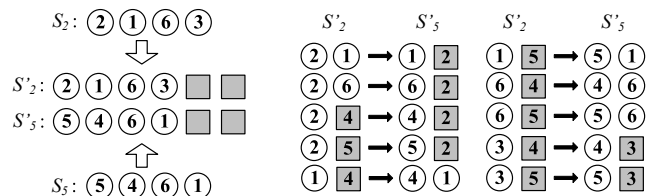

Figure 5. 10 Implicit Node-Pair Flips

by gray squares □ in Figure 5. Formally, for S_i and S_j, a number of $|S_i \cup S_j - S_i|$ wildcards are attached to S_i to make S_i'. In S_i', each wildcard can stand for any node $u \in S_j$ but $\notin S_i$, namely $\forall u \in (S_i \cup S_j - S_i)$. For example, in Figure 5, each gray square in S_2' can stand for either 5 or 4, and a gray square in S_5' can be substituted with either 2 or 3. Since "wildcard nodes" attached to S_i are naturally regarded as further away than neighbors of u_i in S_i, S_i' maintains the features as a location-dependent signature without violating proximity relationships embedded in the original S_i. Figure 5 lists all implicit node pair flips from S_2' to S_5'. We call them *implicit flips* since they are not as obvious as the explicit flips. For example, node pair $\{2,1\}$ in S_2' can only have a counterpart node pair $\{1,\square\}$ in S_5', where □ stands for 2 in this case. $\{2,1\}$ and $\{1,\square\}$ is a flip from S_2' to S_5' because an order reversion occurs no matter which □ in S_5' stands for 2.
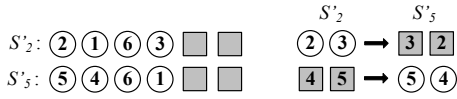

Figure 6. 2 Possible Node-Pair Flips

Figure 6 gives examples of the *possible node-pair flip*. Formally, if a node pair $\{u_m, u_n\}$ appears in S_i but neither u_m nor u_n is in S_j, we consider it *possible* that $\{u_m, u_n\}$ gets reversed in S_j. For example, as shown in Figure 6, $\{2,3\}$ from S_2' can only have a counterpart $\{\square, \square\}$ in S_5'. With no additional information, this node pair gives a possible node-pair flip with 50% probability.

Based on the above explanations, we now define the signature distance between S_i and S_j as follows:

Definition 1: the signature distance $SD(S_i, S_j)$ is equal to the summation of the number of explicit flips $F_e(S_i, S_j)$, implicit flips $F_i(S_i, S_j)$, and possible flips $F_p(S_i, S_j)$ times 0.5 (50% probability of flip for possible node pairs), namely,

$$SD(S_i, S_j) = F_e(S_i, S_j) + F_i(S_i, S_j) + F_p(S_i, S_j) \times 0.5 \quad (1)$$

Taking S_2 and S_5 in Figure 4 as an example, $F_e(S_2, S_5) = 1$ as shown in Figure 4, $F_i(S_2, S_5) = 10$ as listed in Figure 5, and $F_p(S_2, S_5) = 2$ as depicted in Figure 6. According to definition 1, we have $SD(S_2, S_5) = 1 + 10 + 2 \times 0.5 = 12$.

In fact, each node-pair flip from $S(i)$ to $S(j)$ corresponds to one and only one node-pair flip from $S(j)$ to $S(i)$, so the definition of SD guarantees $SD(S_i, S_j) \equiv SD(S_j, S_i)$.

Algorithm 1 illustrates a method for computing the signature distance. First of all, S_i and S_j get sorted at line 1 and 2 with complexity $O(Klog(K))$, where $K = |S_i \cup S_j|$ is the total number of neighbors of two nodes u_i and u_j. The function $wildCard()$ at Line 3 attaches $(K - |S_i|)$ wildcards to S_i and fills these wildcards with elements $(S_i \cup S_j - S_i)$ that are ordered the same as they are in S_j. Line 4 performs similarly to S_j to obtain S_j'. Line 3 and 4 have a cost of $O(K)$ with sorted input $\widetilde{S_i}$ and $\widetilde{S_j}$. Line 5 computes the total number of explicit and implicit node-pair flips using a variant of heap-sort algorithm [51] with complexity $O(Klog(K))$. Line 6 calculates the number of possible flips. Line 7 gives the result of $SD(S_i, S_j)$ based on Equation 1. The time complexity of Algorithm 1 is $O(Klog(K))$. Normally $K \ll n$, where n is the total number of nodes in the network.

Algorithm 1 Signature Distance

Input: S_i and S_j
Output: $SD(S_i, S_j)$

1:	$\widetilde{S_i} = sort(S_i)$;	% $O(Klog(K))$
2:	$\widetilde{S_j} = sort(S_j)$;	% $O(Klog(K))$
3:	$S_i' = wildCard(S_i, \widetilde{S_i}, \widetilde{S_j})$;	% $O(K)$
4:	$S_j' = wildCard(S_j, \widetilde{S_j}, \widetilde{S_i})$;	% $O(K)$
5:	$F_{e+i} = HeapSort(S_i', S_j')$;	% $O(Klog(K))$
6:	$F_p = \frac{(K-|S_i|)(K-|S_i|-1)}{2} + \frac{(K-|S_j|)(K-|S_j|-1)}{2}$;	% $O(1)$
7:	$SD(S_i, S_j) = F_{e+i} + F_p \times 0.5$;	% $O(1)$

4.2.2 Insights into the Signature Distance

In a signature S_i, each ordered node pair contains a proximity relationship. For example, as shown in Figure 7(a), $S_2 = (2, 1, 6, 3)$, the ordered node pair $\{1,3\}$ in S_2 means that from node 2's point of view, node 1 is closer than node 3. In other words, if we divide the plane with $B(1,3)$ (depicted as a dashed line), which is the perpendicular bisector of the line segment $L(1,3)$ connecting node 1 and 3, the ordering of $\{1,3\}$ in S_2 indicates that node 2 is located on the left side of $B(1,3)$.

Based on the above example, we can see that a node-pair flip between two signatures corresponds to *passing a bisector line*. For example, as shown in Figure 7(a), S_2 contains node pair $\{1,3\}$ which gets reversed to $\{3,1\}$ in S_3, meaning that node 2 and node 3 are on different sides of $B(1,3)$. So going from node 2 to node 3 along the straight line segment $L(2,3)$ needs to pass $B(1,3)$, as shown in the figure. Figure 7(b) illustrates an opposite case, in which signature S_2 and S_3 have an accordant node pair $\{2,1\}$, indicating that node 2 and 3 are located at the same side of $B(1,2)$ and $L(2,3)$ does not intersect $B(1,2)$. Figure 7(c) shows an example of the implicit flip. In S_2', node pair $\{6,3\}$ has a counterpart $\{3,\square\}$ in S_3', where □ is a wildcard standing for node 6 in this case. This implicit flip corresponds to an intersection of $L(2,3)$ and $B(3,6)$, as shown in the figure.
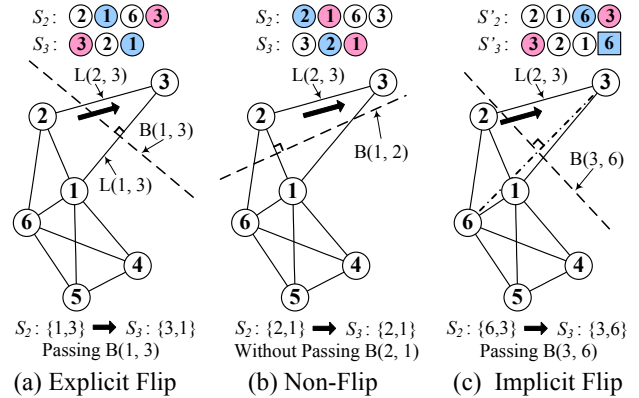

Figure 7. The Insight of a Node-Pair Flip

In general, we have the following observation:

Observation 1: a node-pair flip $\{u_m, u_n\} \Rightarrow \{u_n, u_m\}$ from S_i to S_j indicates that the line segment $L(u_i, u_j)$ passes the perpendicular bisector line $B(u_m, u_n)$.

