

Exploring Landmark Placement Strategies for Topology-Based Localization in Wireless Sensor Networks

Farid Benbadis Katia Obraczka Jorge Cortés Alexandre Brandwajn

Abstract—In topology-based localization, each node in a network computes its hop-count distance to a finite number of reference nodes, or “landmarks”. This paper studies the impact of landmark placement on the accuracy of the resulting coordinate systems. The coordinates of each node are given by the hop-count distance to the landmarks. We show analytically that placing landmarks on the boundary of the topology yields more accurate coordinate systems than when landmarks are placed in the interior. Moreover, under some conditions, we show that uniform landmark deployment on the boundary is optimal. This work is also the first empirical study to consider not only uniform, synthetic topologies, but also non-uniform topologies resembling more concrete deployments. Our simulation results show that, in general, if enough landmarks are used, random landmark placement yields comparative performance to placing landmarks on the boundary randomly or equally spaced. This is an important result since boundary placement, especially at equal distances, may turn out to be infeasible and/or prohibitively expensive (in terms of communication, processing overhead, and power consumption) in networks of nodes with limited capabilities.

I. INTRODUCTION

Sensor networks typically refer to a collection of nodes that have sensing, processing, storage, and (wireless) communication capabilities. In general, because of their small form factor and low cost, sensor network nodes often have limited capabilities; furthermore, as they are frequently battery powered, energy is a premium resource that directly impacts the lifetime of nodes and the sensor network as a whole.

Sensor networks have a wide range of applications with significant scientific and societal relevance [1].

Farid Benbadis is with the Laboratoire LIP6/CNRS, UMR 7606, Université Pierre et Marie Curie – Paris VI, Paris, France, farid.benbadis@lip6.fr

Katia Obraczka and Alexandre Brandwajn are with the Department of Computer Engineering, University of California, Santa Cruz, California, USA, {katia,alex}@soe.ucsc.edu

Jorge Cortés is with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, California, USA, cortes@ucsd.edu

Example applications include environmental monitoring, object tracking, surveillance, and emergency response and rescue operations. While some scenarios allow for manual placement of sensor network nodes in the field, others require “random” deployment, where nodes are simply “dropped” (e.g., from an airplane), and once they land they need to self-organize into a network and start performing the task at hand.

One important step in self-organization is positioning, which refers to having nodes find their physical location [2]. Node positioning is required by sensor network core functions such as topology control, data aggregation, and routing [3], [4], and may also be needed by a number of applications. For instance, the sensor network could be tasked to report the air temperature’s running average by geographic region.

One clear solution to the positioning problem is provided by satellite-based systems [5], [6], [7], among which the Global Positioning System (GPS) is probably the most widely used. However, in some scenarios, satellite-based localization is not possible. This is the case of indoor, underwater, and underground deployments. Furthermore, equipping sensor nodes with GPS receivers might be prohibitive for reasons related to cost, form factor, energy consumption, or a combination of them. A possible alternative is to equip only a subset of the nodes with GPS receivers and have all other nodes compute their position relative to the GPS-capable nodes. The recent work by [8] provides a theoretical study of this problem using graph rigidity. For instance, in a multi-tiered heterogeneous deployment, nodes that have extended life batteries and/or have higher processing power could have GPS capabilities. However, this may still be infeasible in some deployments.

In situations where no GPS anchors can be used, nodes are clueless about their geographic coordinates. Therefore, numerous GPS-less methods have been proposed. Coordinates generated by such methods are known as *virtual* coordinates. Even though some applications may still require real coordinates, virtual coordinate positioning can be used by core functions such as position-based

routing, topology control, and data aggregation.

Motivated by the state-of-the-art on GPS-less positioning systems, this paper aims at evaluating the effect of landmark placement strategies on the quality of the resulting virtual coordinate system. One of our goals is to investigate if landmark placement/election can be either simplified, or, better, avoided altogether by assigning the role of landmarks to any node in the topology.

In the next section, we put our work in perspective by providing some background on GPS-less positioning mechanisms.

II. BACKGROUND AND FOCUS

In general, GPS-less positioning techniques may be classified as (1) using physical measurements (*a.k.a.* range-based positioning) or (2) using topological information (*a.k.a.* range-free positioning). Examples of measurement-based GPS-less techniques include mechanisms that use propagation laws [9] to approximate Euclidean distance using received signal strength (RSS). The RSS can be converted into distance either directly, if the propagation law is uniform and known, or using multiple signals and time difference of arrival (TDoA) [10]. Then, trilateration techniques allow node coordinates to be inferred. The use of directional antennas to triangulate positions has also been proposed [11]. One main drawback of measurement-based mechanisms is that they typically require specialized equipment or capabilities to perform the measurements.

Topology-information based positioning, on the other hand, relies solely on topological information. For example, the approach in [12] first discovers border nodes, then computes their relative coordinates, and finally infers non-border node coordinates relative to border nodes. The correctness of this algorithm has been analyzed in [13]. Alternatively, in GPS-FREE-FREE [14], JUMPS [15], VCAP [16], and BVR [17], the hop distances to reference nodes, or “landmarks,” are transformed into “virtual coordinates.” The hop distance from a node to a landmark is given by the minimum number of hops from that node to the landmark. GPS-FREE-FREE uses trilateration to obtain virtual coordinates from corresponding hop distances, while JUMPS, VCAP, and BVR use the hop distances directly as nodes’ coordinates. Note that the denser the network, the more accurate it is to approximate Euclidean distance using hop distance. However, most existing hop-count based positioning systems make the strong assumption that, for better performance (e.g., accuracy), landmarks need to be placed along the perimeter of the topology at equal distances from one another. To the best of our

knowledge, this assumption is purely intuitive, and has never been justified either empirically, experimentally, or analytically.

Thus, the focus of this paper is to explore the effect of landmark placement on the accuracy of the resulting coordinate system. To our knowledge, our paper is the first to show analytically that, indeed, placing landmarks on the boundary of the topology yields more accurate coordinate systems than when landmarks are placed anywhere in the interior. Moreover, under some conditions, we show that uniform landmark deployment on the boundary is optimal. This is also the first empirical study to consider uniform, synthetic topologies as well as non-uniform topologies resembling more realistic deployments. In our study, we evaluate different landmark placement strategies, namely: (1) “uniform boundary placement” as in JUMPS [15] and VCAP [16], where landmarks are placed at the boundary of the topology at equal distances from one another; (2) “random boundary placement”, where landmarks are placed on the boundary but at random intervals; and (3) “random placement,” which places landmarks anywhere in the topology completely at random, as in BVR [17]. As performance metrics, we consider the ability to uniquely identify a node and how well position-based routing performs over the resulting coordinate system (when compared against routing with real coordinates).

In summary, the contributions of this paper revolve around two main questions: “How does landmark placement affect the accuracy of the resulting hop-count coordinate system?” and “Can landmark placement be avoided altogether?” In answering the first question, our simulation results confirm that placing landmarks on the topology periphery yields more accurate coordinates. The answer to the second question is critical when designing self-organizing networks, since border node selection/placement may be too expensive or even infeasible in some deployments. Our results also show that, in general, landmark placement strategies only have significant performance impact when the number of landmarks is low. In other words, if enough landmarks are used, random landmark placement yields comparative performance to placing landmarks on the boundary (randomly or equally spaced). We contend that the work here is a first step towards the development of reliable and efficient methods for landmark placement in virtual positioning systems.

