GPS-less Low Cost Outdoor Localization For Very Small Devices
Nirupama Bulusu, John Heidemann, Deborah Estrin
USC/Information Sciences Institute
bulusu, johnh, estrin@isi.edu

Abstract— Instrumenting the physical world through large networks of wireless sensor nodes, particularly for applications like marine biology, requires that these nodes be very small, light, un-tethered and unobtrusive, imposing substantial restrictions on the amount of additional hardware that can be placed at each node. Practical considerations such as the small size, form factor, cost and power constraints of nodes preclude the use of GPS (Global Positioning System) for all nodes in these networks. The problem of localization, i.e., determining where a given node is physically located in a network is a challenging one, and yet extremely crucial for many applications of very large device networks. It needs to be solved in the absence of GPS on all the nodes in outdoor environments. In this paper, we propose a simple connectivity-metric based method for localization in outdoor environments that makes use of the inherent radio-frequency (RF) communications capabilities of these devices. A fixed number of reference points in the network transmit periodic beacon signals. Nodes use a simple connectivity metric to infer proximity to a given subset of these reference points and then localize themselves to the centroid of the latter. The accuracy of localization is then dependent on the separation distance between two adjacent reference points and the transmission range of these reference points. Initial experimental results show that the accuracy for 90% of our data points is within one-third of the separation distance.

Keywords—localization, radio, wireless, GPS-less, connectivity, sensor networks.

I. INTRODUCTION

Wireless networks of sensors greatly extend our ability to monitor and control the physical environment from remote locations. Networked sensors can collaborate, aggregate the huge amount of sensed data to provide a rich, multi-dimensional view of the physical environment. However, instrumenting the physical world, particularly for applications such as marine biology, requires that the devices we use as sensor nodes be small, light, unobtrusive and un-tethered. This imposes substantial restrictions on the amount of hardware that can be placed on these devices.

In these large sensor network systems, however, we need nodes to be able to locate themselves in various environments, and on different distance scales. This problem, which we refer to as localization, is a challenging one, and yet extremely crucial for many applications of very large networks of devices.

GPS [8] solves the problem of localization in outdoor environments for PC class nodes. However, for large networks of very small, cheap and low power devices, practical considerations such as size, form factor, cost and power constraints of the nodes preclude the use of GPS on all nodes.

In this paper, we address the problem of localization for such devices, under the following constraints.

- **RF-based:** We assume small nodes will be connected with some kind of short-range radio. Our primary goal is to leverage this radio for localization, thereby eliminating the cost, power and size requirements of a GPS receiver.
- **Ad hoc:** In addition, we desire a solution that does not require pre-planning or extensive infrastructure.
- **Responsiveness:** We need to be able to localize within a fairly low response time.
- **Cost:** We desire to minimize computation and message costs to reduce power consumption.
- **Receiver-based:** In order to scale well to large distributed networks, the responsibility for localization must lie with the receiver node that needs to be localized and not with the reference points.
- **Adaptive Fidelity:** In addition, we wanted our algorithms to be adaptable to the granularity of available reference points.

This paper proposes an idealized radio model and a simple connectivity based localization method for such devices in unconstrained outdoor environments, that

---

1 We borrow the term localization from robotics, where it refers to the problem of determining the position of a mobile robot in some coordinate system.
makes use of the inherent radio-frequency RF communications capabilities of these devices. A fixed number of nodes in the network serve as reference points and transmit periodic beacon signals. Nodes use a simple connectivity metric to infer proximity to a given subset of these reference points and then localize themselves to the centroid of the selected(proximate) reference points.

The paper makes the following key contributions.

- It presents a detailed and systematic exploration and classification of the design space and work done in the area of localization.
- It proposes a method for coarse-grained localization based on an idealized radio model and demonstrates its validity and applicability in outdoor unconstrained environments.
- It describes a simple implementation of the model and presents initial results.

II. RELATED WORK

Essentially any method for localization relies on two features, reference points (either fixed or moving but whose positions are known at any given instant) and communication (unidirectional or bidirectional, single or multiple modalities of communication) between the reference points and the node to be localized.

We classify the approaches in two categories, based on the information inferred during the communication. Approaches which infer fine grained information such as distance to a reference point based on signal strength or timing measurements fall into the category of fine grained localization and those that infer coarse grained information such as proximity or the presence of connectivity to a given reference point are categorized as coarse grained localization methods.

