

Course “Algorithmic Foundations of Sensor Networks”

Lecture 9: Localization in WSNs

Sotiris Nikolettseas

Department of Computer Engineering and Informatics
University of Patras, Greece

Spring Semester 2020-2021

Summary of this Lecture

- Overview (importance, key issues)
- Coarse-grained localization approaches
 - Binary proximity
 - Centroid calculation
 - Geometric constraints
 - APIT
- Fine-grained localization approaches
 - triangulation using distance estimates
 - radio signal strength measurements (RSS)
 - time difference of arrival (TDoA)
 - angle of arrival (AoA)
 - pattern matching
 - RF sequence decoding
- Network-wide localization

Localization definition

- WSNs are intended to provide information about the spatio-temporal characteristics of the observed physical world.
- Each sensor observation can be characterized as a tuple of the form $\langle S, T, M \rangle$ where:
 - S is the spatial location of the measurement
 - T the time of the measurement
 - M the measurement itself

Why do we need localization?

The location information of nodes is fundamental in order to:

- To provide location stamps of sensed data
- To track/localize moving objects and entities
- To monitor the spatial evolution of various phenomena
- To achieve load balancing
- To form clusters
- To facilitate routing
- To perform efficient spatial querying

The issues and localization techniques can be classified on the basis of a number of key questions:

- What to localize? (reference/anchor, mobile cooperative nodes)
- When to localize? (one-shot, on-the-fly, periodically, etc.)
- How well to localize? (absolute or relative coordinates, accuracy)
- Where to localize? (centralized, distributed)
- How to localize? (different methods used)

Localization approaches

- 1 Coarse-grained localization using minimal information
- 2 Fine-grained localization using detailed information

Trade-off between simplicity, hardware costs, resource consumption and accuracy.

A. Coarse-grained localization - Binary proximity

- Simple decision of whether two nodes are within reception range of each other.
- A set of reference nodes are placed in a non-overlapping manner.
- The reference nodes periodically broadcast beacons which include their location IDs.
- The unknown node must determine to which node it is closest to.

Coarse-grained localization - Centroid calculation

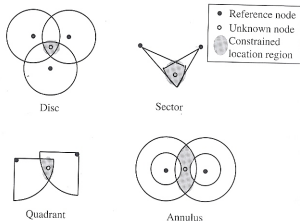
- Consider a two-dimensional scenario.
- Let n reference nodes within the proximity of an unknown node u .
- The location of the i -th reference node is denoted by (x_i, y_i)
- Then, the location of the unknown node is determined as:
 - $x_u = \frac{1}{n} \sum_{i=1}^n x_i$
 - $y_u = \frac{1}{n} \sum_{i=1}^n y_i$

Coarse-grained localization - Geometric constraints

- The region of radio coverage is usually upper-bounded by a circle of radius R_{max} .
- if node B hears A, it knows that it must be no more than a distance R_{max} from A.
- Now, if an unknown node hears from several reference nodes, it can determine that it must lie in the intersection of circles of radius R_{max} centered on these nodes.

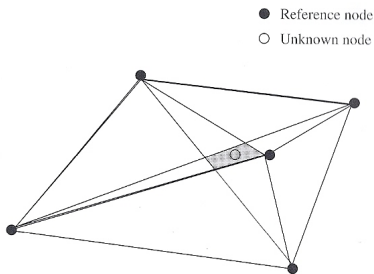
Coarse-grained localization - Geometric constraints

- This can be extended to other scenarios e.g.:
 - Using R_{min} and R_{max} to determine an annulus
 - Using R_{max} and $(\theta_{min}, \theta_{max})$ to determine a cone region



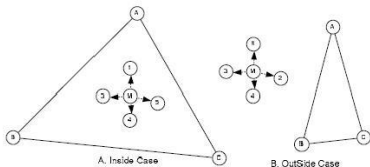
Coarse-grained localization - Approximate point in triangle (APIT)

- It provides location estimates as the centroid of an intersection of regions.
- It uses triangles between different sets of reference nodes in order to localize the unknown node.



Coarse-grained localization - Approximate point in triangle (APIT)

- Each triangle region is defined between different sets of three reference nodes
- If no neighbor of M is farther from/closer to all three anchors A, B and C simultaneously, M assumes that it is inside the triangle.