The remainder of this paper is organized as follows. In Section III, we describe existing hop-count positioning systems for sensor networks in more detail. Section IV shows analytically the effect of landmark placement on the quality of the resulting coordinate system when

uniform topologies are used. The methodology and simulation results on the impact of landmark placement considering both uniform and non-uniform topologies are described in Section V. Finally, Section VI presents our concluding remarks and identifies directions for future work.

III. HOP-COUNT BASED POSITIONING SYSTEMS

The use of topological or hop-count based localization methods in wireless sensor networks is advantageous because they are simple and do not require additional equipment or devices. Below, we describe some notable examples of hop-count based positioning techniques.

GPS-FREE-FREE [14] constructs a two-dimensional coordinate system based on hop-count distances using three landmarks. Landmarks in GPS-FREE-FREE are nodes chosen from the interior of the topology in such a way that they form an equilateral triangle. Each landmark broadcasts a packet in order to allow other nodes to discover their hop-count distance to it. This packet also contains virtual position of the landmark.¹ Thus, each landmark knows its hop-distances to landmarks and their virtual coordinates. Based on this knowledge, and using the hop-distance as a metric, each node calculates its virtual coordinates through trilateration.

VCAP [16] is another hop-count positioning algorithm very similar to GPS-FREE-FREE. VCAP also uses three landmarks at equal distances from each other but instead of a two-dimensional system, VCAP builds a three-dimensional one. In other words, the hop-count distances to the landmarks are directly used as the three coordinates of a node. The advantage of VCAP when compared to GPS-FREE-FREE is that (1) it requires less computation, since the trilateration phase is avoided and (2) it provides better accuracy, since the hop count to the third landmark is used as a real coordinate. Another difference between GPS-FREE-FREE and VCAP is in how they place the landmarks. While both algorithms form an equilateral triangle with the landmarks, VCAP positions them on the boundary of the topology, while GPS-FREE-FREE places them in the interior.

JUMPS [15] is another positioning system based on hop-distances. As VCAP, JUMPS places landmarks on the border of the network at equal distances of one another and uses, as coordinates, hop-count distances

¹By definition, in GPS-FREE-FREE all the nodes consider that the x axis is given by the straight line determined by landmarks 1 and 2, with the convention that they are, respectively, placed at $(0, 0)$ and $(d_{2,1}, 0)$, where $d_{2,1}$ is the hop-distance between landmarks 1 and 2. The third landmark computes its coordinates as any non-landmark node, but setting positive the coordinate on the y axis.

to landmarks. JUMPS utilizes, however, up to ten landmarks instead of the three used in VCAP. It has been shown [15] that adding landmarks increases the accuracy of the resulting coordinate system.

The common point shared by GPS-FREE-FREE, VCAP, and JUMPS is the assumption that landmarks can be manually placed at specific locations. For that to happen, either manual deployment or landmark election mechanisms are required. Many scenarios make manual deployment infeasible (e.g., dropping sensors from a plane in hostile, hard to access regions). In such cases, election algorithms are required to select border nodes with specific placement. The fact that these algorithms may be prohibitively expensive (as they require additional computational and several rounds of communication among nodes), highlights the importance of avoiding landmark placement and election, as done in BVR [17]. However, BVR does not explicitly justify the choice of random landmark placement as well as the reason for using larger numbers of landmarks. The results from our work provide an explanation for these design choices.

IV. THEORETICAL ANALYSIS

In this section, we show that, for uniform topologies, placing landmarks at the boundary of the topology results in a more accurate coordinate system. Under some simplifications, we also show that uniform landmark deployment on the boundary is optimal. In our analysis, the performance metrics used are the *average zone size* and the *maximum zone size*. These metrics, which are also used in the simulation evaluation, are defined below.

Definition 1: A *zone* is a set of nodes sharing the same virtual coordinates. The *zone size* is the largest real distance between two nodes in the same zone.

Figure 1 illustrates this definition.

Consider an environment of interest $Q \subset \mathbb{R}^2$ where n nodes are uniformly deployed. For simplicity, we take $Q = B(0, R)$, the ball of center 0 and radius R . Assume n nodes are uniformly deployed on Q . Consider N landmarks $\lambda_1, \dots, \lambda_N$ placed within Q . Here, we discuss how the configuration of the landmarks affects the number of zones corresponding to the deployment of the nodes.

For each landmark $\lambda_i \in Q$, the *hop-distance* function $h_i : Q \rightarrow \mathbb{N}$ measures the number of hops $h_i(p)$ from a node at $p \in Q$ to the landmark λ_i . Note that this function depends on the specific network topology. Consider the function $h = (h_1, \dots, h_N) : Q \rightarrow \mathbb{N}^N$. For $c \in \mathbb{N}^N$, $\{x \in Q \mid h(x) = c\}$ is the *level set* of h corresponding to c . Note that the level sets of h correspond precisely

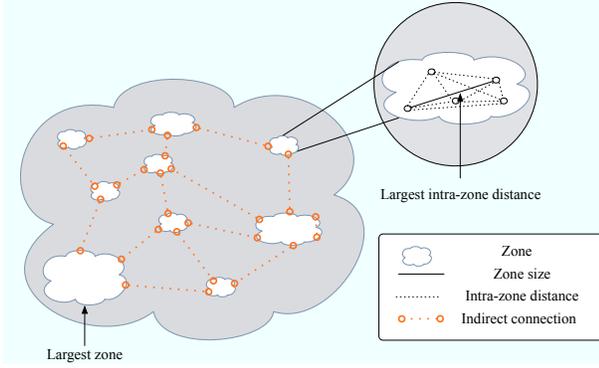


Fig. 1. Schematic representation of zones in a network. Zones are represented by clouds. The distance between any two nodes among the same zone is noted *intra-zone distance*. The largest intra-zone distance is the *zone size*, represented with a plain line.

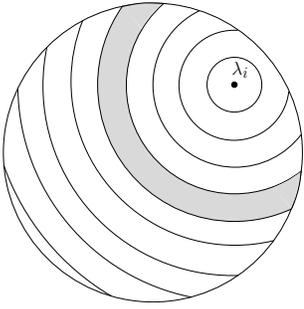


Fig. 2. Level sets of the hop-distance function corresponding to the landmark λ_i . The shaded area represents a sample level set, which is the result of the intersection of the environment with an annulus centered at λ_i and of radii r_1, r_2 differing by r .

to the zones. In other words, $p_1, p_2 \in Q$ are in the same zone if and only if they belong to the same level set of h , i.e., $h_i(p_1) = h_i(p_2)$ for all $i \in \{1, \dots, N\}$.