One remark about the above observation is that there is only one intersection between $L(u_i, u_j)$ and $B(u_m, u_n)$ when node-pair flip happens. This is because two straight lines (segments) have at most one crossing point.

On the other hand, based on the definition of signature distance, $SD(S_i, S_j)$ evaluates the difference between two signatures S_i and S_j by counting the total number of node-pair flips. Thus, we can conclude from Observation 1 that

Observation 2: $SD(S_i, S_j)$ is equivalent to the number of bisector lines we need to pass if going from neighboring node u_i to u_j along the line segment $L(u_i, u_j)$.

Another key observation concerning the physical distance between two neighboring nodes is that

Observation 3: under roughly uniform bisector line density, for neighboring nodes u_i and u_j, the number of bisector lines passed by line segment $L(u_i, u_j)$ is approximately proportional to the physical distance between u_i and u_j, i.e., the length of $L(u_i, u_j)$, denoted as $PD(u_i, u_j)$.

The insight offered by Observation 3 is that longer physical distances provide a higher probability of passing more bisector lines. Here bisector line density is defined as the number of lines exist between two positions with unit distance. Figure 8 shows an example for this observation. Figure 8(a) draws the layout of perpendicular bisector lines (denoted as dashed lines) for all node pairs, and line segments connecting node 1 with its 1-hop neighbors (denoted as solid lines). Figure 8(b) lists line segments, corresponding number of bisector lines they passing, and signature distances between their terminal nodes, respectively. We can see that the number of bisector lines passed by a line segment is approximately proportional to the length of the line segment, i.e., the physical distance between two nodes.
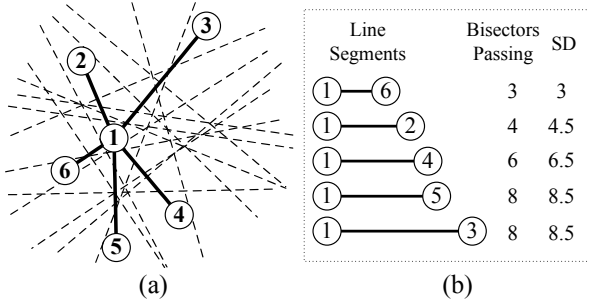

Figure 8. Physical Distance vs. Bisector Lines Passing

Combining the above three observations, we have a heuristic rule as follows: for two neighboring nodes u_i and u_j, their signature distance is approximately proportional to the physical distance between them, namely

$$SD(S_i, S_j) \propto PD(u_i, u_j) \qquad (2)$$

An important remark is that Equation 2 is not valid for non-neighboring node pairs most of the time. This is because $SD(u_i, u_j)$ *only* counts the number of passed bisector lines generated by node pairs from set $S_i \cup S_j$. Figure 9 shows an example for explaining this remark. In this figure, u_i, u_j and u_k are located far from each other, and their neighboring areas (denoted with dashed circles) do not overlap. As a result, $SD(S_i, S_j)$ and $SD(S_i, S_k)$ both are determined only
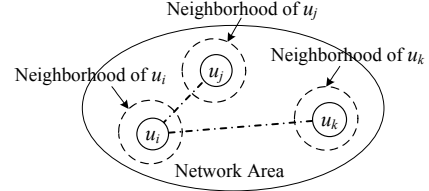

Figure 9. Far-away Node Pairs

by possible flips that depend on the number of nodes in S_i, S_j and S_k. Under similar 1-hop radio range, $SD(S_i, S_j)$ and $SD(S_i, S_k)$ could have similar values, although the physical distances of these two node pairs may be dramatically different. In a word, the heuristic relationship of Equation 2 is meaningful only for 1-hop neighboring nodes.

Based on Equation 2, signature distance can be utilized as a *relative distance* for localization purposes because it approximately reflects the "near-far" relationships among 1-hop neighbors. However, in some cases, SD can be biased due to spatially non-uniform bisector line density throughout the network area, a violation of the condition in Observation 3. In the next section, we propose a more robust relative distance, i.e., Regulated SD, to address this problem.

4.3 RSD: Regulated Signature Distance

This section introduces the *regulated signature distance*, or RSD for short, as a refined version of SD.

4.3.1 The Rationale behind SD Refinement

Spatially non-uniform bisector line density could affect the effectiveness of SD as a relative distance. This problem comes from two aspects: (i) local node placement; and (ii) network wide neighborhood size, both of which are explained in the following with examples.

As shown in Figure 10(a), $L(2, 3)$ passes 4 bisector lines and $L(6, 1)$ intersects 3. However, $L(2, 3)$ is more than twice as long as $L(1, 6)$. This inconsistency is caused by spatially unbalanced bisector line density within the local area. For example, the area close to node 1 has a higher bisector line density, while boundary areas close to node 2 and 3 have a low density. This micro-level observation indicates that SD needs to be refined considering the local bisector line density.

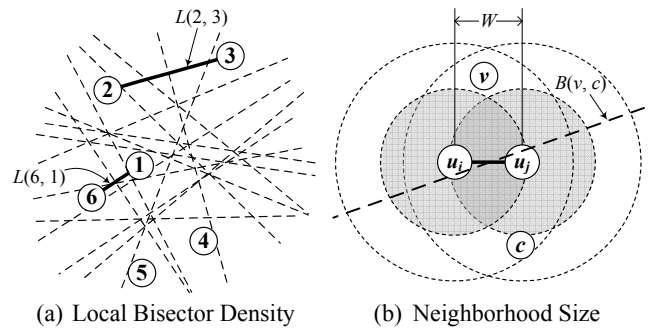

(a) Local Bisector Density   (b) Neighborhood Size

Figure 10. Motivation for SD Regulation

At the macro level, for the same physical distance, the value of signature distance could be different, depending on the neighborhood size. For example, as shown in Figure 10(b), two nodes u_i and u_j has a constant physical distance W. When they have a large 1-hop radio range illustrated by two big circles, both of them have node v and

c as neighbors included in their signatures. In this case, $SD(u_i,u_j)$ counts the passing of $B(v,c)$ which is the bisector for node pair $\{v,c\}$. However, if u_i and u_j have a smaller 1-hop radio range (e.g., due to strong ambient noise) denoted by gray-filled circles. Node v, u are absent from signatures of u_i and u_j, so $SD(u_i,u_j)$ does *not* count $B(v,c)$, resulting a smaller value of $SD(u_i,u_j)$ comparing to the previous case. In fact, under random node deployment with uniform density, the signature distance for node u_i and u_j should be regarded as a relative value in terms of the dimension of their neighborhood area that affects the number of available bisector lines counted for passing in the map.

Based on the above analysis, we propose the *Regulated Signature Distance*, or RSD for short, defined as follows:

$$RSD(u_i,u_j) = SD(S_i,S_j) \cdot \frac{\sqrt{K}}{K(K-1)/2} \quad (3)$$

Equation 3 refines $SD(S_i,S_j)$ with a factor $\frac{\sqrt{K}}{K(K-1)/2}$, where $K = |S_i \cup S_j|$ is the total number of nodes in the neighborhood of node u_i and u_j combined. In this equation, $K(K-1)/2$ calculates the number of local bisector lines, used to normalize $SD(S_i,S_j)$ with the local bisector density; $\sqrt{K}$ estimates the diameter of this neighborhood, which puts the factor of neighborhood size into consideration. We formally derive and explain this equation in the next section.

