Localization in WSN

- WSN provides information about the spatio-temporal characteristics of the observed physical world
- sensor observation as a tuple of the form $< S, T, M >$
  - S is the spatial location of the measurement
  - T the time of the measurement
  - M the measurement itself
- Why is the location of node (or tuple information) in the network so important?
- How can the spatial location of nodes be determined?
- Localization is quite a broad problem domain
- Location information of nodes in the network is fundamental for a number of reasons:

Application of Localization Information

- to provide location stamps for individual sensor measurements that are being gathered
- to locate and track point objects in the environment
- to monitor the spatial evolution of a diffuse phenomenon over time, such as an expanding chemical plume
- for instance, this information is necessary for in-network processing algorithms that determine and track the changing boundaries of such a phenomenon
- to achieve load balancing in topology control mechanisms; if nodes are densely deployed, geographic information of nodes can be used to selectively shut down some percentage of nodes in each geographic area to conserve energy, and rotate these over time to achieve load balancing

Application of Localization Information

- to form clusters; it can be used to define a partition of the network into separate clusters for hierarchical routing and collaborative processing
- to facilitate routing of information through the network; geographic routing algorithms that utilize location information instead of node addresses to provide efficient routing
- to perform efficient spatial querying; a sink or gateway node can issue queries for information about specific locations or geographic regions; it can be used to scope the query propagation instead of flooding the whole network, which would be wasteful of energy
- to determine the quality of coverage; if node locations are known, the network can keep track of the extent of spatial coverage provided by active sensors at any time

Is it so easy ??????

- Ad hoc sensor networks
- various applications: necessary to accurately orient the nodes with respect to a global coordinate system in order to report data that is geographically meaningful
- determine physical coordinates; these coordinates can be global, or relative, they are an arbitrary rigid transformation away from the global coordinate system
- Anchor/Beacon nodes that know their global coordinates a priori, may be hard coded, or acquired through additional hardware like GPS receiver
- At a minimum, three non-collinear beacon nodes are required to define a global coordinate system in two dimensions
- If three dimensional coordinates are required, then at least four non-coplanar beacons must be present

Typical Application and Issues

- Passive habitat monitoring: tracking movements and population counts of animals
- currently employed techniques do not scale; no automated system
- passive source localization system - localizing sound, camera system
- nodes with microphones distributed; estimate location of source using times of arrival of signal
- structured - location information can be obtained using existing infrastructure GPS or Cellular phone positioning techniques
- GPS receiver - expensive, works if good satellite coverage
- later provides poor location accuracy - subset of nodes have know location a priori and positions of other nodes must be determined using some localization technique

Basic Issues in Localization

- what to localize? reference node: having a priori known locations, unknown nodes
- these nodes could be static or mobile; unknown nodes may be cooperative or non-cooperative (can not participate actively)
- when to localize? one time process - static environment, on-the-fly or refresh location information as objects or nodes move around
- how well to localize? resolution of location information desired
- where to localize? at central location or distributed iterative manner within reference nodes or unknown nodes
- how to localize? different signal measurements: RF signal strength, packet-loss statistics, acoustic/ultrasound signals, infrared
- the signals may be emitted and measured by the reference nodes, by the unknown nodes, or both

Issues in Localization

- no A-priori knowledge of sensor locations in large scale and ad-hoc deployments
- must be able to operate outdoors in various weather conditions
- power efficiency - battery lifetime
- outdoor vs. indoor environment, smart environment
- cooperative vs. non-cooperative target - animal
- accuracy, active vs. passive infrastructure, dynamic vs. static
- granularity and scale of measurement, cost involved
- coordinate system: global, absolute, relative
- Form factor: size of node, multiple sensors, communications requirements, time synchronization

Issues in Localization Algorithm Design

- Resource constraints
- Node density
- Environmental obstacles and terrain irregularities
- Non-convex topologies
  - having trouble positioning nodes near the edges
  - fewer range measurements for border nodes
  - sensors outside the main convex body of the network can often prove unlocalizable

Categories of Localization System

- active localization: emit signals into the environment that are used to measure range to the target; signals emitted by infrastructure components or by targets
- active non-cooperative and cooperative target
- cooperative infrastructure
- passive localization: discover ranges and locations by passively monitoring existing signals in a particular channel (ToF, TDoA)
- blind Source Localization: a signal source is localized without any a priori knowledge of the type of signal emitted; done by blind beam-forming which effectively cross-correlates the signals from different receivers
- these techniques generally only work so long as the signals being compared are coherent