Let us therefore study the level sets of the individual hop-distance functions h_i . Since the nodes are uniformly deployed, we make the simplifying assumption that n is sufficiently large so that the hop-distance function h_i can be approximated by the Euclidean distance between p and λ_i divided by the communication radius. Specifically, $h_i(p) = \|p - \lambda_i\|/r$. Under this assumption, the level sets of h_i are the intersection of the environment Q with the annuli

$$B(\lambda_i, r_1, r_2) = \{x \in \mathbb{R}^2 \mid r_1 \leq \|x - \lambda\| \leq r_2\},$$

centered at λ_i and with radii r_1, r_2 differing by exactly r (the communication radius between agents). Figure 2 illustrates this.

A. Optimality of landmark placement on the boundary

From the previous discussion, it is clear that placing the landmarks at the boundary of the environment is advantageous for our two topological measures (average

zone size and maximum zone size). We formalize this observation in the following proposition. In the statement, ∂Q denotes the boundary of Q .

Proposition 1: Consider the hop-distance function $h_i : Q \rightarrow \mathbb{N}$ associated to a landmark $\lambda_i \in Q$. If $\lambda_i \in \partial Q$, then both the number of level sets of h_i and their area are optimized.

Proof: The number m_i of level sets associated with h_i is lower bounded by $\lceil R/r \rceil$ (when λ_i is placed at the center of the environment) and upper bounded by $\lceil 2R/r \rceil$ (when λ_i is placed at the boundary of the environment). Moreover, for each $k \in \{1, \dots, \lceil R/r \rceil\}$, the area of the intersection $Q \cap B(\lambda_i, (k-1)r, kr)$ is upper bounded by $(2k-1)\pi r^2$ (when λ_i is placed at the center of the environment) and lower bounded by $k^2 r^2 \arccos(\frac{kr}{2R}) + R^2 \arccos(1 - \frac{k^2 r^2}{2R^2}) - krR \sqrt{1 - \frac{k^2 r^2}{4R^2}} - (k-1)^2 r^2 \arccos(\frac{(k-1)r}{2R}) - R^2 \arccos(1 - \frac{(k-1)^2 r^2}{2R^2}) + (k-1)rR \sqrt{1 - \frac{(k-1)^2 r^2}{4R^2}}$ (when λ_i is placed at the boundary of the environment). Finally, note that, as one moves the location of λ_i from the center of the environment to the boundary along a straight line, the areas of the level sets corresponding to $k \in \{1, \dots, \lceil R/r \rceil\}$ are monotonically non-increasing. This lost area goes to the level sets corresponding to $k \in \{\lceil R/r \rceil + 1, \dots, \lceil 2R/r \rceil\}$, which appear successively as λ_i approaches the boundary. ■

Note that the area of each level set is inversely proportional to the number of nodes contained in the level set. Therefore, the smallest the area, the fewer the number of nodes with the same hop coordinate with respect to λ_i , which in turn makes smaller the zone size. Regarding average zone area, since the sum of the areas of the zones is equal to the area of the environment, we deduce

$$\text{Average zone area} = \frac{\pi R^2}{m},$$

where m is the number of zones corresponding to the landmark placement $\lambda_1, \dots, \lambda_N$. These results lead us to conjecture that the uniform landmark placement on ∂Q is optimal for the average zone size, because it maximizes the number of intersection between the annuli of the various landmarks, and therefore, maximizes the number of zones.

B. Optimality of uniform landmark placement for maximum zone size

Next, we examine the optimality of the uniform landmark placement on the boundary of the environment with regards to the maximum zone size measure. We start by introducing some basic notation.

1) *Geodesic distance on the circle*: Without loss of generality, we take $R = 1$ (the arguments below can be carried out analogously for arbitrary R). Let \mathbb{S}^1 denote the circle of radius 1. Normally, we refer to points in \mathbb{S}^1 using angle notation, $\theta \in [0, 2\pi)$. Alternatively, one could use Euclidean coordinates $(x, y) \in \mathbb{R}^2$, with $x^2 + y^2 = 1$. Both systems of coordinates are related by

$$(x, y) = (\cos \theta, \sin \theta), \quad \theta = \arctan\left(\frac{y}{x}\right).$$

Given two points $\theta_1, \theta_2 \in \mathbb{S}^1$, let $\text{dist}_g(\theta_1, \theta_2)$ be the *geodesic distance* between θ_1 and θ_2 defined by $\text{dist}_g(\theta_1, \theta_2) = \min\{\text{dist}_c(\theta_1, \theta_2), \text{dist}_{cc}(\theta_1, \theta_2)\}$, where

$$\begin{aligned} \text{dist}_c(\theta_1, \theta_2) &= (\theta_1 - \theta_2) \pmod{2\pi}, \\ \text{dist}_{cc}(\theta_1, \theta_2) &= (\theta_2 - \theta_1) \pmod{2\pi}, \end{aligned}$$

are the path lengths from θ_1 to θ_2 traveling clockwise and counterclockwise, respectively. Here $\theta \pmod{2\pi}$ is the remainder of the division of θ by 2π . Given two points in \mathbb{S}^1 , the relationship between their Euclidean and geodesic distances is given by

$$\text{dist}_g(\theta_1, \theta_2) = 2 \arcsin\left(\frac{\|(x_1, y_1) - (x_2, y_2)\|}{2}\right). \quad (1)$$

2) *Re-phrasing the “minimize-maximum-zone-size” optimization problem*: In our forthcoming discussion, we make two important simplifications: (i) we restrict our attention to the boundary of Q , and consider the intersection of the zones with ∂Q , instead of considering the zones in the full environment Q , and (ii) we consider the geodesic distance on ∂Q , rather than the Euclidean one. To emphasize the latter fact, we denote by $B_g(\lambda, r)$ the ball in ∂Q centered at θ with radius r with the geodesic distance. Two reasons justify (ii). On the one hand, from (1), one can see that this approximation is quite accurate on ∂Q for points that are up to an Euclidean distance $R = 1$. On the other hand, (ii) is reasonable when considering the problem of minimizing the maximum zone size in ∂Q with uniform landmark deployments. This is so because, given $i \in \{1, \dots, N\}$, any point in ∂Q that is more than an Euclidean distance $R = 1$ apart from λ_i must be less than an Euclidean distance $R = 1$ apart from some other λ_j , where the approximation of the Euclidean distance by the geodesic distance is accurate.