We discuss the techniques and work that has been done so far in each of these categories, and classify them further.

A. Fine-grained localization

Fine-grained localization methods can be further classified into range-finding and directionality based methods, depending on whether ranges or angles to reference points are being inferred.

A.1 Range-finding

The range(distance) of the mobile node to several reference points is determined by one of the several techniques enumerated below. The position of the mobile node can be computed based on trilateration or multilateration.

- **Timing:** The distances from the mobile host to the reference points can be inferred from the times-of-flight of their respective communication signals.

The time-of-flight, again, may be calculated using the timing advance technique which measures the amount that the timing of the mobile has to be advanced in order for the received signal to fit into the correct time slot. This technique is used in GPS [8] and Pinpoint’s Local Positioning System(LPS) [7]. GPS measures one-way flight time whereas LPS measures round-trip-time (thereby eliminating the need for time synchronization).

GPS [8] is a wide-area radio positioning system. In GPS each satellite transmits a unique code, a copy of which is created in real time in the user-set receiver by the internal electronics. The receiver then gradually time shifts its internal clock till it corresponds to the received code: an event called lock-on. Once locked-on to satellite, the receiver can determine the exact timing of the received signal in reference to its own internal clock. If that clock were perfectly synchronized with the satellite’s atomic clocks, the distance to each satellite could be determined by subtracting a known transmission time from the calculated receive time. In real GPS receivers, the internal clock is not quite accurate enough. An inaccuracy of a mere microsecond corresponds to a 300-meter error.

Pinpoint’s 3D-iD system [7] is a Local Positioning System(LPS) that covers an entire three-dimensional indoor space and is capable of determining the 3-D location of items within that space. The LPS subdivides the interior of the building into cell areas that vary in size with the desired level of coverage. The cells are each handled by a cell controller which is attached by a coaxial cable to up to 16 antennas. It provides an accuracy of 10 meters for most indoor applications, though some may require accuracy of 2 meters. The main drawback of this system is that it is centralized, and requires a lot of infrastructural set up effort.
Alternately, the time of flight can be calculated by making explicit time-of-arrival measurements based on two distinct modalities of communication, ultrasound and radio, as in the Active Bat [6]. These two different modalities travel at vastly different speeds (350 m/s and 3 × 10^{-8} m/s), enabling the radio signal to be used for synchronization between the transmitter and the receiver, and the ultrasound signal to be used for ranging. The Active Bat system however requires significant effort for deployment indoors. Ultrasound systems cannot be used outdoors because they all use a single transmission frequency (40 kHz) and hence there is a high probability of interference from other ultrasound sources.

- **Signal Strength:** In their recent paper [5], Bahl et al., suggest estimating distance based on signal propagation characteristics in indoor environments. They compute distance from measured signal strength by applying a Wall Attenuation Factor (WAF) based signal propagation model. Their approach is effective indoors, unlike ours, but it requires extensive pre-planning, making it unsuitable for rapid or ad hoc deployment.

### A.2 Directionality

Another way of estimating location is to compute the angle of each of the reference points with respect to the mobile node in some reference frame. The position of the mobile node can then be computed using triangulation methods.

One such technique used in cellular networks, is known as small aperture direction finding. It requires a complex antenna array at each of the cell site locations. The antenna arrays can in principle work together to determine the angle (relative to the cell site) from which the cellular signal originated. When several cell sites can determine their respective angles of arrival, the cell phone location can be estimated from the intersection of projected lines drawn out from the cell site at the angle corresponding to the signals origin. There are two drawbacks of this approach. The cost of the complex antenna array implies that it can be placed only at the cell sites. Secondly the cell sites are responsible for determining the location of the mobile node which will not scale well when we have a large number of such nodes. This approach too, cannot be used in indoor environments.

Another example of directionality based systems are the VOR/VORTAC stations [10], which were used for long distance aviation navigation prior to GPS. The VOR station transmits a unique omnidirectional signal that allows an aircraft aloft to determine its bearing relative to the VOR station. The VOR signal is electrically phased so that the received signal is different in various parts of the 360 degree circle. By determining which of the 360 different radials it is receiving, the aircraft can determine the direction of each VOR station relative to its current position.