Broad Category of Localization approaches

- the basic localization algorithm based on techniques: proximity, calculation of centroids, constraints, ranging, angulation, pattern recognition, multi-dimensional scaling
- Coarse-grained localization: minimal information used to compute location
  - discrete measurements
  - binary proximity (nodes near to each other)
  - cardinal direction (node in the north, east etc.)
- Fine-grained localization: detailed information
  - RF power, signal waveform, time stamps etc.
- tradeoff
  - lower resource consumption and equipment cost
  - provide lower accuracy than the detailed information techniques

Coarse-grained localization

- set of reference nodes, emit beacon include location IDs and unknown node determines which node it is closet to;
- unknown node emits a beacon then reference node that hears the beacon uses its own location to determine the location of the unknown node
- proximity detection e.g. active badge in an indoor office environment, or passive RF identification detected by reader within a short range
- inventory tracking
- key difference in RFID proximity detection compared with active badges is that the unknown nodes are passive tags, being queried by the reference nodes

Coarse-grained localization

- Centroid calculation: used when density of reference nodes is high within the range of the unknown node
- Let there be n reference nodes detected within the proximity of the unknown node
- the location of the ith such reference denoted by (x_i, y_i)
- the location of the unknown node (x_u, y_u) is determined as

$$x_u = \frac{1}{n}\sum_{i=1}^{n} x_i \quad y_u = \frac{1}{n}\sum_{i=1}^{n} y_i$$

- each node having a simple circular range R in an infinite square mesh of reference nodes spaced a distance d apart
- shown through simulations that, as the overlap ratio R/d is increased from 1 to 4, the average RMS error in localization is reduced from $0.5d$ to $0.25d$

Coarse-grained localization

- Geometric constraints
- radio signal coverage described by a geometric shape
- location estimates by determining which geometric regions that node is constrained to be in, because of intersections between overlapping coverage regions
- region of radio coverage using a circle of radius R_{max}
- if an unknown node hears from several reference nodes, it can determine that it must lie in the geometric region described by the intersection of circles of radius R_{max} centered on these nodes
- rectangular bounding boxes: node determines bounds $x_{min}, y_{min}, x_{max}, y_{max}$ on its position
- the unknown nodes use the centroid of the overlapping region as a specific location estimate if necessary, determine a bound on the location error using the size of this region
- arbitrary shapes can be potentially computed in this manner

Coarse-grained localization

- Approximate point in triangle (APIT)
- localization using geometric constraints is the APIT technique
- provides location estimates as the centroid of an intersection of regions
- novelty lies in how the regions are defined as triangles between different sets of three reference nodes (rather than the coverage of a single node)
- Identifying codes
- utilizes overlapping coverage regions to provide localization
- referred to as the identifying code construction (ID-CODE) algorithm
- the sensor deployment is planned in such a way as to ensure that each resolvable location is covered by a unique set of sensors

Coarse-grained localization

- algorithm runs on a deployment region graph $G = (V, E)$ vertices V represent the different regions, and the edges E represent radio connectivity between regions
- goal is to construct an identifying code for any distinguishable graph, with each vertex in the code corresponding to a region where a reference node must be placed
- Once this is done, by the definition of the identifying code, each location region in the graph will be covered by a unique set of reference nodes
- obtaining a minimal cardinality identifying code is known to be NP-complete

Fine-grained localization

- basic time-of-flight techniques using RF signals are not capable of providing precise distance estimates over short ranges typical of WSN because of synchronization limitations
- other techniques such as radio signal strength (RSS) measurements, time difference of arrival (TDoA) and Angle of Arrival (AoA) are used for distance-estimation
- triangulation using distance estimates, pattern matching, and sequence decoding used in the large-scale GPS

Received Signal Strength Indication (RSSI)