B. Fine-grained localization

These techniques, based on detailed information, include:

- triangulation using distance estimates
- radio signal strength measurements (RSS)
- time difference of arrival (TDoA)
- angle of arrival (AoA)
- pattern matching
- RF sequence decoding

Fine-grained localization - Radio signal-based distance-estimation (RSS)

- Mean radio signal strengths diminish with distance according to a power law e.g.

$$P_{r,dB}(d) = P_{r,dB}(d_0) - \eta \cdot 10 \cdot \log\left(\frac{d}{d_0}\right) + X_{\sigma,dB}$$

- where $P_{r,dB}(d)$ is the received power at distance d
 - where $P_{r,dB}(d_0)$ is the received power at some reference distance d_0
 - η is the path loss exponent (ranging usually from 2 to 6)
 - $X_{\sigma,dB}$ a log-normal random variable with variance σ^2 that accounts for fading effects
- The fading term often has a large variance and that is why RSS-based techniques may offer low accuracy.

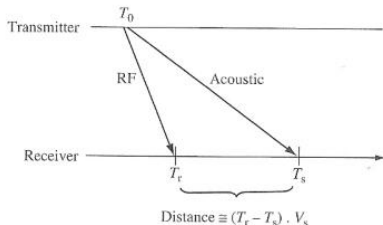
Fine-grained localization - Distance-estimation using time differences (TDoA)

- Limitations of time-of-flight techniques: precision constraints
- Limitations of RSS techniques: fading effects

A more promising technique is the combined use of ultrasound/acoustic and radio signals to estimate distances by determining the TDoA of these signals.

Distance-estimation using time differences (TDoA)

- The idea is to simultaneously transmit both the radio and acousting signals (audible or ultrasound) and measure the time T_r and T_s of the arrival of these signals respectively at the receiver.
- Since the speed of the radio signal is much larger than the speed of the acoustic signal, the distance is then simply estimated as $(T_s - T_r) \cdot V_s$, where V_s is the speed of the acoustic signal.



Distance-estimation using time differences (TDoA)

Limitations of this technique:

- Acoustic ranging requires the nodes to be in fairly close proximity to each other and in line of sight.
- There is uncertainty in the calculation because the speed of sound varies depending on many factors such as altitude, humidity and air temperature.
- Acoustic signals also show multi-path propagation effects that may impact the accuracy of signal detection.
- The sensor nodes must be equipped with acoustic transceivers.

Fine-grained localization - Triangulation using distance estimates

- The location of the unknown node (x_0, y_0) can be determined based on measured distance estimates $\overline{d}_i$ to n reference nodes $(x_1, y_1), \dots, (x_n, y_n)$.
- This can be formulated as a least squares minimization problem.

Least squares method:

- A standard approach to the approximate solution of overdetermined systems, i.e. sets of equations in which there are more equations than unknowns.
- "Least squares" means that the overall solution minimizes the sum of the squares of the errors made in solving every single equation.

Triangulation using distance estimates

- The least squares minimization problem is then to determine the (x_0, y_0) that minimizes

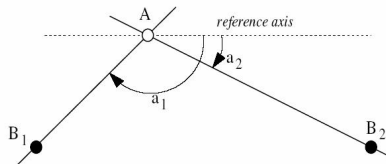
$$\sum_{i=1}^n (\rho_i)^2$$

Fine-grained localization - Angle of arrival (AoA)

- Another possibility for localization is the use of angular estimates instead of distance estimates
- Angles can potentially be estimated by using rotating directional beacons
- A very simple localization technique involves three rotating reference beacons at the boundary of the sensor network providing localization for all interior nodes

Fine-grained localization - Angle of arrival (AoA)

- The angular information can be combined with a good distance estimate to the reference node
- A measures the direction of an incoming link from one anchor and combines it with distance estimation from that anchor or
- By using 2 anchors, A can determine its position without the need of any distance information



- An alternative to measuring distances or angles that is possible in some contexts is to use a **pre-determined map** of signal coverage in different locations of the environment
- Use this map to determine where a particular node is located by performing **pattern matching** on its measurements

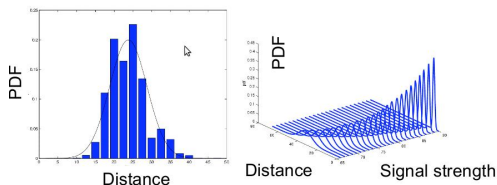
Pattern matching (RADAR)

- This technique requires the prior collection of empirical measurements of signal strength statistics (**training phase**) from different reference transmitters at various locations
- It is also important to take into account the directional orientation of the receiving node
- Once this information is collected, any node in the area is localized by comparing its measurements from these references to determine which location matches the received pattern best

