Localization Algorithms in Wireless Sensor Networks: Current Approaches and Future Challenges

Amitangshu Pal

Department of Electrical and Computer Engineering

The University of North Carolina at Charlotte

9201 University City Blvd, Charlotte, North Carolina 28223-0001

apal@uncc.edu

Tel: 1-980-229-3383

Abstract

Recent advances in radio and embedded systems have enabled the proliferation of wireless sensor networks. Wireless sensor networks are tremendously being used in different environments to perform various monitoring tasks such as search, rescue, disaster relief, target tracking and a number of tasks in smart environments. In many such tasks, node localization is inherently one of the system parameters. Node localization is required to report the origin of events, assist group querying of sensors, routing and to answer questions on the network coverage. So, one of the fundamental challenges in wireless sensor network is node localization. This paper reviews different approaches of node localization discovery in wireless sensor networks. The overview of the schemes proposed by different scholars for the improvement of localization in wireless sensor networks is also presented. Future research directions and challenges for improving node localization in wireless sensor networks are also discussed.

Keywords: Centralized Localization, Distributed Localization, Beacon-based distributed algorithms, Relaxation-based distributed algorithms, Coordinate system stitching based distributed algorithms, Diffusion, Bounding Box, Gradient, Wireless Sensor Networks.

1. Introduction

The massive advances of microelectromechanical systems (MEMS), computing and communication technology have fomented the emergence of massively distributed, wireless sensor networks consisting of hundreds and thousands of nodes. Each node is able to sense the environment, perform simple computations and communicate with its other sensors or to the central unit. One way of deploying the sensor networks is to scatter the nodes throughout some region of interest. This makes the network topology random. Since there is no a priori communication protocol, the network is ad hoc. These networks are tremendously being implemented to perform a number of tasks, ranging from environmental and natural habitat monitoring to home networking, medical applications and smart battlefields. Sensor network can signal a machine malfunction to the control centre in a factory or it can warn about smoke on a remote forest hill indicating that a forest fire is about to start. On the other hand wireless sensor nodes can be designed to detect the ground vibrations generated by silent footsteps of a burglar and trigger an alarm.

Since most applications depend on a successful localization, i.e. to compute their positions in some fixed coordinate system, it is of great importance to design efficient localization algorithms. In large scale ad hoc networks, node localization can assist in routing [1], [2], [3]. In the smart kindergarten [4] node localization can be used to monitor the progress of the children by tracking their interaction with toys and also with each other. It can also be used in hospital environments to keep track of equipments, patients, doctors and nurses [1].

For these advantages precise knowledge of node localization in ad hoc sensor networks is an active field of research in wireless networking. Unfortunately, for a large number of sensor nodes, straightforward solution of adding GPS to all nodes in the network is not feasible because:

- In the presence of dense forests, mountains or other obstacles that block the line-of-sight from GPS satellites, GPS cannot be implemented.
- The power consumption of GPS will reduce the battery life of the sensor nodes and also reduce the effective lifetime of the entire network.
- In a network with large number of nodes, the production cost factor of GPS is an important issue.
- Sensor nodes are required to be small. But the size of GPS and its antenna increases the sensor node form factor.

For these reasons an alternate solution of GPS is required which is cost effective, rapidly deployable and can operate in diverse environments.

The paper is organized as follows. Section 2 presents the formulation of localization problem in wireless sensor networks. Related work has been discussed in section 3. In section 4, presents different location discovery approaches to solve the problem of localization in wireless sensor networks (WSN). In section 5 different proposals to improve localization in WSN are discussed. Section 6 states the summary of all proposals. Section 7 concludes the paper where the future challenges and directions to improve localization in WSN technology are described.

2. Problem Definition

Consider the case when we have deployed a sensor network consist of N sensors at locations S = {S_1, S_2,......,S_N}. Let S_x^i refer to the x-coordinate of the location of sensor i and let S_y^i and S_z^i refer to the y and z coordinates, respectively. Constraining S_z^i to be 0 suffices the 2D version of this problem. Determining these locations constitutes the localization problem. Some sensor nodes are aware of their own positions, these nodes are known as *anchors* or *beacons*. All the other nodes localize themselves with the help of location references received from the anchors. So, mathematically the localization problem can be formulated as follows: given a multihop network, represented by a graph G = (V, E), and a set of beacon nodes B, their positions {x_b, y_b} for all b ε B, we want to find the position {x_u, y_u} for all unknown nodes u ε U.