4.3.2 Regulation Factor Derivation

For neighboring nodes u_i and u_j, let $|S_i \cup S_j| = K$. There are a total of $n_B = K(K-1)/2$ bisector lines generated by node pairs in $S_i \cup S_j$. According to the Pie-Cutting Theorem [50], n_B bisector lines divide the local area into n_R small regions, where

$$n_R = O((n_B^2 + n_B + 2)/2) = O(n_B^2) \quad (4)$$

Let S be the size of the local area occupied by the neighborhoods of u_i and u_j. The expected size and diameter of each small region, denoted as $E[s_R]$ and $E[d_R]$ respectively, are

$$E[s_R] = \frac{S}{n_R} = \frac{S}{O(n_B^2)}, \quad E[d_R] = \alpha \cdot \sqrt{E[s_R]} = \frac{\alpha\sqrt{S}}{O(n_B)} \quad (5)$$

where α is a constant factor determined by the shape modeling of the small region.

Passing $NB(u_i,u_j)$ Bisectors

Residual

$L(u_i,u_j)$

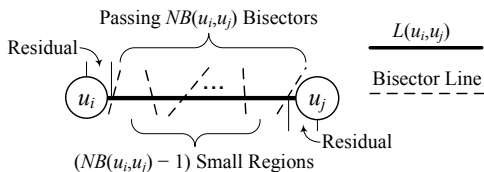Bisector Line

$(NB(u_i,u_j) - 1)$ Small Regions

Residual

Figure 11. Bisector Lines and Small Regions

Suppose that line segment $L(u_i,u_j)$ intersects $NB(u_i,u_j)$ bisector lines, then $L(u_i,u_j)$ passes $(NB(u_i,u_j) - 1)$ small regions as shown in Figure 11. Adding residuals at both ends (each counted as a half region), we get an approximation

$$PD(u_i,u_j) \approx NB(u_i,u_j) \cdot E[d_R] \quad (6)$$

meaning that the distance between u_i and u_j approximately equals the number of small regions times expected diameter.

$SD(u_i,u_j)$ estimates the number of bisector lines that $L(u_i,u_j)$ passes, i.e., $SD(u_i,u_j) \approx NB(u_i,u_j)$. So we have

$$PD(u_i,u_j) \approx SD(S_i,S_j) \cdot E[d_R] = SD(S_i,S_j) \cdot \frac{\alpha\sqrt{S}}{O(n_B)} \quad (7)$$

For uniform random node deployment, the expected number of nodes in an area is proportional to the area size, namely $E[K] = \phi \cdot S$ where ϕ is the node density. Since $n_B = K(K-1)/2$, we can rewrite Equation 7 as

$$PD(u_i,u_j) \approx SD(S_i,S_j) \cdot \frac{\phi\sqrt{K}}{K(K-1)/2}, \text{ where } \phi = \frac{\alpha}{\sqrt{\phi}} \quad (8)$$

ϕ is a constant scaling factor that can be eliminated without violating near-far relationship among different neighboring nodes. Thus, we obtain the proposed *relative distance* RSD:

$$RSD(u_i,u_j) = SD(S_i,S_j) \cdot \frac{\sqrt{K}}{K(K-1)/2} \quad (9)$$

4.3.3 RSD vs. SD

Figure 12 compares SD and RSD serving as a relative distance. This figure is obtained from the simulation of a network composed of 150 randomly deployed nodes covering a 500×500 square feet area with 100-foot radio range. For each node pair in the network, both SD and RSD are computed, and plotted against their physical distance in Figure 12(a) and Figure 12(b) respectively. Figure 12 conveys two major points: (i) within 1-hop radio range, RSD offers a much better linear correlation with physical distance than SD; (ii) both figures confirms our earlier remark about Equation 2 that after a 1-hop radio range, signature distance is no longer correlated with physical distance. From this example, we can see that RSD is a better choice than SD to serve as a metric of *relative distance*.
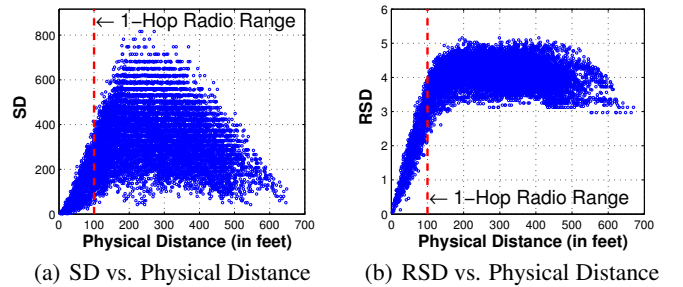

(a) SD vs. Physical Distance   (b) RSD vs. Physical Distance

Figure 12. Correlation with Physical Distance

5 Design as a Supporting Layer

The design of RSD can be implemented as a supporting layer that is transparent to the localization algorithms. As shown in Figure 13, we simply use the smallest accumulated RSD along a path between two nodes instead of the shortest-path hop count as the estimated relative distance. Specially, the accumulated RSD between two nodes is defined as

- For 1-hop neighboring nodes u_i and u_j, accumulated RSD equals $RSD(u_i,u_j)$ computed with Equation 9;
- For non-neighboring nodes u_i and u_j, accumulated RSD is calculated as the summation of the RSD values of neighboring nodes along a path between u_i and u_j.
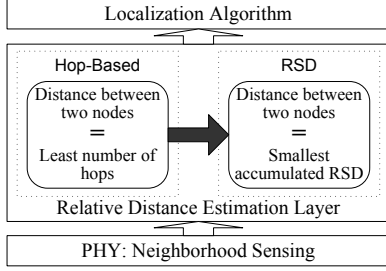
Figure 13. RSD Design Embedding

In the following, we briefly describe three connectivity-based localization schemes studied in our evaluation.

5.1 Connectivity-Based Schemes

The key idea in connectivity-based localization is to use the number of communication hops between two nodes to evaluate the physical distance between them. The following three schemes are typical examples:

- MDS-MAP [26] by Y. Shang, W. Ruml, et al.

- DV-Hop [27] by D. Niculescu and B. Nath.

- RPA [33] by C. Savarese, J. M. Rabaey, et al.

Here we provide brief descriptions of these algorithms and details can be found in [26, 27, 33].

MDS-MAP [26] first forms a distance matrix A in which the value at the ith row and the jth column is the shortest hop-based distance between node u_i and u_j. The algorithm computes a "relative map" of the network with the MDS (multi-dimensional scaling) [52] technique given distance matrix A as input. After that, an "absolute map" can be obtained by scaling and rotating the "relative map" according to the absolute coordinates built with at least three anchor nodes.

DV-Hop [27] uses a mechanism that is similar to classical distance vector routing. After beacon flooding from more than three anchors in the network. The algorithm estimates the expected physical distance for 1-hop with

$$HopSize = \frac{\sum_{i \neq j} Distance(v_i, v_j)}{\sum_{i \neq j} Hops(v_i, v_j)} \qquad (10)$$

where v_i, v_j are anchors. $Distance(v_i, v_j)$ and $Hops(v_i, v_j)$ are the physical distance and the least number of hops between v_i and v_j, respectively. For each sensor node u_i, it estimates its distance to anchor v_i with $HopSize \times Hops(u_i, v_i)$. Finally, each node's location is computed with least-square multilateration on available anchors.

RPA [33], which is proposed independently from DV-Hop, uses a similar mechanism of hop-based distance estimation called Hop-TERRAIN for its first step. Besides, it introduces an iterative refinement step for position adjustment based on local sensing results. Basically, at iteration k, the position of node u_i is recomputed based on the estimated positions of its neighbors obtained from iteration $k-1$. More sophisticated approaches such as confidence based filtering are also proposed in RPA.

5.2 Design Embedding

With local RSS sensing results, RSD can be calculated at each node or in a localization server. For range-free connectivity based localization algorithms such as MDS-MAP,

DV-Hop, and RPS, applying RSD is convenient and incurs little overhead. Without modifying the major design of the original algorithm, the RSD value can be used instead of "1" (indicating connection) for neighboring nodes. Specifically, in MDS-MAP, for example, everything remains the same except that the distance matrix now holds the smallest accumulated RSD values along a path instead of the shortest hop distance. For DV-Hop, similarly, the relative distance turns to the smallest accumulated RSD instead of shortest-path hops. Expected physical distance for 1 *unit* of RSD is given by

$$RSD_{unit}Size = \frac{\sum_{i \neq j} Distance(v_i, v_j)}{\sum_{i \neq j} RSD(v_i, v_j)} \qquad (11)$$

For RPA, besides embedding RSD in the Hop-TERRAIN step, the refinement step also benefits from applying RSD for local iterative position adjustment.