Pattern matching (RADAR)

- + As a pure RF technique has the potential to perform better than the RSS-based distance-estimation and the triangulation approach
- It is very location specific and requires intensive data collection prior to operation
- It may not be useful in settings where the radio characteristics of the environment are highly dynamic

- Received signal strength indicator (RSSI) is a measurement of the power present in a received radio signal.
- Problem: Highly error-prone process
- Shown: PDF for a fixed RSSI



Fine-grained localization - RF sequence decoding (ecolocation)

- This technique uses the relative ordering (wrt Euclidean distance) of received radio signal strengths for different references as the basis for localization.
- It works as follows:
 - 1 The unknown node broadcasts a localization packet
 - 2 Multiple references record their RSSI reading for this packet and report it to a common calculation node
 - 3 The multiple RSSI readings are used to determine the ordered sequence of references from highest to lowest RSSI
 - 4 The region is scanned for the location for which the correct ordering of references has the **best match** to the measured sequence. This is considered the location of the unknown node.

RF sequence decoding (ecolocation)

- In an ideal environment, the measured sequence would be error free
- However in real environments because of multi-path fading effects we have errors
- Some anchors that are closer to the unknown node may show lower RSSI, while others that are farther away may appear earlier in the sequence.

RF sequence decoding (ecolocation)

- The best match is quantified by deriving the $\frac{n \cdot (n-1)}{2}$ pair-wise ordering constraints (e.g. anchor A is closer than B, B is closer than C, etc.) at each location
 - and determining how many of these constraints are satisfied/violated in the measured sequence
 - The location which provides the maximum number of satisfied constraints is the best match
-
- + More accurate compared with other RF-only schemes (e.g. triangulation using distance estimates)
 - + The ordered sequence of RSSI values provides robustness to fluctuations in the absolute RSSI value

RF sequence decoding (ecolocation)

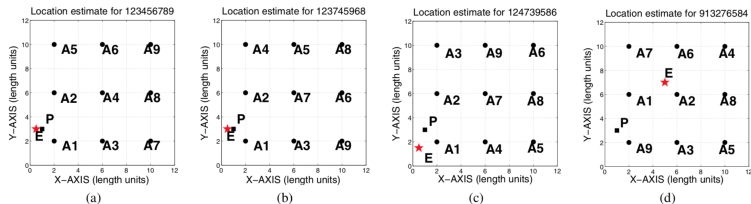


Fig. 3. Ecolocation location estimate (E) for the unknown node (P) at (1, 3) for a grid layout of 9 reference nodes (A). The reference nodes are numbered according to their rank in the ordered sequence. (a) Sequence: 123456789 (no erroneous constraints) [Estimate: (0.5, 3)] (b) Sequence: 123745968 (13.9% erroneous constraints) [Estimate: (0.5, 3)] (c) Sequence: 124739586 (22.2% erroneous constraints) [Estimate: (0.5, 1.5)] (d) Sequence: 913276584 (47.2% erroneous constraints) [Estimate: (5, 7)].

C. Network-wide localization - Issues

- So far, we have focused on the problem of *node localization* for determining the location of a single unknown node given a number of anchor nodes
- A broader problem is the *network localization* where several unknown nodes have to be localized in a network with a few reference nodes

- The performance of localization depends on the resources and information available within the network
- Several scenarios are possible:
 - there may be no reference nodes at all
 - there may be just a single mobile reference
 - the number of reference nodes may vary
 - different kind of information may be available (distance, angle, etc.)

Network-wide localization - Issues

- Some localization approaches are centralized
 - the data is provided to a central node where the solution is computed
 - + this approach may be sufficient in moderate-sized network
- Other localization approaches are distributed

Measuring performance of localization techniques:

- Distribution of location errors
- The mean location error
- The percentage of unknown nodes that can be located within a desired accuracy

Network-wide localization - Constraint-based approaches

Geometric constraints can often be expressed in the form of linear matrix inequalities and linear constraints. This applies to:

- radial constraints (two nodes are determined to be within range R of each other)
- annular constraints (a node is determined to be within ranges $[R_{min}, R_{max}]$ of another)
- angular constraints (a node is determined to be within particular angular sector of another)
- convex constraints

Constraint-based approaches

- Information about a set of reference nodes together with the above constraints describes a feasible set of constraints for a semidefinite program
- By selecting an appropriate objective function for the program, the constraining rectangle which bounds the location for each unknown node can be determined

Semidefinite programming definition:

- Semidefinite programming (SDP) is a subfield of convex optimization concerned with the optimization of a linear objective function over the intersection of the cone of positive semidefinite matrices with an affine space, i.e., a spectrahedron.