3. Related Work

Localization in WSN is an active area of research and so there are some existing literature surveys [23], [24] on this topic. In these literatures the authors discuss most important localization techniques and critique those techniques. But there are some existing techniques which use two localization techniques such as multidimensional scaling (MDS) and proximity based map (PDM) [16] or MDS and Ad-hoc Positioning System (APS) [17]. These techniques have not been mentioned in any literature but these techniques give new directions in WSN localization as these schemes give high accuracy in low communication and computation cost. On the other hand interferometric ranging based localization has been proposed in [18], [19], [20] which have not been discussed by any existing literature. Moreover due to channel fading and noise corruption error propagation comes in picture. To suppress this error propagation a localization scheme has been proposed in [21] which was not been discussed by any literature. This literature gives comprehensive summary of these techniques along with other existing localization schemes. At the same time this paper also compares all localization techniques and also provides future research directions in this area.

4. Different Location Discovery Approaches

Existing location discovery approaches basically consists of two basic phases: (1) distance (or angle) estimation and (2) distance (or angle) combining. The most popular methods for estimating the distance between two nodes are described below:

Received Signal Strength Indicator (RSSI): RSSI measures the power of the signal at the receiver and based on the known transmit power, the effective propagation loss can be calculated. Next by using theoretical and empirical models we can translate this loss into a distance estimate. This method has been used mainly for RF signals. RSSI is a relatively cheap solution without any extra devices, as all sensor nodes are likely to have radios. The performance, however, is not as good as other ranging techniques due to the multipath propagation of radio signals. In [26], the authors characterize the limits of a variety of approaches to indoor localization using signal strengths from 802.11 routers. They also suggest that adding additional hardware or altering the model of the environment is the only

alternative to improve the localization performance.

Time based methods (ToA, TDoA): These methods record the time-of-arrival (ToA) or time-difference-of-arrival (TDoA). The propagation time can be directly translated into distance, based on the known signal propagation speed. These methods can be applied to many different signals, such as RF, acoustic, infrared and ultrasound. TDoA methods are impressively accurate under line-of-sight conditions. But this line-of-sight condition is difficult to meet in some environments. Furthermore, the speed of sound in air varies with air temperature and humidity, which introduce inaccuracy into distance estimation. Acoustic signals also show multi-path propagation effects that may impact the accuracy of signal detection.

Angle-of-Arrival (AoA): AoA estimates the angle at which signals are received and use simple geometric relationships to calculate node positions. Generally, AoA techniques provide more accurate localization result than RSSI based techniques but the cost of hardware of very high in AoA.

For the combining phase, the most popular alternatives are:

Hyperbolic trilateration: The most basic and intuitive method is called hyperbolic trilateration. It locates a node by calculating the intersection of 3 circles as shown in Fig. 1(a).
Triangulation: This method is used when the direction of the node instead of the distance is estimated, as in AoA systems. The node positions are calculated in this case by using the trigonometry laws of sines and cosines (shown in Fig. 1(b)).
Maximum Likelihood (ML) estimation: ML estimation estimates the position of a node by minimizing the differences between the measured distances and estimated distances (shown in Fig. 1(c)).

5. Different Proposals For Network Management And Control Issues

This section presents different proposals put forward by the research community in the areas of localization in wireless sensor networks and critiques their contributions.
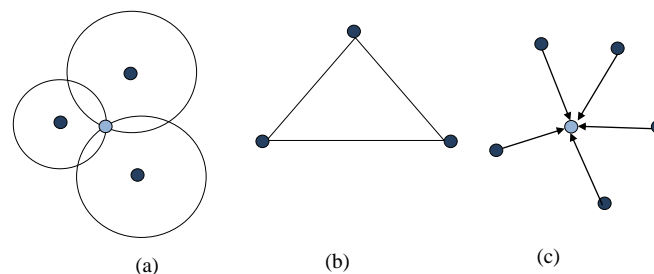Research on localization in wireless sensor networks can be classified into two broad categories.