5.3 Complexity of RSD Embedding

Involving RSD in localization introduces little additional cost. Algorithm 1 costs $O(Klog(K))$ for SD calculation, where K is the maximum length of a signature sequence in the network, and normally $K \ll n$ where n is the total number of nodes. Even in a centralized localization system, additional overhead of RSD calculation is $O(nK^2log(K))$ (n nodes and each has at most $K-1$ neighbors), given the complexity of the other computational components. For instance, MDS has a complexity of $O(n^3)$ for the step of matrix decomposition alone [44]. About communication, the only additional overhead is for signature exchange among neighboring nodes. Signatures are very short and can be piggybacked economically on messages during the network initialization phase. Importantly, signature exchange only occurs within 1-hop neighborhood without multi-hop flooding. Therefore, embedding RSD in a connectivity-based localization does not affect the scalability of the system.

6 Test-bed Experimentation

In this section, we report two types of outdoor experiments with 54 and 49 MICAz motes respectively.

6.1 Experiment I: Linear Network

We start our test-bed evaluation from a linear network widely applied in transportation-related road networks.

6.1.1 Experiment Setup

As shown in Figure 14, 54 MICAz motes were deployed on the grass covered curb along a road (with surrounding obstacles including trees, metal fence and parked vehicles). Each node was left 8 inches above the ground and the distance between two immediate nodes was about 16 feet. Every node broadcasts 100 packets with carrier-sensing and back-off time intervals set from 200ms (millisecond) to 1690ms individually for collision avoidance. The radio sending power was 0dBm at channel 26 (F_c = 2480MHz) to avoid possible WiFi interference from the surroundings. Each packet contained the sender's node ID and a sequence number of the packet. When a node received a packet, it logged the sensed signal strength, sender's ID, and the sequence number of the packet into its flash memory.

Table 2 lists some of the collected information about the linear network. The whole length of the network was about 848 feet with 15 hops between two terminal nodes (node#1

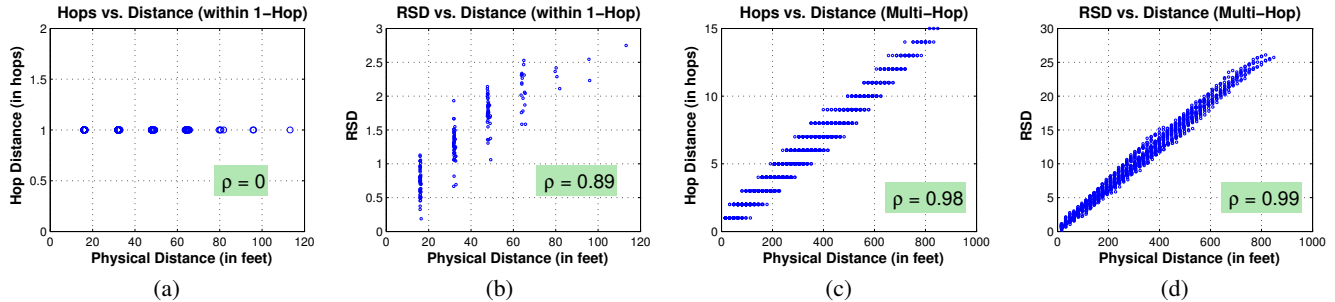**Figure 14. Test-bed Experiments I: Linear Network**

Figure 15. Distance Correlation Comparison: RSD vs. Hop (Linear Network)

and node#54). The maximum 1-hop radio communication range in this experiment was about 144 feet and the average 1-hop neighborhood size was 6.7 (nodes).

Table 2. Statistics of the Linear Network

Scale	Max. Hops	Max. 1-Hop RF Range	Neighborhood Size
≈ 848(feet)	15	≈144(feet)	6.7 (nodes)

6.1.2 Distance Correlations

For either the traditional hop-based distance or the newly proposed RSD, their effectiveness as a relative distance in localization is determined by the correlation between their values and the physical distance.

Figure 15 illustrates the correlations between hop-based distance, RSD, and the physical distance. Figure 15(a) and 15(b) plot all the node pairs within 1-hop communication range in the linear network. In both figures, the X-axis is the physical distance between two nodes, and the Y-axis is the hop-based distance and RSD, respectively. Figure 15(a) reveals that all 1-hop node pairs have an identical hop distance of 1, thus the correlation coefficient is $\rho = 0$ within 1-hop range. On the contrary, as shown in Figure 15(b), most of the node pairs with different distances can be differentiated from each other according to their RSD value. The empirical data shown in Figure 15(b) have a correlation coefficient $\rho = 0.89$. Comparing Figure 15(a) and 15(b), we can conclude that RSD owns a *sub-hop resolution* that is not available from the traditional hop-based distances.

Figure 15(c) plots all the node pairs (both 1-hop and multi-hop node pairs) according to their hop distance and physical distance. Similarly, Figure 15(d) plots RSD (accumulated RSD for multi-hop nodes) against physical distance for all node pairs. Comparing these two figures, although the correlation coefficients are close ($\rho = 0.98$ for hop distance and $\rho = 0.99$ for RSD), RSD provides better resolution. In Figure 15(c), a physical distance can only be mapped

to an integer hop distance in a discrete manner, while in Figure 15(d), the mapping is continuous.

Localization results in next section show that RSD's sub-hop resolution can nicely solve the *ambiguity problem*, i.e., mapping closely located nodes to identical positions.

6.1.3 Localization Performance

We use the terminology "MDS-Hop", "DV-Hop" and "RPA-Hop" for the original hop-based approaches and "MDS-RSD", "DV-RSD" and "RPA-RSD" for corresponding methods embedded with RSD.

Let us first look at DV-Hop and DV-RSD. Figure 16(a) shows the localization results from both algorithms with two terminal nodes of the linear network as anchors. In the figure, black line segments are plotted starting from nodes' deployed positions (depicted as blue dots) and pointing to corresponding estimated locations, for clear observation. In the left subfigure for DV-Hop, many nodes are mapped to identical estimated locations, i.e., the ambiguity problem; while DV-RSD does not encounter this issue as shown in the right subfigure. This result confirms that RSD offers a unique sub-hop resolution while hop-based distance does not.

Figure 16(b) illustrates the results of RPA-Hop and RPA-RSD, both with two iterative refinement steps. Both RPA-Hop and RPA-RSD achieved unique position estimation for each node. For RPA-Hop, this is credited to the refinement step that smooths the estimated position of each node over its neighbors, naturally solving the problem of clustered mapping. From Figure 16(b), we can observe that RPA-RSD achieves better localization accuracy than RPA-Hop.