### B. Coarse Grained Localization

Our work is perhaps most similar to earlier work done in coarse-grained localization using Infra Red (IR) technology.

In the *Active Badge* system [1] system, a badge worn by a person emits a unique IR signal every 10 seconds. Sensors are placed at fixed positions within a building and as they receive the unique identifiers, the location manager software is able to provide information about the person’s location to the requesting services and applications. While the performance of this system is quite good, a major drawback is that the range of the IR system is fairly small, and consequently the building has to be wired up with a significant number of sensors. In the few places where such systems have been deployed, sensors have been physically wired in every room of the building. Such a system scales poorly, and incurs significant installation, configuration and maintenance cost.

Another system that is based on IR technology is described in [3]. This system requires IR transmitters to be located at fixed positions inside the ceiling of the building. An optical sensor sitting on a head mounted unit senses the IR beacons and system software determines the position of the person. This system suffers form similar drawbacks as the Active Badge system.

IR tends to perform poorly in the presence of direct sunlight and hence cannot be used outdoors.
III. IDEALIZED RADIO MODEL AND LOCALIZATION ALGORITHM

A. Idealized Radio Model

We have found an idealized radio model useful for predicting bounds on the quality of localization. This section presents this idealized model. To our surprise, this model compares quite well to outdoor radio propagation as we explore in Section IV.

We make two assumptions in our idealized model:
• Perfect spherical radio propagation.
• Identical transmission range(power) for all radios.

B. Localization Algorithm

Multiple nodes in the network serve as reference points(named $R_1$ to $R_n$). They are situated at known positions($X_1, Y_1$)to($X_n, Y_n$), that form a regular mesh and transmit periodic beacon signals(period = $T$) containing their respective positions. We assume that the reference points can be synchronized so that their beacon signal transmissions of neighboring reference points do not overlap in time. Furthermore, in any time interval $T$, each of the reference points would have transmitted exactly one beacon signal.

First, we define a few terms.

$d$ Separation distance between adjacent reference points
$R$ Transmission range of the reference point
$T$ Time interval between two successive beacon signals transmitted by a reference point
$t$ Receiver sampling or data collection time
$Nsent(i, t)$ Number of beacons that have been sent by $R_i$ in time $t$
$Nrecv(i, t)$ Number of beacons sent by $R_i$ that have been received in time $t$
$CM$ Connectivity metric
$S$ Sample size for connectivity metric
$CMthresh$ Threshold for $CM$
$(Xest, Yest)$ Estimated Location of the receiver
$(Xa, Ya)$ Actual Location of the receiver

Each mobile node listens for a fixed time period $t$ and collects all the beacon signals that it receives from various reference points.

We characterize the information per reference point $R_i$ by a connectivity metric($CM_i$), defined as

$$CM_i = \frac{Nrecv(i)}{Nsent(i)} \times 100$$

In order to improve the reliability of our connectivity metric, we would like to base our metric on a sample of at least $S$ packets, where $S$ is the sample size, a tunable parameter of our method (i.e., $Nsent(i) = S$). Since we know $T$ to be the time period between two successive beacon signal transmissions, we can set $t$, the receiver’s sampling time as:

$$t = (S + 1 - \epsilon)T \quad (0 < \epsilon \ll 1)$$

From the beacon signals that it receives, the receiver node infers connectivity to a collection of reference points for which the respective connectivity metric exceeds a certain threshold, $CMthresh$say 90\%). We denote the collection of reference points by $R_{i1}, R_{i2}, \ldots, R_{ik}$. The receiver localizes itself to the region which coincides to the intersection of the connectivity regions of this set of reference points, which is defined by the centroid of these reference points.

$$(Xest, Yest) = \left( \frac{X_{i1} + \cdots + X_{ik}}{k}, \frac{Y_{i1} + \cdots + Y_{ik}}{k} \right)$$

We characterize the accuracy of the estimate by the error distance $ED$ defined as,

$$ED = \sqrt{(Xest - Xa)^2 + (Yest - Ya)^2}$$

By increasing the density of beacons that populate the grid (i.e increasing $\frac{R}{d}$, the granularity of the localization regions becomes finer, and hence the accuracy of the location estimate improves. This is illustrated in figure 1.

IV. VALIDATION

Since our localization model is dependent on the spherical radio propagation assumption, described in the previous section; we checked the validity of our assumption in both outdoor and indoor environments.