<center>(a)   (b)   (c)</center>

Fig.1. Localization techniques a) Hyperbolic trilateration b) Triangulation c) Maximum Likelihood Estimation

Centralized Localization: Centralized localization is basically migration of inter-node ranging and connectivity data to a sufficiently powerful central base station and then the migration of resulting locations back to respective nodes. The advantage of centralized algorithms are that it eliminates the problem of computation in each node, at the same time the limitations lie in the communication cost of moving data back to the base station. As

representative proposals in this category [5], [6], [7] are explained in greater detail.

Distributed Localization: In Distributed localizations all the relevant computations are done on the sensor nodes themselves and the nodes communicate with each other to get their positions in a network. Distributed localizations can be categorized into three classes.

- *Beacon-based distributed algorithms:* Beacon-based distributed algorithms start with some group of beacons and nodes in the network to obtain a distance measurement to a few beacons, and then use these measurements to determine their own location. Some of the proposals [8], [9], [10], [11], in this category are described below.

- *Relaxation-based distributed algorithms:* In relaxation-based distributed algorithms use a coarse algorithm to roughly localize nodes in the network. This coarse algorithm is followed by a refinement step, which typically involves each node adjusting its position to approximate the optimal solution. Some of the proposals [12], [13] in this category are discussed in greater details.

- *Coordinate system stitching based distributed algorithms:* In Coordinate system stitching the network is divided into small overlapping subregions, each of which creates an optimal local map. Next the scheme merges the local maps into a single global map. Some approaches [14], [15] of this category are examined in the next section.

- *Hybrid localization algorithms:* Hybrid localization schemes use two different localization techniques such as : multidimensional scaling (MDS) and proximity based map (PDM) or MDS and Ad-hoc Positioning System (APS) to reduce communication and computation cost. Such kinds of approaches are depicted in [16], [17].

- *Interferometric ranging based localization:* Radio interferometric positioning exploits interfering radio waves emitted from two locations at slightly different frequencies to obtain the necessary ranging information for localization. Such types of localization techniques are proposed in [18], [19] and [20].

- *Error propagation aware localization:* When sensors communicate with each other, error propagation can be caused due to the undesirable wireless environment, such as channel fading and noise corruption. To suppress error propagation [21] has proposed a scheme called error propagation aware (EWA) algorithm.

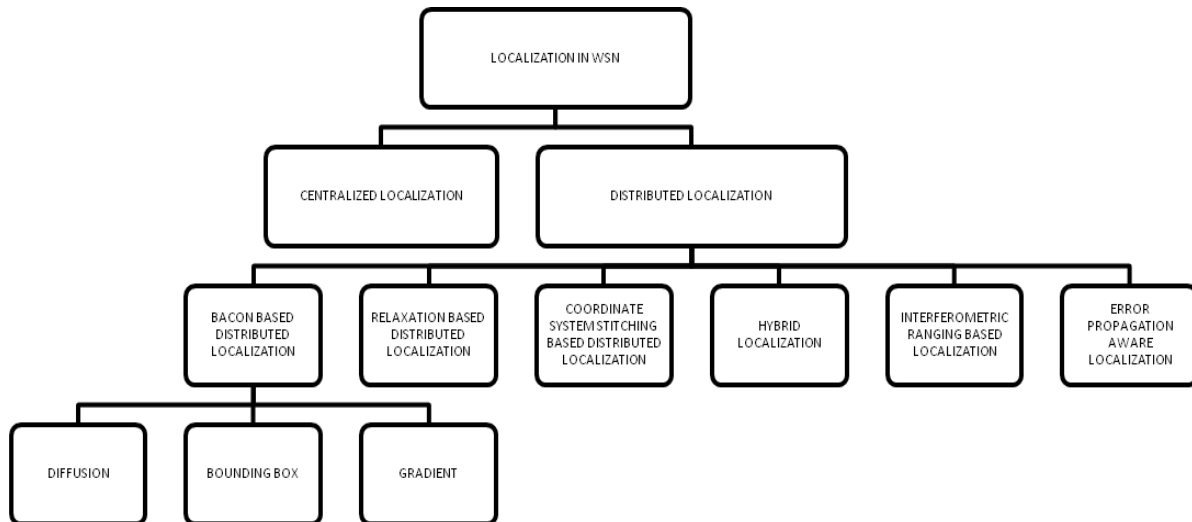A classification of various schemes is shown in Fig. 2.