Constraint-based approaches

Using bounding rectangles a distributed iterative solution can be used:

- At each step nodes broadcast to their neighbors their current constrained region, which is calculated based on the overheard information about their neighbors constrained region at the previous step
- This iteration continues until there is no longer a significant improvement in the bounds and can provide a solution that is near-optimal (or optimal)

Network localization can be performed in the presence of mobile reference/target nodes

- If the mobile node is a reference and able to provide an accurate location beacon, then it can substantially improve localization over time because each new observation of the moving beacon introduces additional constraints

Network-wide localization - Rss-based joint estimation

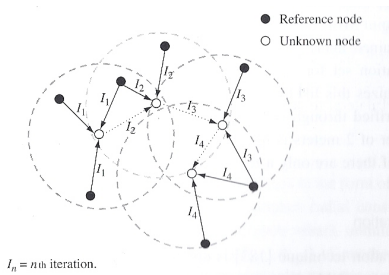
- We can measure the radio signal strengths between all pairs of nodes in the network that are within detection range and use a joint maximum likelihood estimation (MLE) technique to determine the location of unknown nodes.
- In the MLE firstly, an expression is derived for the likelihood that the obtained matrix of RSSIs would be received given a particular location set for all nodes
- the objective is to find the location that maximizes this likelihood

Network-wide localization - Iterative multialteration

- Applicable when inter-node distance info is available
- Applies the triangulation technique in an iterative manner to determine the location of all nodes
- Firstly, determines the location of the unknown node with the most reference nodes in its neighborhood
- This node is then added to the reference nodes and the process is repeated

Network-wide localization - Iterative multialteration

- It may not be applicable if there is no node that has sufficient reference nodes in its neighborhood
- The use of localized unknown nodes as reference nodes can introduce cumulative error in the process



Network-wide localization - Multi-hop distance-estimation approaches

Alternatively, we can estimate distances to reference nodes that may be several hops away. There are three variants of this approach:

- 1 DV-hop: each unknown node determines its distance from various anchor nodes by multiplying the least number of hops to the anchor nodes with an estimated average distance per hop.
- 2 DV-distance: distance estimates are used to determine the shortest distance path between the unknown nodes and reference nodes.
- 3 Euclidean propagation: in this case geometric relations are used in addition to distance estimates. E.g. consider a quadrilateral ABCR, where A and R are at opposite ends; if A knows AB, AC, BC and B and C have estimates of their distance to R, then A can use geometric relations to calculate AR.

Network-wide localization - Force-calculation approach

In this approach a physics-based analogy is used:

- Each node picks a reasonable initial guess as to its location
- $d_{i,j}$ is the calculated distance between two nodes
- $\bar{d}_{i,j}$ the estimated distance
- $\vec{u}_{i,j}$ the unit vector between them
- H_i is the set of all neighboring nodes of i

Network-wide localization - Force-calculation approach

Then a vector force on a link and the resultant force on a node can be defined as

$$\vec{F}_{i,j} = (d_{i,j} - \bar{d}_{i,j}) \cdot \vec{u}_{i,j}$$

$$\vec{F}_i = \sum_{j \in H_i} \vec{F}_{i,j}$$

Each unknown node updates its position in the direction of the resulting vector force in small increments over several iterations.

- May be susceptible to local minima

Network-wide localization - Reference-less localization

- If we have GPS information and no reference nodes we can obtain **relative**, instead of absolute, coordinates
- While such a map is not useful for location stamping of data, it can be useful for geographic routing schemes, etc.

Three scenarios with successively fewer assumptions for the virtual coordinate system:

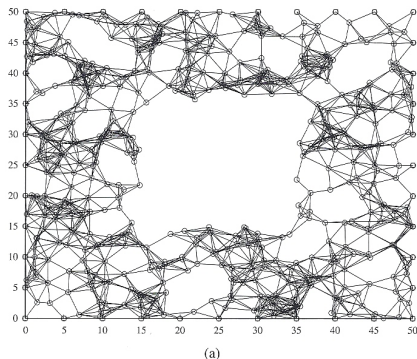
- All (and only these) nodes at the boundary of the network are reference nodes
- Nodes at the boundary are aware that they are at the boundary, but are not reference nodes
- There are no reference nodes in the network, and no nodes are aware that they are at the boundary