Figure 16(c) shows the localization results from MDS-Hop and MDS-RSD. We found that 2-dimensional MDS is not stable for linear or close-to-linear networks because of singularity issues in matrix decomposition. Here, 1-dimensional MDS is applied first, and then two terminal nodes are used as anchor nodes for map scaling and rota-
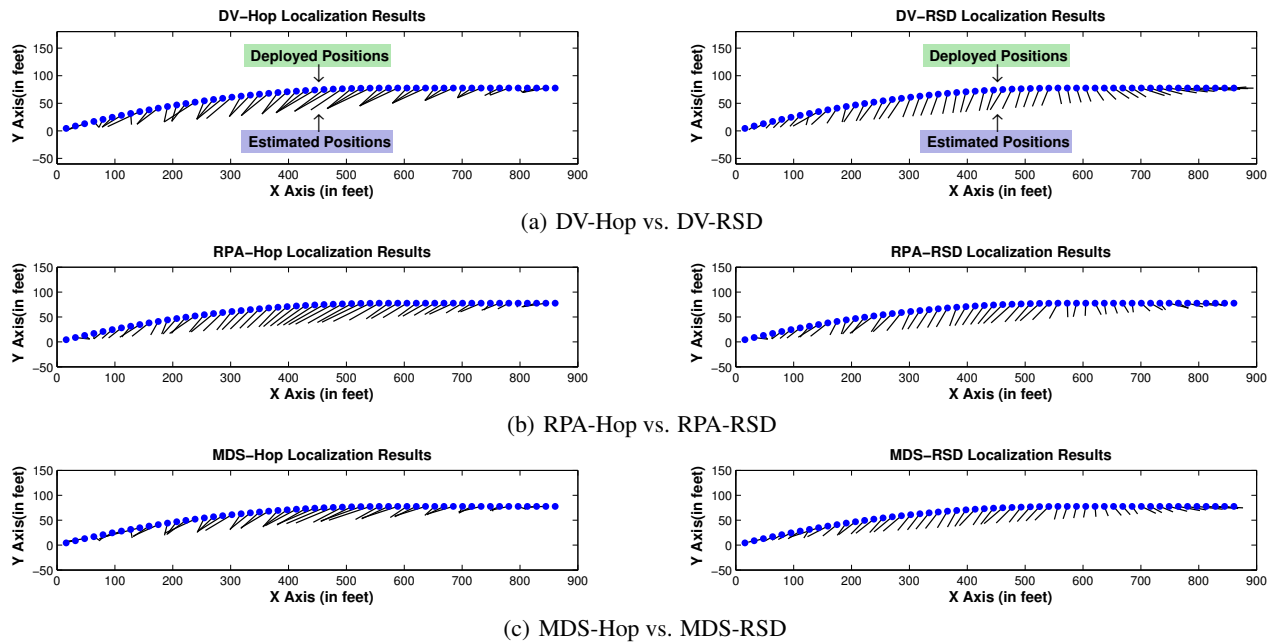
(a) DV-Hop vs. DV-RSD

(b) RPA-Hop vs. RPA-RSD

(c) MDS-Hop vs. MDS-RSD

Figure 16. Localization in Linear Networks: Hop-Based Distance vs. RSD

tion to obtain the 2-dimensional absolute map. Figure 16(c) tells that MDS algorithm itself is not able to solve the ambiguity problem; while MDS-RSD shows better performance without clustered mapping.
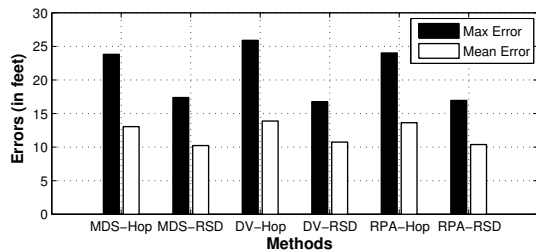

Figure 17. Comparison: RSD vs. Hop Distance

In our evaluation, the localization error is defined as the distance from the true position of a node to its estimated location. Figure 17 shows the maximum and mean localization errors from the results in Figure 16 for all six methods. We can clearly observe that all the RSD-embedded methods ("*-RSD") have smaller errors than their original "*-Hop" versions. Specifically, the maximum and mean errors of "MDS-Hop", "DV-Hop" and "RPA-Hop" get reduced by 27%, 22%, 35%, 23%, 29% and 24%, respectively.

6.2 Experiment II: Regular 2D Network

The second test-bed evaluation targets a 2-dimensional grid-shaped network with consideration of investigation on the impact of irregular network topology to localization performance. A grid can be easily transformed to irregular shapes by simply removing "pixels" in the grid.

6.2.1 Experiment Setup

Figure 18 shows the experiment scenario of a large open-air parking lot at night. We placed 49 MICAz sensor nodes in a 7×7 grid shape, as shown in the left of Figure 19. The network covered an area of about 100×100 square feet. Rows

Figure 18. Test-bed Experiments II: Regular 2D Network
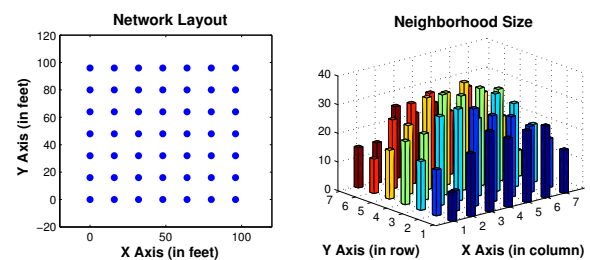

Figure 19. Network Layout and Neighborhood Size

and columns were approximately 16 feet apart. Similarly to Experiment I, each node was left 8 inches above the ground with antennas pointing to the sky.

Table 3 lists some key statistics regarding the experiment. This 49-node system is a 3-hop network. 1-hop radio range varies from about 20 feet to 100 feet among different node pairs along diverse directions. The right subfigure in Figure 19 shows the neighborhood size of each node. The X-axis and Y-axis in this figure depict the index of rows and columns in the network. The height of each bar indicates the 1-hop neighborhood size of the node at that position. This figure verifies that nodes in the center of the network have more 1-hop neighbors while the boundary nodes have smaller neighborhood size.
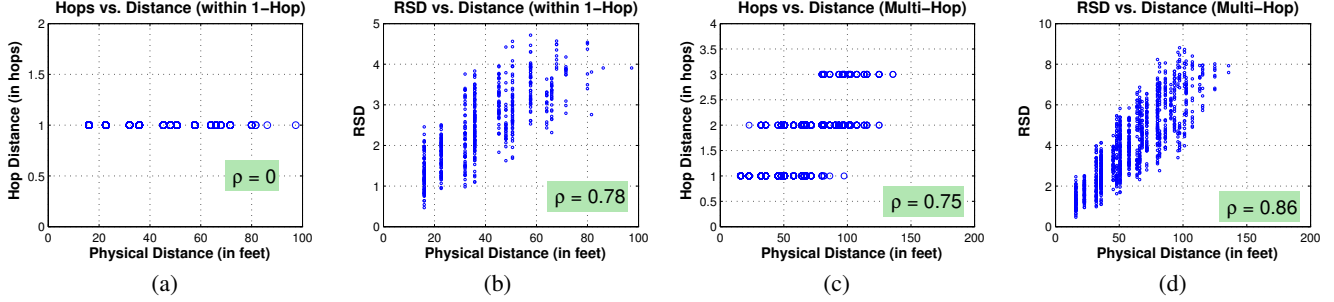
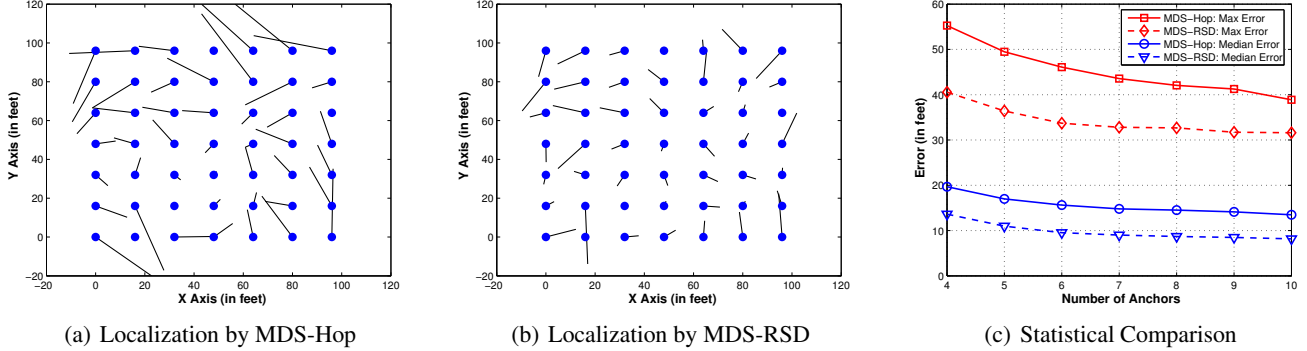Figure 20. Distance Correlation Comparison: RSD vs. Hop (Regular 2D Network)

(a) Localization by MDS-Hop (b) Localization by MDS-RSD (c) Statistical Comparison

Figure 21. Localization Results of MDS-RSD and MDS-Hop

Table 3. Statistics of the 2D Network

Statistics	1-Hop	2-Hop	3-Hop
Number of Node Pairs	575	519	82
Minimum Distance (in feet)	16	22.63	80
Median Distance (in feet)	35.78	71.55	97.32
Maximum Distance (in feet)	97.32	124.96	135.76

6.2.2 *Distance Correlations*

Figure 20 shows the correlations between hop-based distance, RSD, and physical distance in this grid-shaped 2D network. Figure 20(a) and 20(b) confirms that RSD provides a better sub-hop resolution than hop-based distance. Furthermore, Figure 20(b) also verifies that for MICAz motes using monopole whip antenna with radiation pattern close to isotropic [53], neighborhood ordering based on RSS can be a good heuristic indicator for physical distance. For multi-hop distance, Figure 20(c) and Figure 20(d) show that RSD provides a higher correlation coefficient than hop-based distance, verifying that RSD is superior to hop-based distance as a relative distance for proximity expression.