Fig.2. Classification of various proposals for Localization in WSN

4.1. Centralized Localization

4.1.1. MDS-MAP

In [5] the authors present a centralized algorithm called MDS-MAP (mentioned in Appendix A) which basically consists of three steps.

1. First the scheme computes shortest paths between all pairs of nodes in the region of consideration by the use of all pair shortest path algorithm such as Dijkstra's or Floyd's algorithm. The shortest path distances are used to construct the distance matrix for MDS.

2. Next the classical MDS is applied to the distance matrix, retaining the first 2 (or 3) largest eigenvalues and eigenvectors to construct a 2-D (or 3-D) relative map that gives a location for each node. Although these locations may be accurate relative to one another, the entire map will be arbitrarily rotated and flipped relative to the true node positions.

3. Based on the position of sufficient anchor nodes (3 or more for 2-D, 4 or more for 3-D), transform the relative map to an absolute map based on the absolute positions of anchors which includes scaling, rotation, and reflection. The goal is to minimize the sum of squares of the errors between the true positions of the anchors and their transformed positions in the MDS map.

The advantage of this scheme is that it does not need anchor or beacon nodes to start with. It builds a relative map of the nodes even without anchor nodes and next with three or more anchor nodes, the relative map is transformed into absolute coordinates. This method works well in situations with low ratios of anchor nodes. A drawback of MDS-MAP is that it requires global information of the network and centralized computation.

4.1.2. Localize node based on Simulated Annealing

In [6] the authors propose an innovative approach based on Simulated Annealing to localize the sensor nodes in a centralized manner. Since the algorithm is centralized, it enjoys

the access to estimated locations and neighbourhood information of all localizable nodes in the system. Let us consider a sensor network of m anchor nodes with known locations and n-m sensor nodes with unknown locations. As the proposed algorithm is implemented in a centralized architecture, it has access to estimated locations and neighborhood information of all localizable nodes in the system. The proposed scheme is based on two stages. In the first stage simulated annealing is used to obtain an estimate of location of the localizable sensor nodes using distance constraints. Let us define the set Ni as a set containing all one hop neighbors of node i. The localization problem can be formulated as:

$$\text{Min} \sum_{i=m+1 \text{ to } n}\sum_{j \in Ni} (d^{\wedge}_{ij} - d_{ij})^2 \qquad (1)$$

In equation (1), d_{ij} is the measured distance between node i and its neighbor j; $d^{\wedge}_{ij} = \sqrt{\{(x^{\wedge}_i - x^{\wedge}_j)^2 + (y^{\wedge}_i - y^{\wedge}_j)^2\}}$ is the estimated distance; $(x^{\wedge}_i, y^{\wedge}_i)$ and $(x^{\wedge}_j, y^{\wedge}_j)$ are the estimated coordinates of node i and its one hop neighbor j respectively and the cost function CF = $\sum_{i=m+1 \text{ to } n}\sum_{j \in Ni} (d^{\wedge}_{ij} - d_{ij})^2$. Then according to Simulated Annealing coordinate estimate $(x^{\wedge}_i, y^{\wedge}_i)$ of any chosen node i is given a small displacement in a random direction and the new value of the cost function is calculated for the new location estimate. If $\Delta(CF) \leq 0$, $(\Delta(CF) = CF_{new} - CF_{old})$ then the perturbation is accepted and the new location estimate is used as the starting point of the next step. Otherwise the probability that the displacement is accepted is $P(\Delta(CF)) = \exp(-\Delta(CF)/T)$. Here T is a control parameter and P is a monotonically increasing function of T.

In the next stage of the algorithm the authors eliminate the error caused by flip ambiguity. Flip ambiguity occurs when a node's neighbors are placed in positions such that they are approximately on the same line, this node can be reflected across the line of best fit produced by its neighbors with essentially no change in the cost function. In Fig. 3, the neighbors of node A are nodes B, C, D and E which are almost collinear and the node A could be flipped across the line of best fit of nodes B, C, D and E to location A$^{/}$ with almost no change in the cost function. But we should note from Fig. 3 that the flipped position A$^{/}$ has gone into the wrong neighborhood of nodes H and I. Based on this observation the authors define a complement set comp(Ni) of the set N_i as a set containing all nodes which are not neighbors of node i. If R is the transmission range of the sensor node and the estimated coordinate of node j $\in$ comp(Ni) is such that $d^{\wedge}_{ij} < R$, then the node j has been placed in the wrong neighborhood of node i, resulting in both nodes i and j having each other as wrong neighbors. So the minimum error due to the flip is $d^{\wedge}_{ij} - R$ and the new localization problem can be formulated as in equation (2).