1st scenario: All (and only these) nodes at the boundary of the network are reference nodes

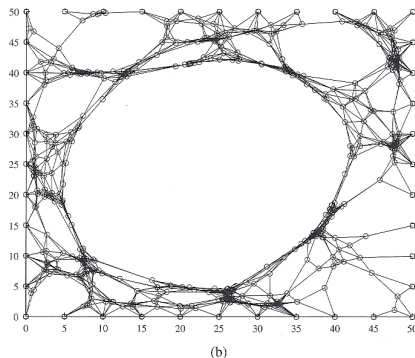
- All interior nodes begin by assuming a initial coordinate (0,0)
- At each step each node determines its location as the centroid of the locations of all its neighbors

Network-wide localization - Reference-less localization

This algorithm tends to stretch (see next slide) the locations of nodes through the location region



Original map



Obtained relative map

- It results in only slightly longer routing paths for greedy geographic routing and even slightly better routing success rates

2nd scenario: Nodes at the boundary are aware that they are at the boundary, but are not reference nodes

- Firstly, the border nodes first flood messages to communicate with each other and determine the pair-wise hop-counts between themselves
- These hop-counts are used in a triangulation algorithm to obtain virtual coordinates for the set B of all border nodes by minimizing

$$\sum_{i,j \in B} (\text{hops}(i,j) - \text{dist}(i,j))^2$$

where $\text{hops}(i,j)$ is the number of hops between border nodes i,j and $\text{dist}(i,j)$ is their Euclidean distance for given virtual coordinates.

3rd scenario: There are no reference nodes in the network, and no nodes are aware that they are at the boundary

- Any node that is farthest away from a common node in terms of hop-count with respect to all its two-hop neighbors can determine that it is on the border
- This hop-count determination is performed through a flood from one of the bootstrap nodes

- Ananth Rao, Sylvia Ratnasamy, Christos Papadimitriou, Scott Shenker, and Ion Stoica. *Geographic routing without location information*. MobiCom 2003. 393-422.
- Andreas Savvides, Chih-chieh Han and Mani B. Strivastava. *Dynamic Fine-Grained Localization in Ad-Hoc Networks of Sensors*. MobiCom 2001.
- Paramvir Bahl and Venkata N. Padmanabhan. *RADAR: an in-building RF-based user location and tracking system*. INFOCOM 2000.
- Kiran Yedavalli, Bhaskar Krishnamachari, Sharmila Ravula and Bhaskar Srinivasan . *Ecolocation: A sequence based technique for RF localization in wireless sensor networks*. IPSN 2005.

- Neal Patwari, Robert J. O Dea, Yanwei Wang and Robert J. O dea. *Relative Location in Wireless Networks*. VTC-Spring 2001.
- Ananth Rao, Sylvia Ratnasamy, Christos Papadimitriou, Scott Shenker, and Ion Stoica. *Geographic routing without location information*. MobiCom 2003.

Triangulation using distance estimates

- By subtracting out the n -th equation from the rest, we would have $n - 1$ equations of the following form:

$$x_i^2 + y_i^2 - x_n^2 - y_n^2 - d_i^2 + x_n^2 = 2 \cdot x_0 \cdot (x_i - x_n) + 2 \cdot y_0 \cdot (y_i - y_n)$$

- which yields the linear relationship

$$A \cdot \bar{x} = B \quad (2)$$

- where A is an $(n - 1) \cdot 2$ matrix such that the i -th row of A is $[2 \cdot (x_i - x_n) \quad 2 \cdot (y_i - y_n)]$
- $\bar{x}$ is the column vector representing the coordinates of the unknown location $x_0 y_0]^T$
- B is the $(n - 1)$ element column vector whose i -th term is the expression $x_i^2 + y_i^2 - x_n^2 - y_n^2 - d_i^2 + x_n^2$.

Triangulation using distance estimates

- In practice we cannot determine B , since we have access to only the estimated distances
- So we can calculate instead the elements of the related vector $\bar{B}$ which is the same as B with $\bar{d}$ substituted for d_i
- Now the least squares solution to equation (2) is to determine an estimate for $\bar{x}$ that minimizes $\|A \cdot \bar{x} - \bar{B}\|_2$.
- Such an estimate is provided by

$$\bar{x} = (A^T \cdot A)^{-1} \cdot A^T \cdot \bar{B}$$