6.2.3 *Localization Performance*

Since the network has only 3 hops that is not suitable for DV-Hop and RPS, to be fair, we only applied MDS-based algorithms in this evaluation. Performance comparisons for DV-Hop and RPS in non-linear-shape networks will be investigated in Section 7 via large-scale simulation.

Figure 21(a) and 21(b) depict localization results from MDS-Hop and MDS-RSD with 4 randomly selected anchors in the system, respectively. In both figures, the blue dots are the true positions of nodes and a line segment originating from each dot point to its estimated position. We can see from these two figures that MDS-RSD gives better localization accuracy than MDS-Hop in this case.

In order to eliminate possible bias caused by anchor selection, we randomly picked different numbers of anchors, from 4 to 10 in step of 1, and tried each for 1000 runs. Figure 21(c) plots the averaged maximum and median errors for both methods, from which we can conclude that (i) MDS-RSD offers a significantly better performance over MDS-Hop across all numbers of anchors, and (ii) more anchor gives slight gain in localization accuracy. By embedding RSD, both the maximum and median errors are reduced greatly. For example, for 4 anchor nodes, RSD reduces the maximum and median errors by about 27% and 30%, respectively.

Radio is notorious for irregularity, however, we are able to achieve better localization accuracy than considering connectivity alone. Test-bed experiments show that RSD provides sub-hop resolution and correlates more with physical distance than the traditional hop distance.

7 Simulation Evaluation

In this section, we report selected simulation results about the performance gain from the RSD design for large-scale networks under different system settings.

7.1 The Noise Model

In the simulation, we applied the widely used logarithmic attenuation model [43, 44, 47, 54] for RSS sensing:

$$P_{i,j}(t) = P_0 - 10\beta \log(\frac{PD(u_i, u_j)}{d_0}) + X_i(t) \qquad (12)$$

where $P_{i,j}(t)$ stands for the sensing result at node u_i for node u_j at time instance t. P_0 is the received power at a short reference distance d_0. $PD(u_i, u_j)$ is the physical distance between u_i and u_j. β is the path loss factor (also called fading factor in some literatures) and $X_i(t)$ is a random noise at time t for node u_i following $X_i(t) \sim N(0, \sigma_X^2)$. We set a reference
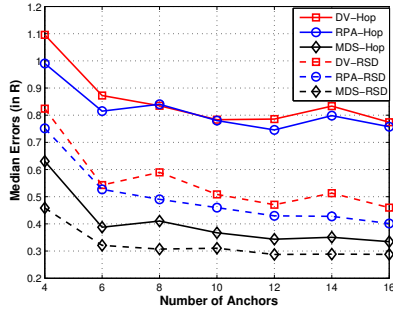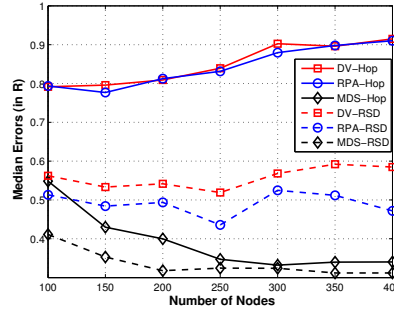
Figure 22. Different No. of Anchors

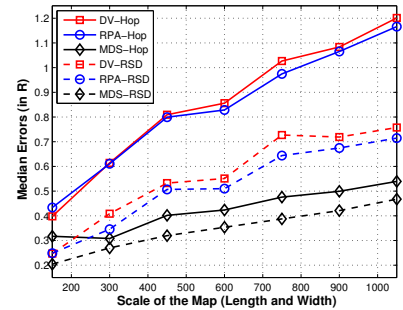**Figure 23. Different Node Densities**

Figure 24. Different System Scales

1-hop radio range of $R = 100$ feet. The corresponding signal strength was set as the receiver sensitivity threshold.

In the simulation, we modeled the area of interest as a square map without holes where radio can not reach. More complicated maps can be used with works [29, 31, 34]. Unless otherwise mentioned, Table 4 lists the default simulation configurations for the following sections. All the statistics reported were averaged over 50 runs for high confidence.

Table 4. Default Configurations in Simulation

Parameter	Default Values & Description
Field Area	500 (in feet)×500 (in feet)
Noise Model	$\beta = 4$, $\sigma_X = 6$ for the whole map
Number of Sensor Nodes	200, randomly deployed with uniform distribution
Number of Anchor Nodes	8, randomly deployed

7.2 The Effectiveness of RSD

This section evaluates the effectiveness of RSD by comparing localization errors between connectivity-based methods (MDS-Hop, DV-Hop, and RPA-Hop) with corresponding RSD-embedded versions (MDS-RSD, DV-RSD, and RPA-RSD) under various system configurations. We normalized the localization errors respect to a reference radio range $R = 100$ (feet) for consistency with the evaluations in previous literatures [26].

Impact of Anchor Density

In this experiment, we increased the number of anchor nodes in the network from 4 to 16 in steps of 2. As expected, Figure 22 shows that more anchors help improve the localization accuracy for all methods. More notable is that RSD embedding is more effective than adding several anchors, especially after 8 anchors when curves become flat. By embedding RSD, localization error gets reduced constantly by around 30% for DV-Hop and RPA-Hop, and by about 10% for MDS-Hop.

Impact of Node Density

In this experiment, we increased the number of nodes in the map from 100 to 400 in steps of 50. Figure 23 shows that: (i) RSD-applied methods always showed better performance (e.g., about 30% and 10% performance gain from DV-Hop/RPA-Hop and MDS-Hop, respectively); (ii) for DV-Hop and and RPA-Hop, increasing the node number from 100 to 200 did not affect the system accuracy too much. This is because at this stage, higher node density helps estimate the hop-based distance. While after 200 nodes, the hop distance can hardly be improved but more nodes are mapped to

identical estimated positions, bringing up the error statistically; (iii) for MDS-Hop and MDS-RSD, higher node density mostly gives smaller localization error. This is because the system becomes more robust to a single node's sensing error with more 1-hop neighbors.

Impact of System Scale

In this experiment, we enlarged dimension of the map from 150 feet (in length and width) to 1050 feet in steps of 150 feet. The number of nodes in the network was increased proportionally to maintain the same node density. The number of anchors was kept constant. Figure 24 shows the results. We can see that (i) RSD-applied methods always achieve better localization accuracy than their original versions based on hop distance; (ii) localization performance gets worse in larger systems because the number of anchor nodes were not increased proportionally; and (iii) MDS-based methods are more robust than DV-based approaches in terms of system scales. This is because MDS-based methods utilize both local proximity and the estimated distance to remote nodes, while DV and RPA depend more on the estimated distance to remote anchors, the error of which accumulates easily in larger networks.