• Outdoors: In an empty parking lot; we did a systematic traversal of about 121 grid points in a 10m*10m grid with the radio transmitter placed at $(0,0)$. There were no holes or non-linearities in the transmission. We calculate the range sample as the distance between the transmitter $(0,0)$ and the farthest point from it along each X ISO-line(points with same X coordinate value)
whose \( CM \) exceeds \( CM_{thresh} \), yielding 11 range samples in all. The maximum variance in the range with \( CM_{thresh} = 90\% \) was 2 meters, as can be seen in figure 2. The median range was 8.94m, which is used to plot the theoretical range in the graph. Among the 78 data points that appear in the graph, there is a mismatch at 10 points and match at 68 points between theory and experimental values, i.e., there is an 87\% correlation between the two.

- **Indoors**: We conducted a similar experiment indoors. The range with a 90\% connectivity metric varied from about 73 feet (in a corridor with direct Line Of Sight) to about less than 15 feet (with walls in between).

Hence the idealized radio model may be considered valid for outdoor unconstrained environments only.

## V. EXPERIMENTAL RESULTS

### A. Experimental Testbed

Our experimental testbed consisted of 4 reference points, located at the corner of a 10m*10m grid. This grid was further subdivided into 100 1m*1m smaller grids and we collected data at each of the 121 small grid corners.

The transmitters at each reference point and the receiver are the Radiometrix-RPC 418 (radio packet controller) modules connected to Toshiba Librettos running RedHat Linux 6.0. A 3 inch antenna is used for the experimental purposes. We used about 5 RPC modules in all (including 1 for the receiver).

#### A.1 Software

The software was written for the Radiometrix RPC-418 modules and consists of two components.

1. **Beacon**: The reference point periodically transmits a packet (every 2 seconds in our experiment) containing its ID and position.

2. **Receiver**: The receiver obtains its current measured position based on an input from the user. For each measured position, it samples for a time period \( t \) determined by the sample size \( S \), and logs the set of reference points it hears from and its current localization estimate.

There are several implementation issues that our software did not address, since we were interested in an empirical study rather than actual deployment. These issues are mentioned below, and discussed more fully in Section VI.

- **Collision Avoidance**: Multiple reference points must transmit their beacon signals periodically without collision or with minimum collision. This is acutely tied to time synchronization. Currently we start the transmission at the reference points in a linear sequence to achieve this, which will clearly not scale when we have larger numbers of reference points.

- **Power Consumption**: Packet reception is expensive power wise and we want our receivers to ideally sample
B. Results

In this section, we discuss the experimental results based on our implementation. Our experimental parameters are $T = 2$ seconds, $S = 20$, $t = 41.9$ seconds.\footnote{Although our experimental parameter values for $T$, and hence $t$ are high, we can substantially scale them down without violating the integrity of the experiment.}

In figure 3, we see the areas of connectivity of the 4 reference points in the grid. We collected data at 121 points in all, wherein adjacent points are separated by 1m each. We see several distinct regions in the grid, based on the areas of overlap. Each distinct region constitutes an equivalence class, defined by the centroid of the reference points in the region. These can be contrasted with the theoretically predicted overlap regions, seen in figure 4.

The location estimate at each grid point is the centroid. We use the error-distance metric defined in Section III to characterize the performance.

In figure 5, the error-distance is plotted as a function of the position. Clearly, the error-distance is lowest at the the position corresponding to the centroid of the region and increases towards the edges of the region. The average error-distance was 1.83m and the standard deviation was 1.07m. The minimum error was 0m and the maximum error was 4.12 across 121 error-distance samples.

Figure 6 shows the cumulative error-distance distribution across all the grid points. For over 90% of the data points, the error-distance falls within 3.0 meters i.e within 30% of the separation-distance between two adjacent reference points. This result is based on 4 reference points only. Since we observed a high correlation between our model and experiment, improved granularity can be expected with a higher density of reference points.

Figure 7 presents a simulation based scaling result of the error-distance behavior. We present it to predict how the granularity of localization can be expected to vary when the density of reference points is increased.

In our simulation, we assume an infinite two-dimensional mesh of reference points, with any two adjacent reference points spaced a distance $d$ apart and transmission range $R$. Our coordinate system is centered at one such reference point, which is assumed to be at $(0,0)$.