Note that the zones on ∂Q correspond to the level sets of $h_{|\partial Q} : \partial Q \rightarrow \mathbb{N}^N$. Each of these zones is an arc segment whose boundary points correspond to some landmark, see Figure 3. Therefore, for each landmark $\lambda_i \in \partial Q$, consider the intersection points between ∂Q and the boundary of the balls $B_g(\lambda_i, kr)$, with

$k \in \{1, \dots, \lceil 2R/r \rceil\}$. Note that any two consecutive intersection points are exactly at a geodesic distance r from each other. This implies that the zones appear

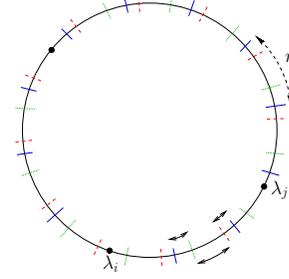


Fig. 3. Sample plot of the zones on ∂Q determined by an arbitrary placement of three landmarks (under the geodesic distance). Note that the zones appear periodically.

periodically at intervals of length r along ∂Q . Thus, in order to study the zone size, we identify points that are exactly r -apart, i.e.,

$$\theta_1 \sim \theta_2 \quad \text{iff} \quad \text{dist}_g(\theta_1, \theta_2) = r.$$

To obtain a simple representation of the quotient space $\partial Q / \sim$ under this identification, assume for simplicity that $2\pi R/r \in \mathbb{N}$, and fix any point $O \in \partial Q$ as a reference. Then, we have

$$\frac{\partial Q}{\sim} \equiv \mathbb{S}^1. \quad (2)$$

Under this identification, for each $i \in \{1, \dots, N\}$, the landmark $\lambda_i \in \partial Q$, λ_i and all the intersection points $\partial Q \cap \partial B_g(\lambda_i, kr)$, $k \in \{1, \dots, \lceil 2R/r \rceil\}$, get mapped to the same point in \mathbb{S}^1 . Also, under this identification, the zones in ∂Q correspond to the segments between two landmark locations in \mathbb{S}^1 . As a consequence, the problem of minimizing the maximum zone size in ∂Q , translated into \mathbb{S}^1 , becomes the disk-covering optimization problem discussed in the next section.

3) *Disk-covering optimization problem*: Given N points $\theta_1, \dots, \theta_N$ in \mathbb{S}^1 , consider the following *disk-covering* optimization problem:

For any θ in \mathbb{S}^1 , let $\min_{i \in \{1, \dots, N\}} \text{dist}_g(\theta, \theta_i)$ be the minimum distance of θ to the set of locations $\{\theta_1, \dots, \theta_N\}$. We refer to this distance as the *coverage* of θ provided by $\theta_1, \dots, \theta_N$. Larger values correspond to worse coverage. Consider the worst possible coverage provided by $\theta_1, \dots, \theta_N$ at a point of \mathbb{S}^1 , that is,

$$\mathcal{H}(\theta_1, \dots, \theta_N) = \max_{\theta \in \mathbb{S}^1} \min_{i \in \{1, \dots, N\}} \text{dist}_g(\theta, \theta_i).$$

We are interested in finding the minimizers of \mathcal{H} .

Interestingly, the function \mathcal{H} can be rewritten using the notion of Voronoi partition. The *Voronoi partition* of \mathbb{S}^1 generated by $\theta_1, \dots, \theta_N$ is the collection of sets V_i, \dots, V_N defined by

$$V_i = \{\theta \in \mathbb{S}^1 \mid \text{dist}_g(\theta, \theta_i) \leq \text{dist}_g(\theta, \theta_j) \text{ for } j \neq i\}.$$

In other words, V_i is the set of points that are closer to θ_i than to any of the other locations θ_j , $j \neq i$. In our case, V_i is a segment centered at θ_i , with boundary points determined by the mid points with its immediate clockwise and counterclockwise neighbors. Figure 4 illustrates this notion. Note that

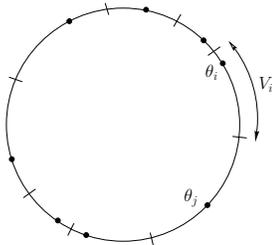


Fig. 4. Voronoi partition of \mathbb{S}^1 generated by $\theta_1, \dots, \theta_N$.

$$\mathcal{H}(\theta_1, \dots, \theta_N) = \max_{i \in \{1, \dots, N\}} \max_{\theta \in V_i} \text{dist}_g(\theta, \theta_i).$$

We are now ready to prove the following result.

Proposition 2: Any uniform deployment of N points on \mathbb{S}^1 is a global minimizer of \mathcal{H} .

Proof: Since \mathcal{H} is invariant under permutations, we assume without loss of generality that the locations $\theta_1, \dots, \theta_N$ are ordered in counterclockwise order in increasing order according to their index. Let $(\theta_1^*, \dots, \theta_N^*)$ be a uniform deployment on \mathbb{S}^1 , that is, $\text{dist}_g(\theta_i^*, \theta_{i+1}^*) = 2\pi/N$, where we define for convenience $\theta_{N+1}^* = \theta_1^*$. Note that $\mathcal{H}(\theta_1^*, \dots, \theta_N^*) = \pi/N$. Now, the result follows from noting that for any non-uniform configuration $(\theta_1, \dots, \theta_N)$, there must exist $i \in \{1, \dots, N\}$ such that $\text{dist}_g(\theta_i, \theta_{i+1}) > 2\pi/N$, and hence $\max_{\theta \in V_i} \text{dist}_g(\theta, \theta_i) > \pi/N$. Consequently, $\mathcal{H}(\theta_1, \dots, \theta_N) > \pi/N = \mathcal{H}(\theta_1^*, \dots, \theta_N^*)$. ■

Recall the equivalence between the disk-covering optimization problem and the problem of minimizing the maximum zone size in ∂Q discussed in Section IV-B2. In particular, note that the size of each segment (which is the image of a zone under the identification (2)) is twice the distance from the boundary point of the corresponding Voronoi cell to each of its generators. Given Proposition 2, we conclude that the uniform landmark deployment is optimal with regards to maximum zone size.

V. SIMULATION ANALYSIS

For the simulation experiments, we have written our own simulator since existing network simulators work at the packet-level and are too fine-grained for our purpose. Using a “higher-level” simulator allows us to simulate much larger topologies of up to 5,000 nodes.

Indeed, the simulator we conceived only (1) places nodes according to the distributions described in Section V-A, (2) determines hop-distances to landmarks by successive neighborhood discoveries and uses them as coordinates, and (3) discovers paths, based on the hop-count coordinate system, between randomly selected sources and destinations.

For simplicity, we simulated a perfect MAC layer, which means that 1) two nodes are neighbors if the distance between them is less than r , the radio coverage range described in Section V-A, and that 2) there is no packet loss during transmissions. Even though assuming a perfect MAC layer is not realistic, we claim it does not affect our comparative analysis, as all the strategies studied were subject to the same conditions.

As previously pointed out, unlike previous studies which only considered uniform network topologies, i.e., topologies where nodes are placed uniformly over the field, we also consider topologies with non-uniform node placement. Such topologies are motivated by more realistic scenarios such as campuses (e.g., universities) where nodes (users) tend to gather around access points. Our simulation experiments employing uniform topologies also validate our theoretical analysis. We use JUMPS [15] as the hop-count based positioning system.

A. Parameters

The environment considered is a circle of radius 1,000 meters, and the radio coverage range r of the nodes is 60 meters. We assume that nodes are homogeneous, i.e., they all have the same capabilities, and that neighborhood discovery is provided by the MAC layer.

- 1) *Number of Landmarks:* the simulated number of landmarks ranges from 3 to 10. Thus, we can evaluate the performance of both JUMPS [15] and VCAP [16].
- 2) *Landmark Placement:* the different landmark placement strategies are outlined below and illustrated in Figure 5.

- **UniBound** places landmarks on the boundary of the topology, at equal distances from each other. One possible landmark election algorithm to be used in a scenario where manual placement is not possible is described in VCAP [16].