7.3 The Robustness of RSD

In previous evaluations, the path loss factor β keeps a uniform value in the map. In real system deployments, radio path loss factor β is more than unknown but temporally dynamic and spatially unevenly distributed in the map [47, 54]. In this experiment, we generate a β distribution model according to the state-of-the-art empirical study [47, 54], for evaluating the robustness of localization algorithms in the case of spatially unevenly distributed β. The simulation randomly deployed 300 nodes with uniform distribution in a map of size 1000 feet × 1000 feet. As shown in Figure 25(a), β was around 4 but varied gradually across the map, with a "hill" located close to $X = 800, Y = 500$ and a "valley" located around $X = 200, Y = 500$. Basically, a hill indicates bigger β and a valley depicts smaller β.

Figure 25(b) shows the topology of deployed nodes. Each node is colored according to its true coordinates in the map, so similar colors indicate proximity. Figure 25(c) illustrates the 1-hop links in network. Each line segment (edge) stands for a 1-hop link. The path loss factor β of each link is assigned based on the patches of the spatial model it traverses. A link exists only when the signal strength at both terminal nodes are higher than the default RF sensitivity threshold.
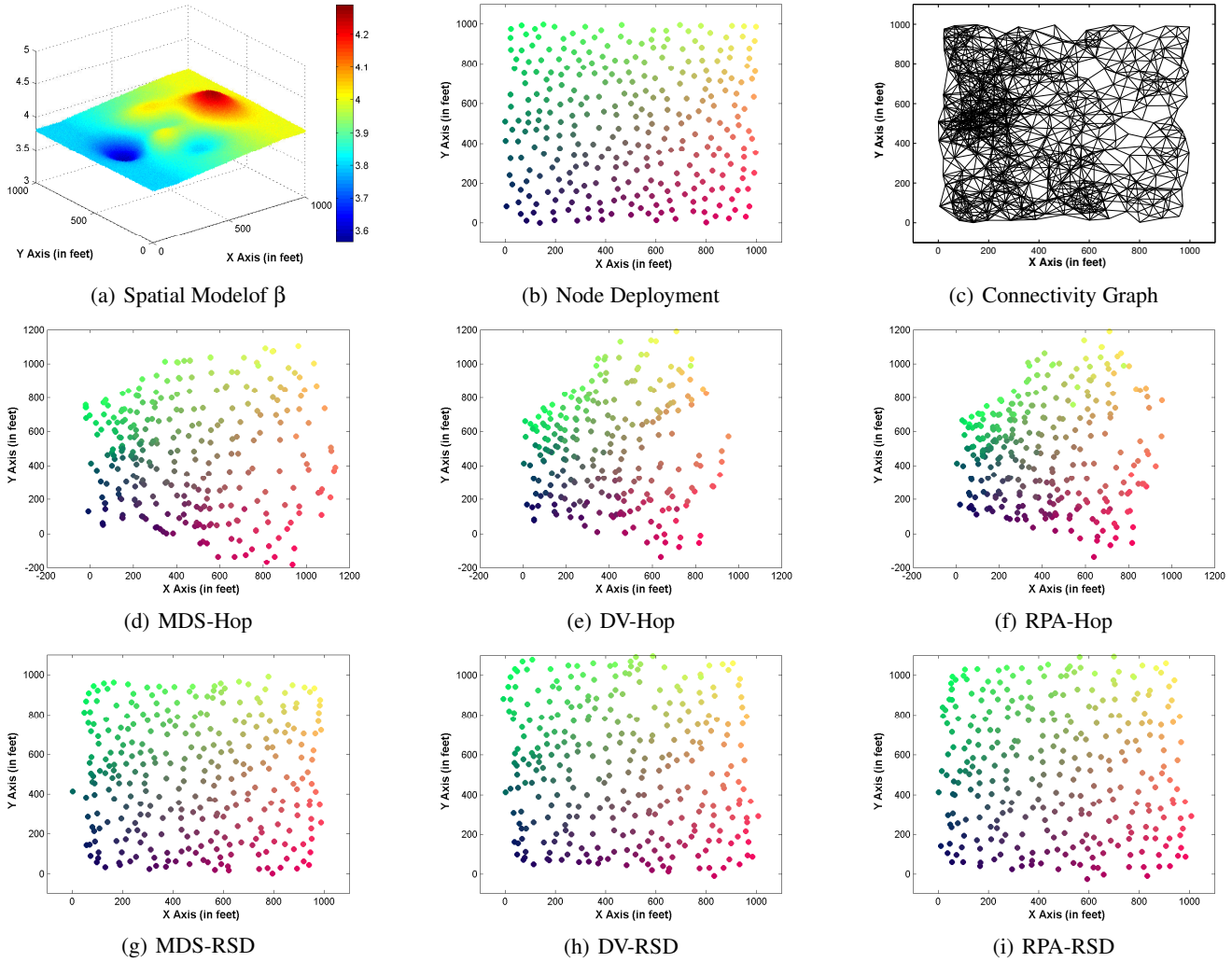
(a) Spatial Modelof β   (b) Node Deployment   (c) Connectivity Graph

(d) MDS-Hop   (e) DV-Hop   (f) RPA-Hop

(g) MDS-RSD   (h) DV-RSD   (i) RPA-RSD

Figure 25. Robustness of RSD for Unknown Spatial Fluctuation of Path Loss Factor β

We can see from the graph that the link density is higher in the left part of the map, especially near $X = 200, Y = 500$, where β "valley" exists in Figure 25(a). On the other hand, links are more sparse in the right part, especially close to $X = 800, Y = 500$, where β "hill" is located. This result is expected since a larger β creates a shorter communication range, thus nodes close to the "hill" are less connected.

Figures 25(d), 25(e), and 25(f) show the localization results given by MDS-Hop, DV-Hop, and RPA-Hop, respectively. We can see that the overall shape of the network gets distorted. Many nodes are closely clustered toward the position $X = 200, Y = 500$, while others are sparsely spread out from the point $X = 800, Y = 500$. This interesting phenomenon is consistent with the connectivity graph shown in Figure 25(c). In fact, lower β value allows nodes to have a longer 1-hop radio range, so from the viewpoint of hop distance, it creates *an illusion of shorter physical distances* between two nodes; while bigger β values have the opposite effects. RSD does not fluctuate with the radio range due to the regulation step in the design. This nice feature gets confirmed from localization results shown in Fig-

ures 25(g), 25(h) and 25(i) where the overall shape of the network is correct for all RSD-applied methods. Therefore, RSD provides the important feature that it is robust for uneven path loss factors across the area.

Simulation results tells that (i) RSD offers a nice feature of robustness to the spatially unevenly distributed radio path loss; (2) embedding RSD helps greatly improve the system accuracy of connectivity-based localization algorithms.

8 Conclusions

This paper proposes a relative distance for achieving range-free localization beyond connectivity. Starting from the concept of neighborhood ordering as a high-dimensional location-dependent signature for each node in the network, we presented the design of RSD, which quantifies the difference between signatures to capture distance relationships among neighboring nodes with sub-hop resolution. With little overhead, RSD can be conveniently embedded in many connectivity-based localization algorithms to improve localization accuracy. System evaluations on test-beds demonstrated that applying RSD helps solve the ambiguity problem

and considerably enhance localization accuracy. In addition, extensive simulation reveals that RSD offers a nice feature of robustness for spatially unevenly distributed radio path loss, preventing from outputting an distorted network topology.

9 Acknowledgments

10 References

[1] D. Culler, D. Estrin, M. Srivastava. Guest Editors' Introduction: Overview of Sensor Networks, Computer, 37(8), 2004.

[2] R. Want and B.N. Schilit. Guest Editors' Introduction: Expanding the Horizons of Location-Aware Computing. Computer, 34(8), 2001.

[3] G. Simon, M. Maróti, Á. Lédeczi, G. Balogh, B. Kusý, A. Nádas et al. Sensor Network-based Countersniper System. In SenSys'04.

[4] G. Werner-Allen, J. Johnson, M. Ruiz, J.Lees. Monitoring Volcanic Eruptions with a Wireless Sensor Network. In EWSN'05.

[5] A. Terzis, A. Anandarajah, K. Moore, I-J. Wang. Slip Surface Localization in Wireless Sensor Networks for Landslide Prediction. In IPSN'06.