The localization estimate of any point $(X,Y)$ in the mesh can be obtained in two steps.

**Step 1:** Determine all the reference points which are

![Fig. 3. Experimental 90% connectivity ranges for the 4 reference points](image1)

![Fig. 4. Theoretical 90% connectivity ranges for the 4 reference points](image2)
within range $R$ of $(X,Y)$, by considering all the reference points between $(X - R, Y - R)$ and $(X + R, Y + R)$.

Step 2: Localize $(X,Y)$ to the centroid of the selected reference points and compute the corresponding error-distance.

For a given $d$, we increase $R/d$ from 1 to 4. We consider the average and maximum error-distances of the localization estimates for 10201 uniformly spaced points within one grid in the mesh, for each $R/d$ value. Although the maximum and average error do not decrease monotonically, non-trivial increments to $R/d$, (for instance, an increment of 1) lead to lower maximum and average error-distances on the whole.

In particular, the maximum error-distance experiences a substantial drop (from 0.5$d$ to 0.25$d$) when the density $R/d$ is increased from 1 to 4.

VI. DISCUSSION AND FUTURE WORK

In this section, we will discuss some general problems that arise in deploying our localization method and present some of our ideas on solving them.

- Collision Avoidance or Reference Point Synchronization

In order for our method to work well, reference points must be synchronized so that their beacon signal transmissions do not overlap in time. This synchronization needs to be local and not necessarily global. This synchronization can be achieved by using some kind of a randomized back-off scheme, wherein the time interval $T$ is subdivided into several smaller slots and each reference point randomly picks a slot to transmit or by applying a hash function on its node (reference point) ID, if available. Of the two alternatives, the former seems more attractive since it eliminates the need for node IDs.

- Reference Point Configuration

We have left the issue of how the reference point coordinates are configured and how they are deployed open. This could be achieved through limited human intervention. The reference points themselves could use GPS since they do not have the same constraints as other nodes. If the number of reference points is too large, then we need to develop some self-configuration schemes in future work.

- Placement Heterogeneity of Reference points

Our localization method assumes that the reference points are placed at the intersections of a regular mesh. We controlled the placement, to bound the quality of localization. However it may not always be feasible to control the reference point placement with such uniformity. Preliminary analysis with a 1D model (linear placement of reference points with non-uniform distances of separation) seems to suggest that it does not impact the localization accuracy too adversely,
provided the maximum degree of non-uniformity is bounded. This is an area for future work.

- **Robustness**
  Since the success of our localization method depends on the node reliably inferring connectivity, and hence proximity to its neighbouring reference points, it must be tolerant of reference point failures. Reference points should monitor themselves and fail-stop when their battery power drops down. Some amount of redundancy (additional nodes that can serve as reference points, if need be) should be incorporated into the system to tolerate reference point failures.

- **Efficiency and Parameter Tuning**
  In order to avoid collisions, we need $T$ to be high. In order to ensure the consistency of our connectivity metric, we need $S$ to be high. However, in order to reduce power consumption at the receiver, we need to reduce $t$ i.e., we would like to use smaller values of $S$ and $T$. Since, we use the connectivity metric as a coarse-grained measure, our experience seems to suggest that a small value of $S$, such as 10 would suffice. The value of $T$ would be determined by the reference point density $R/d$ and the efficacy of the collision avoidance scheme.

VII. CONCLUDING REMARKS

This paper addresses localization in unconstrained, outdoor environments for networks of low-cost, very small devices where GPS is not available on all nodes. We suggested a connectivity-metric based localization method based on an idealized radio model where the receiver localizes itself with high confidence to the centroid of a set of reference points. In outdoor environments, our model correlated very well with reality (87%).

Our approach is simple, adaptive to the granularity of reference points available and lends itself easily to a distributed implementation, and is hence scalable to large, distributed networks of devices. Initial experimentation has shown promising results, with our simple scheme, for a small number of reference points. Simulation suggests that the granularity of error-distance can be improved by further increasing the density of reference points.

We also outlined some general problems which need to be tackled for large scale deployment and presented our ideas for solving them.

ACKNOWLEDGMENTS

The authors would like to thank Lewis Girod and Jeremy Elson for their suggestions and feedback.

REFERENCES