- hardware methods of computing distance measurements between nearby sensor nodes (i.e. ranging)
- every sensor has a radio and it helps localize the network
- two important techniques for using radio information to compute ranges
- hop count and Received Signal Strength Indication (RSSI)
- the energy of a radio signal diminishes with the square of the distance from the signal's source
- a node listening to a radio transmission should be able to use the strength of the received signal to calculate its distance from the transmitter
- RSSI ranging measurements contain noise on the order of several meters
- this noise occurs because radio propagation tends to be highly non-uniform in real environments

Received Signal Strength Indication (RSSI)

- radio propagates differently over asphalt than over grass, physical obstacles such as walls, furniture, etc. reflect and absorb radio waves
- first-order approximation, mean radio signal strengths diminish with distance according to a power law
- model used for wireless radio propagation

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

- $P_{\mathrm{r,dB}}(d)$ is the received power at distance d
- $P(d_0)$ is the received power at some reference distance d_0
- η the path-loss exponent
- $X_{\sigma,\mathrm{dB}}$ a log-normal random variable with variance σ^2 that accounts for fading effects

Radio signal-based distance-estimation (RSS)

- if the path-loss exponent for a given environment is known the received signal strength can be used to estimate the distance
- the fading term often has a large variance, which can significantly impact the quality of the range estimates
- this is the reason RF-RSS-based ranging techniques may offer location accuracy only on the order of meters or more
- RSS-based ranging may perform much better in situations where the fading effects can be combated by diversity techniques that take advantage of separate spatio-temporally correlated signal samples

Radio Hop Count

- even though RSSI is too inaccurate for many applications, the radio can still be used to assist localization
- hop count to be a useful way to compute inter-node distances
- the local connectivity information provided by the radio defines an unweighted graph, where the vertices are sensor nodes, and edges represent direct radio links between nodes
- the hop count h_{ij} between sensor nodes s_i and s_j is then defined as the length of the shortest path in the graph between s_i and s_j
- if the hop count between s_i and s_j is h_{ij} then the distance between s_i and s_j, d_{ij}, is less than $R * h_{ij}$, where R is again the maximum radio range

Time Difference of Arrival (TDoA)

- time-of-flight techniques show poor performance due to precision constraints
- RSS techniques, although somewhat better, are still limited by 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
- in this technique: simultaneously transmit both the radio and acoustic signals (audible or ultrasound) and measure the times T_r and T_s of the arrival of these signals respectively at the receiver
- the speed of the radio signal is much larger than the speed of the acoustic signal, the distance is then simply estimated as $(T_s - T_r)$ V_s, where V_s is the speed of the acoustic signal

Time Difference of Arrival (TDoA): Issues

- limitation of acoustic ranging is that it generally requires the nodes to be in fairly close proximity to each other (within a few meters) and preferably in line of sight
- there is also some 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 basic idea is to send a pseudo-random noise sequence as the acoustic signal and use a matched filter for detection
- acoustic TDoA ranging techniques can be very accurate in practical settings

Distance-estimation using TDoA

- each node is equipped with a speaker and a microphone
- some systems use ultrasound while others use audible frequencies
- the transmitter first sends a radio message, waits some fixed interval of time, t_{delay} (which might be zero), and then produces a fixed pattern of chirps on its speaker
- listening nodes hear the radio signal, they note the current time, t_{radio}, then turn on their microphones
- when their microphones detect the chirp pattern, they again note the current time, t_{sound}
- once they have t_{radio}, t_{sound}, and t_{delay}, the listeners can compute the distance d between themselves and the transmitter using the fact that radio waves travel substantially faster than sound in air

$$d = (s_{radio} - s_{sound}) * (t_{sound} - t_{radio} - t_{delay})$$

Triangulation using distance estimates

- the location of the unknown node (x_0, y_0) can be determined based on measured distance estimates $\hat{d}_i$ to n reference nodes $\{(x_1, y_1), \ldots, (x_i, y_i), \ldots, (x_n, y_n)\}$
- this can be formulated as a least squares minimization problem
- let d_i be the correct Euclidean distance to the n reference nodes,

$$d_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$$

- the difference between the measured and actual distances can be represented as

$$\rho_i = \hat{d}_i - d_i$$

- the least squares minimization problem is then to determine the (x_0, y_0) that minimizes $\sum_{i=1}^{n} (\rho_i)^2$
- it can be solved by the use of gradient descent techniques or by iterative successive approximation techniques
- alternative approach provides a numerical solution to an over-determined $(n \geq 3)$ linear system