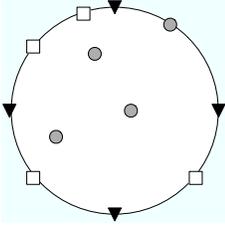


Fig. 5. This figure represents a 4-landmark circular topology. Triangles, squares and circles, respectively, represent the **UniBound**, **RandBound**, and **Rand** landmark placement strategies.

- In **RandBound**, landmarks are randomly placed on the boundary of the topology.
- **Rand** randomly places landmarks anywhere in the topology. Their location might be on the boundary or inside the disc area. In order to select N landmarks according to this strategy, techniques such as random selection, or choosing N nodes with the highest/lowest IDs can be employed. This strategy is used in the BVR algorithm [17]. These sample landmark selection mechanisms make it clear that **UniBound** is by far the most complex and costly, followed by **RandBound**. **Rand** is the simplest and least expensive. This means that doing away (completely or partially) with sensor selection can save significant network resources.

Recall that any node in the topology can be considered a landmark, i.e., no special capability is required to play this role. In our simulations, nodes are designated as landmarks depending on the specific landmark placement strategy employed.

- 3) *Number of Nodes*: The overall number of nodes, including landmarks, changes from 1,000 to 5,000, in steps of 2,000. Note that considering different number of nodes in a fixed environment and with a constant radio coverage range is equivalent to considering scenarios where the size of the environment and the radio coverage changes, but the number of nodes is held constant.
- 4) *Node Distribution*: As previously pointed out, two kinds of topologies are considered.
 - **Uniform topologies**: nodes are uniformly distributed over the field.
 - **Non-uniform topologies**: nodes are placed around “concentration points” according to a normal distribution. The number of concentration points ranges from 1% to 20% of the total number of nodes. The greater number of concentration points, the more uniform the topology.

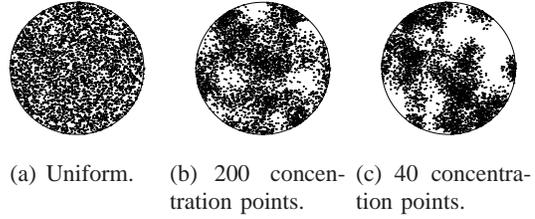


Fig. 6. Representation of a 4,000 nodes topology with three different node distributions. Only the first one is uniform.

We should point out that, unlike the studies conducted in VCAP and JUMPS, we also consider the case of disconnected networks. This means that nodes with no direct neighbors may exist. Such nodes can obtain coordinates from a subset of landmarks only, or do not obtain any coordinate at all.

For every scenario (i.e., combination of node distribution, number of landmarks, number of nodes, and landmark placement strategy), we execute 50 runs.

B. Performance Metrics

Zones: In order to evaluate the accuracy of a localization algorithm, researchers usually measure the distance error, which represents the Euclidean distance between the real position and the computed one. Such a measurement requires that both positions – real and virtual – are correlated. Note that the coordinates assigned to sensor nodes by JUMPS [15] and VCAP [16] do not express their geographical positions. Therefore, we cannot use the distance error to evaluate the accuracy of these localization systems.

Thus, similarly to VCAP, most of our performance metrics are based on the concept of *zones*. As described in Section IV, a zone is the set of nodes sharing the same virtual coordinates. The zone size is thus the maximum Euclidean distance, measured using real coordinates, between two nodes within the same zone. Thus, it provides a measure of the coordinate system’s ambiguity. In other words, the smaller the zone size, the more accurate the coordinates. A succinct pictorial description of zones is given in Figure 1.

In this paper, we consider three zone-related metrics. First, we evaluate, the *average zone size* for each scenario. Then, we measure the *maximum zone size*, i.e., the largest zone in a scenario. Note that if the maximum zone size is smaller than the node’s radio range, nodes sharing the same coordinates are physically neighbors and thus communicate directly. Finally, we count the *number of nodes per zone*. The lower this number, the more accurate the coordinate system. Ideally, we obtain

one node per zone, which means that no coordinate ambiguity exists.

Route computation: Another important criterion we use in our experimental evaluation is how well route discovery performs over the resulting virtual coordinate system when compared to using real coordinates. To evaluate routing performance, we consider the rate of successful route discovery. We ran our routing experiments as follows. For every simulation run, we picked 1,000 random source-destination pairs and performed simple greedy route computation. In other words, the next hop decision is solely based on the positions of the node and its neighbors and tries to select as next hop the closest neighbor to destination. It cannot, however, guarantee route discovery due to local minima situations where no neighbor is closer to the destination than the node where the route ends. In such a situation, the route computation procedure is considered as failed.

C. Results

In this section, we present results from our simulation experiments. Every data point is obtained as the average over fifty simulation runs.² The reader is referred to [18] for all our simulation results.

Average Zone Size: Figure 7 shows the average zone size as a function of number of landmarks for the different strategies. We can observe that the shape of the curves is similar irrespective of the strategy, showing that as the number of landmarks increases, the benefits of placing landmarks at the boundary of the topology (equally spaced or randomly) decrease. For this particular experiment, for example, while there are clear performance differences between the three strategies for five or less landmarks, the average zone size does not change significantly when seven or more landmarks are used even under different placement strategies. This observation remains valid for both uniform and non-uniform topologies.

Note that the only exception appears in the case of the topology with 1,000 nodes using only 2% of concentration points. This is due to the fact that the topology is very sparse and nodes may not be connected to all the landmarks in all the simulations.

Maximum Zone Size: VCAP [16] proposes the combination of position-based and proactive routing. Indeed, VCAP generates zones with size of up to two radio ranges. Therefore, a packet can reach a node 2-hops distant from the intended destination. Adding 2-hop neighborhood knowledge is then required so that, when a

node receives a message intended to another node with the same virtual coordinates, it uses proactive routing within the 2-hop neighborhood to forward the packet to its intended destination. Thus, the maximum zone size is an important metric, since it determines what kind (and how expensive) of proactive forwarding method must be used in addition to the position-based one.

In Figure 8, we show the maximum zone size (in radio coverage units) as a function of the number of landmarks and their placement strategies. We observe that, confirming our theoretical analysis, placing landmarks on the boundary results in smaller maximum zones, independent of the number of landmarks, number of nodes, or node distribution. For instance, lower numbers of landmarks randomly placed generate zones of up to ten radio range units. This requires a 10-hop proactive routing protocol, which will be extremely expensive in terms of overhead. As before, the difference between landmark placement strategies, however, becomes less significant when topologies are more uniform and the number of landmarks increases.

We should point out that the results reported in Figure 8 are different than the results presented in JUMPS [15]. The reason for this difference is that, as noted earlier, here we also consider disconnected networks. In JUMPS, before obtaining a coordinate, a node considers itself positioned ∞ hops from the respective landmark. Consider two nodes placed far from each other with no direct neighbors, in a three landmarks coordinate system. These two nodes are not connected to any landmark, thus do not obtain any coordinates. Both will have (∞, ∞, ∞) as virtual coordinates. In our simulations, we consider those nodes as belonging to the same zone. The distance between them is then taken into account to measure the average and maximum zone sizes. Note that these measurements would be reduced if such nodes were not considered.