[6] J. Hicks, A. Christian, B. Avery. HPL-2005-115 Integrating Presence and Location Services using SIP. HP Labs Tech Report. June 2005.

[7] B. Karp and H. T. Kung. GPSR: Greedy Perimeter Stateless Routing for Wireless Networks. In MobiCom'00.

[8] Y. J. Kim, R. Govindan, B. Karp, S. Shenker. Geographic Routing Made Practical. In NSDI'05.

[9] P. Bahl and V. N. Padmanabhan. Radar: An In-building RF-based User Location and Tracking System. In InfoCom'00.

[10] N. B. Priyantha, A. Chakraborty, H. Balakrishnan. The Cricket Location-support System. In MobiCom'00.

[11] A. Savvides, C. C. Han, M. B. Strivastava. Dynamic Fine-grained Localization in Ad-hoc Networks of Sensors. In MobiCom'01.

[12] X. Z. Cheng, A. Thaeler, G. L. Xue and D. C. Chen. TPS: A Time-Based Positioning Scheme for Outdoor Wireless Sensor Networks. In InfoCom'04.

[13] S. Lanzisera, D. T. Lin and K. S. J. Pister. RF Time of Flight Ranging for Wireless Sensor Network Localization. In WISES'06.

[14] J. Liu, Y. Zhang and F. Zhao. Robust Distributed Node Localization with Error Management. In MobiHoc'06.

[15] D. Moore, J. Leonard, K. S. J. Pister. Robust Distributed Network Localization with Noisy Range Measurements. In SenSys'04.

[16] X. Z. Cheng, H. Shu, Q. L. Liang, D. H-C Du. Silent Positioning in Underwater Acoustic Sensor Networks. IEEE TVT, 57(3), 2008.

[17] M. Maróti, B. Kusý, G. Balogh, P. Völgyesi, A. Nádas, K. Molnár, S. Dóra, Á Ledeczi. Radio Interferometric Geolocation. In Sensys'05.

[18] H-L. Chang, J-B. Tian, T-T Lai, H-H Chu, P. Huang. Spinning Beacons for Precise Indoor Localization. In SenSys'08.

[19] Amitabh Basu, Jie Gao, J. S. B. Mitchell, G. Sabhnani. Distributed Localization by Noisy Distance and Angle Information, In MobiHoc'06.

[20] Z. Yang, and Y. H. Liu, Quality of Trilateration: Confidence-based Iterative Localization, In ICDCS'08.

[21] K. Chintalapudi, R. Govindan, R. Govindan, G. Sukhatme. Ad-hoc Localization Using Ranging and Sectoring. In InfoCom'04.

[22] D. Niculescu and B. Nath. Ad Hoc Positioning System (APS) using AOA. In InfoCom'03.

[23] J. Bruck, J. Gao, A. Jiang. Localization and Routing in Sensor Networks by Local Angle Information. In MobiHoc'05.

[24] D. K. Goldenberg, P. Bihler, M. Cao, M. Cao, J. Fang et al. Localization in Sparse Networks using Sweeps. In MobiCom'06.

[25] N. Bulusu, J. Heidemann, and D. Estrin, GPS-less Low Cost Outdoor Localization for Very Small Devices, IEEE Per. Com. Mag. 7(4), 2000.

[26] Y. Shang, W. Ruml, Y. Zhang, and M. P.J.Fromherz. Localization from Mere Connectivity. In MobiHoc '03.

[27] D. Niculescu and B. Nath, DV Based Positioning in Ad Hoc Networks, Journal of Telecommunication Systems. 22(4), 2003.

[28] T. He, C. Huang, B. M. Blum, J. A. Stankovic. Range-free Localization Schemes in Large-scale Sensor Networks. In MobiCom'03.

[29] Sol Lederer, Yue Wang, Jie Gao, Connectivity-based Localization of Large Scale Sensor Networks with Complex Shape, In InfoCom'08

[30] N. Bulusu, J. Heidemann, D. Estrin and T. Tran. Self-configuring localization systems: Design and Experimental Evaluation. ACM Trans. on Embedded Comp. Sys., 2004, 3(1).

[31] M. Li and Y. H. Liu, Rendered Path: Range-Free Localization in Anisotropic Sensor Networks with Holes, In MobiCom'07.

[32] R. Nagpal, H. Shrobe, J. Bachrach. Organizing a Global Coordinate System from Local Information on An Ad hoc Sensor Network. In IPSN'03.

[33] C. Savarese,J. M. Rabaey, K. Langendoen. Robust Positioning Algorithms for Distributed Ad-Hoc Wireless Sensor Networks. In USENIX'02.

[34] C. Wang and L. Xiao. Locating Sensors in Concave Environments. In InfoCom'06.

[35] L. Doherty, K. S. J. Pister, L. El Ghaoui. Convex Position Estimation in Wireless Sensor Networks. In InfoCom'01.

[36] L. Römer. The Lighthouse Location System for Smart Dust. In MobiSys'03.

[37] R. Stoleru, T. He, J. A. Stankovic, D. Luebke. A High-accuracy, Low-cost Localization System for Wireless Sensor Networks. In SenSys'05.

[38] Z. G. Zhong, T. He. MSP: Multi-Sequence Positioning of Wireless Sensor Nodes. In SenSys'07.

[39] K. Whitehouse, C. Karlof, A. Woo et al. The Effects of Ranging Noise on Multihop Localization: An Empirical Study. In IPSN'05.

[40] K. Whitehouse, C. Karlof and D. Culler. A Practical Evaluation of Radio Signal Strength for Ranging-based Localization. SIG. Mob. Comp. Com. Rev., 11(1), 2007.

[41] K. Srinivasan. Understanding the Causes of Packet Delivery Success and Failure in Dense Wireless Sensor Networks. Technical report SING-06-00.

[42] E. Miluzzo, X. Zheng, K. Fodor and A. T. Campbell. Radio Characterization of 802.15.4 and its impact on the Design of Mobile Sensor Networks. In EWSN'08.

[43] J. A. Costa, N. Patwari, A. O. Hero III. Distributed Weighted-Multidimensional Scaling for Node Localization in Sensor Networks. ACM Trans. on Sensor Networks, 2(1), 2006.

[44] N. Patwari1, A. O. Hero III, J. A. Costa. Learning Sensor Location from Signal Strength and Connectivity. Advances in Information Security series, Vol. 30, Springer, 2006, ISBN 978-0-387-32721-1.

[45] J. Hightower, G. Borriello and R. Want. SpotON: An indoor 3D Location Sensing Technology Based on RF Signal Strength, U of Washington CSE Report.

[46] A. Varshavsky, E. d. Lara et al. GSM Indoor Localization, Pervasive and Mobile Computing Journal (PMC), Elsevier, 3(6),2007

[47] T.S. Rappaport, Wireless Communications, Principles and Practice, Prentice Hall, 1996.

[48] K. Yedavalli and B. Krishnamachari. Sequence-Based Localization in Wireless Sensor Networks. IEEE Tran. on Mob. Comp., 7(1), 2008.

[49] Yu Zhang, Lin Zhang, Xiuming Shan, Ranking-Based Statistical Localization for Wireless Sensor Networks. IEEE WCNC 2008.

[50] E.W. Weisstein, Circle Division by Lines. MathWorld online, http://mathworld.wolfram.com/CircleDivisionbyLines.html

[51] S. Carlsson. A Variant of Heapsort with Almost Optimal Number of Comparisons. Inf. Process. Lett. 24(4), 1987.

[52] M. J. Greenacre. Theory and Applications of Correspondence Analysis. Academic Press Inc., London, UK, 1984.

[53] Kent Smith. Antennas for Low Power Applications. RFM Corp., Dallas. http://www.rfm.com/corp/appdata/antenna.pdf

[54] The ns Manual, Chapter 18: Radio Propagation Models. Editor: Kevin Fall et al. http://www.isi.edu/nsnam/ns/doc/index.html