Number of nodes per zone: A single zone for the whole topology is the worst possible case one can obtain – it means that all nodes have the same coordinates. On the other hand, the ideal case is when there are as many zones as nodes. Thus, the lower the number of nodes per zone, the more accurate the coordinate system.

We show in Figure 9 the average number of nodes per zone. We observe that the difference between the strategies becomes less important when the number of landmarks increases. This agrees with the trend shown by Figures 7 and 8.

Route computation: Figure 10 shows that different landmark placement strategies have significant impact on routing performance. We observe that placing landmarks on the boundary yields the best results, especially when

²Because the confidence interval is negligible, compared to the average value, we do not represent it on these figures.

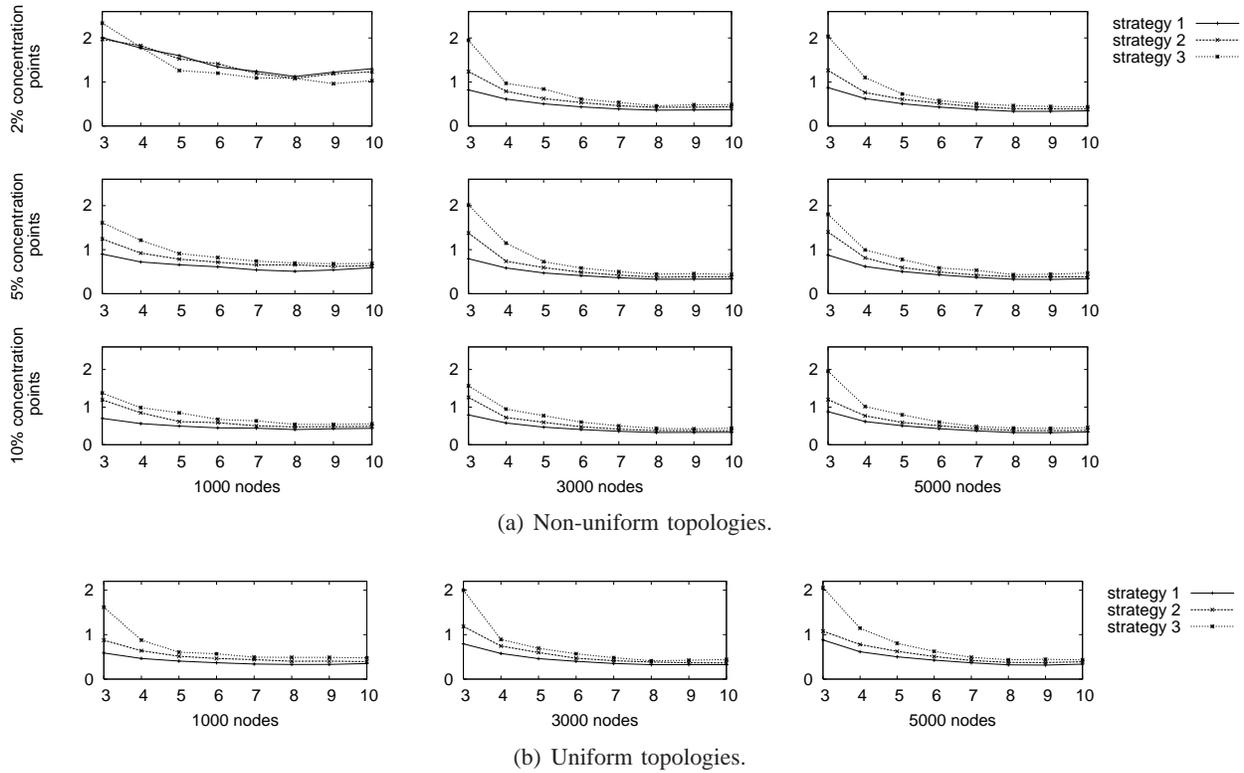


Fig. 7. Average zone size in radio range units (y axis) as a function of number of landmarks (x axis) for different landmark placement strategies.

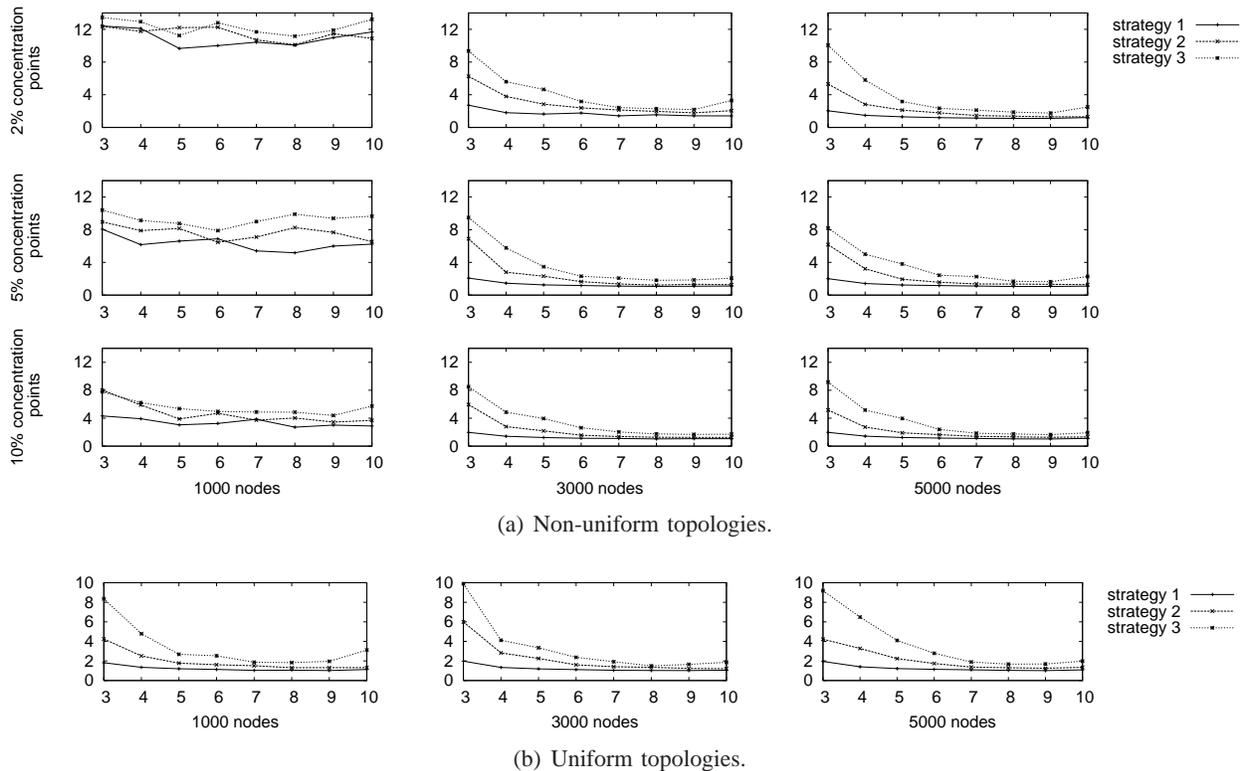


Fig. 8. Maximum zone size (y axis) in radio range units as a function of number of landmarks (x axis) for different scenarios.

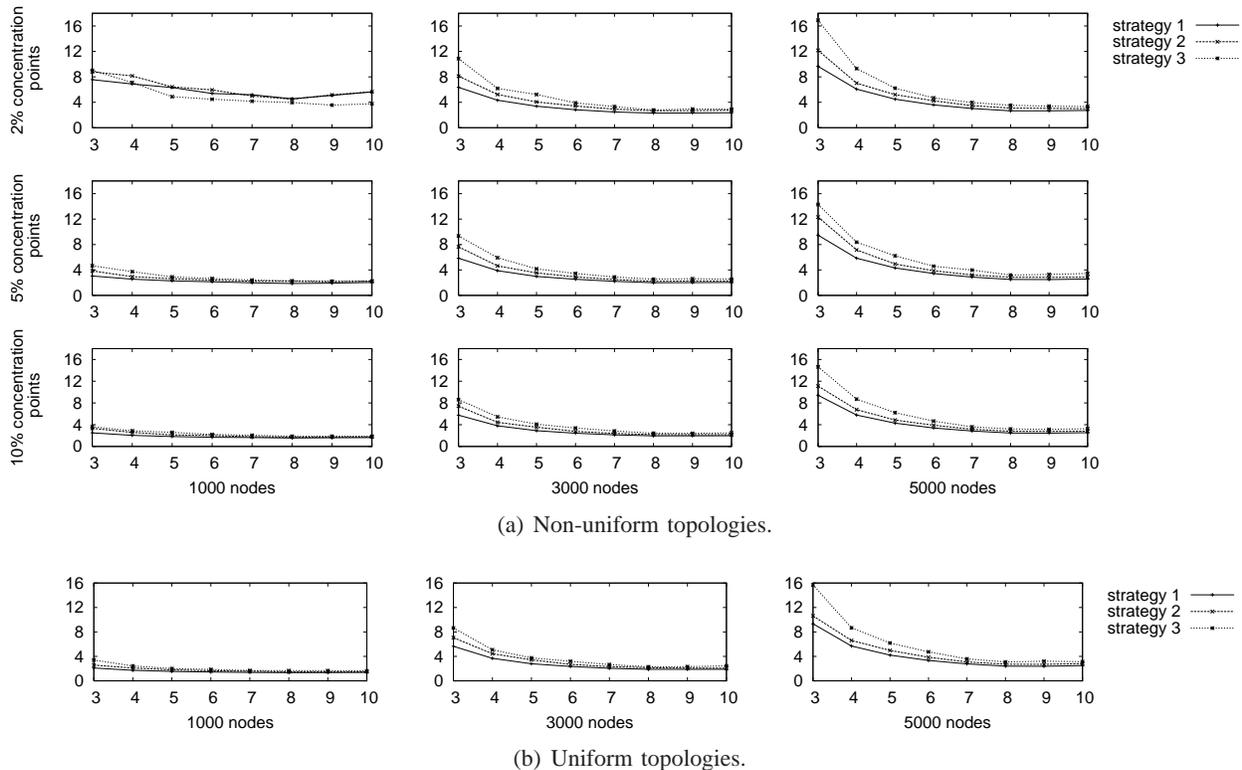


Fig. 9. Average number of nodes per zone (y axis), as a function of number of landmarks (x axis).

they are at equal distances from one another.

This behavior is closely related to the number of nodes per zone represented in Figure 9. Indeed, when a node receives a packet to forward, it chooses, depending on the virtual coordinates, which neighbor is the more appropriate to be the next hop. If two nodes or more share the same coordinates, the forwarding node chooses one of them randomly. If the average number of nodes among a zone is high, then the probability of choosing the right next hop is lower. Thus, routing is more efficient in scenarios where the average number of nodes sharing the same coordinates is lower.

Routing over coordinates obtained using **UniBound** or **RandBound** landmark placement, however, leads to similar performance when compared to routing over real coordinates, provided that sufficient landmarks are employed. This is an important observation as it shows that **RandBound**, i.e., placing landmarks (randomly) on the periphery, is enough to achieve adequate routing performance, avoiding the need of equally distant landmark placement.

We also notice again that as the number of landmarks increases up to a certain threshold, considerable performance gains are achieved. However, beyond the threshold, the gains are not very significant. For the scenarios we ran, seven landmarks seems to be the

threshold for achieving adequate packet delivery.

D. Discussion

In this section, we highlight the insights provided by our experimental study on how landmark placement affects the performance of topology-based self-localization systems.

First, the experimental results we obtained verify our mathematical analysis and show that, indeed, placing the landmarks on the topology boundary, according to the **UniBound** or **RandBound** strategies improves the performance of the coordinate system when compared to **Rand**. However, our simulation study provides us with insight on the performance trends for different types of topologies, at different scales and node densities. For instance, we confirm the results obtained in JUMPS [15], showing that increasing the number of landmarks increases the accuracy of the underlying coordinate system. However, we go beyond that result and show that, if enough landmarks are used, random landmark placement yields comparative accuracy to placing landmarks on the topology boundary (equally spaced or randomly). This is an important result for energy-constrained network designers, planners, and providers, since boundary placement can be prohibitively resource consuming.

We also evaluate the performance of routing over

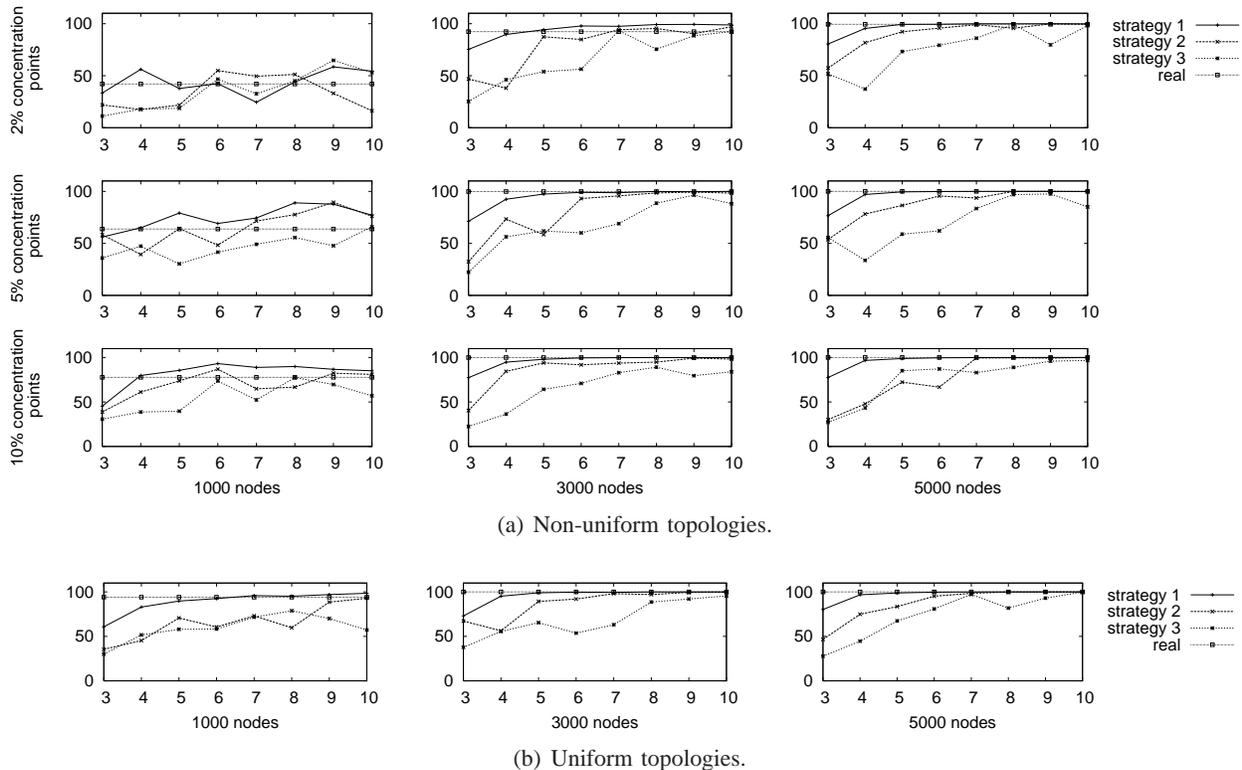


Fig. 10. Route computation success rate (y axis) as a function of number of landmarks (x axis). Here we observe that the landmark placement strategy used is not negligible, but becomes less important as the number of landmarks increases.

the resulting topology-based positioning system against routing using real coordinates and show a similar trend, i.e., that the benefits of boundary placement decreases as the number of landmark increases. We should point out that, except for sparse topologies where a large number of nodes are disconnected, these trends hold for both uniform- and non-uniform topologies.

Another interesting, but not surprising result, is the “diminishing returns” behavior we observed in all our experiments. In other words, our results show that as the number of landmarks increases up to a certain threshold more significant performance improvements can be observed. However, beyond that point, the curve “flattens out,” i.e., the gains of adding more landmarks decrease as the number of landmarks increases. As part of future work, we plan to analyze this behavior analytically.

VI. CONCLUSION

In this paper, we have tackled the problem of landmark placement for hop-count based positioning systems. While previous studies choose as landmarks nodes consistently distributed on the boundary of the topology, we show here that such a criterion does not necessarily yield sufficient performance benefits that warrant its cost. Our mathematical analysis, confirmed by our simulation results, shows that, indeed, placing landmarks on the

topology boundary increases the accuracy of the resulting coordinate system. Furthermore, under some conditions, we also show that uniform landmark placement is optimal. However, extensive simulations using different types of topologies with varying node densities and number of landmarks show that these performance benefits (including packet delivery ratio achieved by greedy routing) decrease as the number of landmarks increases. This means that if enough landmarks are deployed, random landmark placement, which is considerably less resource consuming, yields comparative performance to boundary placement. We also show that, after a certain threshold, additional landmarks provide increasingly less performance gains.

As directions for future work, we plan to prove the optimality of uniform boundary placement under general conditions, including non-uniform topologies. We also plan to investigate the explicit characterization of the threshold beyond which random and uniform deployments have comparable performances and to formally analyze the “diminishing returns” performance trend observed in the simulations. Additionally, with the insight gained in this work, we plan to propose mechanisms that dynamically determine the number of landmarks needed to obtain the most accurate coordinate system. These mechanisms should also be able to identify, given a

certain node distribution, the optimal landmark locations.

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
- [2] M. Dias de Amorim, M. L. Sichitiu, F. Benbadis, Y. Viniotis, and S. Fdida, "Dissecting the Routing Architecture of Self-Organizing Networks," in *IEEE Wireless Communications*, vol. 13, no. 6, 2006, pp. 98–104.
- [3] M. Mauve, J. Widmer, and H. Hartenstein, "A survey on position-based routing in mobile ad hoc networks," *IEEE Network Magazine*, vol. 15, no. 6, pp. 30–39, Nov. 2001.
- [4] S. Giordano, I. Stojmenovic, and L. Blazevie, "Position based routing algorithms for ad hoc networks: a taxonomy," 2001. [Online]. Available: citeseer.ifi.unizh.ch/giordano01position.html
- [5] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System: Theory and Practice, Fourth Edition*. New York: Springer-Verlag, 1997.
- [6] "Glonass," <http://www.glonass-center.ru/>.
- [7] CNES, ESA, "The galileo project," http://www.cnes.fr/html/_455_471_1679_.php.
- [8] J. Aspnes, T. Eren, D. K. Goldenberg, A. S. Morse, W. Whiteley, Y. R. Yang, B. D. O. Anderson, and P. N. Belhumeur, "A theory of network localization," *IEEE Transactions on Mobile Computing*, vol. 5, no. 12, pp. 1663–1678, 2006.
- [9] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005. [Online]. Available: <http://wsl.stanford.edu/andrea/wireless/>
- [10] F. Gustafsson and F. Gunnarsson, "Positioning using time-difference of arrival measurements," in *Proceedings of ICASSP*, Hongkong, Apr 2003.
- [11] D. Niculescu and B. Nath, "Ad hoc positioning system APS using AoA," in *Proceedings of IEEE INFOCOM'03*, San Francisco, CA, USA, 2003.
- [12] A. Rao, S. Ratnasamy, C. Papadimitriou, S. Shenker, and I. Stoica, "Geographic routing without location information," in *Proceedings of ACM MOBICOM'03*, San Diego, CA, USA, 2003.
- [13] A. Jadbabaie, "On geographic routing without location information," in *Proceedings of the IEEE Conference on Decision and Control*, Paradise Island, Bahamas, 2004, pp. 4764–4769.
- [14] F. Benbadis, T. Friedman, M. Dias de Amorim, and S. Fdida, "GPS-Free-Free positioning system for sensor networks," in *Proceedings of WOCN 2005 Conference*, Dubai, UAE, 2005.
- [15] F. Benbadis, J.-J. Puig, T. Friedman, M. D. de Amorim, D. Symplot-Ryl, and S. Fdida, "JumPS: Enhancing hop-count positioning in sensor networks using multiple coordinates," LIP6, Tech. Rep., 2006, <http://rp.lip6.fr/~benbadis/jumps/>.
- [16] A. Caruso, S. Chessa, S. De, and A. Urpi, "GPS-Free Coordinate Assignment and Routing in Wireless Sensor Networks," in *Proceedings of IEEE INFOCOM'05*, Miami, FL, USA, 2005.
- [17] R. Fonseca, S. Ratnasamy, D. Culler, S. Shenker, and I. Stoica, "Beacon vector routing: Scalable point-to-point in wireless sensor networks," in *Proceedings of NSDI'05*, Boston, MA, 2005.
- [18] F. Benbadis, K. Obraczka, J. Cortés, and A. Brandwajn, "Exploring landmark placement strategies for self-localization in wireless sensor networks," 2006, http://rp.lip6.fr/~benbadis/landmarks_placement/.