$$\text{Min} \sum_{i=m+1 \text{ to } n} (\sum_{j \in Ni} (d^{\wedge}_{ij} - d_{ij})^2 + \sum (d^{\wedge}_{ij} - R)^2) \qquad (2)$$

The paper presented a novel simulated annealing based localization algorithm which mitigates the flip ambiguity problem. By simulations the authors the authors show that the proposed algorithm gives better accuracy than the semi-definite programming localization.
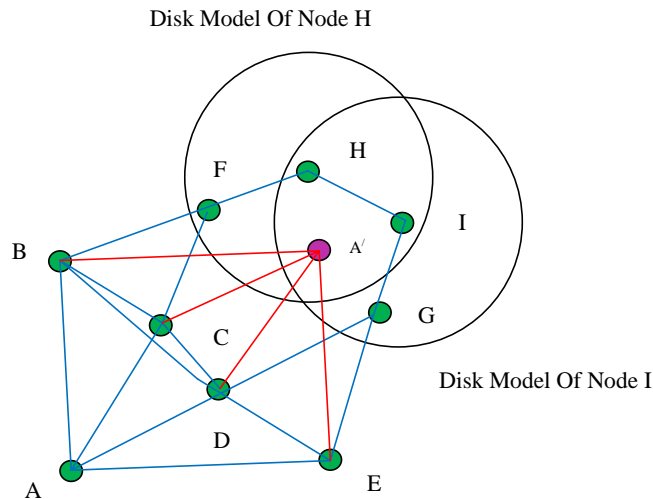
Disk Model Of Node H

Fig.3. Illustration of Flip Ambiguity [6]

They show that the proposed algorithm does not propagate error in localization. The proposed flip ambiguity mitigation method is based on neighborhood information of nodes and it works well in a sensor network with medium to high node density. However when the node density is low, it is possible that a node is flipped and still maintains the correct neighborhood. In this situation, the proposed algorithm fails to identify the flipped node.

4.1.3. A RSSI-based centralized localization technique

In [7] the authors propose a scheme which localizes nodes through RF attenuation in Electromagnetic waves. The scheme basically consists of three stages:

1) *RF mapping of the network*: It is obtained by conveying short packets at different power levels through the network and by storing the average RSSI value of the received packets in memory tables.

2) *Creation of the ranging model*: All the tuples recorded between the two anchors are processed at the central unit to compensate the non linearity and calibrate the model. Let a generic tuple (i, j, P_{tx}, P_{rx}) comes from the RF mapping characterizing stage, where i is the transmitting node and j is the receiving node. Now first the algorithm corrects the received power as $P_{rx}' = f(P_{rx}, P_{tx})$, f() is a function which takes into account the modularity effects. So, the estimated distance between the nodes will be $r_{ij}^0 = m^{-1}(P_{rx}')$

3) *Centralized localization model:* An optimization problem is solved and provides the position of the nodes. The final result can be obtained by minimizing the function

$$E = \sum_{i=1 \text{ to } n} \sum_{j=1 \text{ to } n} (k_{i,j} a_{i,j} (r_{ij} - r_{ij}^0)^2) , r_{ij} = d(i, j) \text{ when i and j are anchors.}$$

Where N is the number of nodes, $a_{i,j}$ is 1 when the link is present and 0 otherwise. Once the distance between the nodes r_{ij} can be expressed in terms of their coordinates $(x, y)_i$ and $(x, y)_j$ the authors solve the minimizing problem by sequential quadratic programming (SQP) method.

The advantage of this scheme is that it is a practical, self-organizing scheme that allows addressing any outdoor environments. The limitation of this scheme is that the scheme is

power consuming because it requires extensive generation and need to forward much information to the central unit.

4.2. Distributed Localization

4.2.1. Beacon based distributed localization

Beacon based approaches can be categorized in Diffusion, Bounding Box and Gradient which are described as follows:

4.2.1.1. Diffusion

In diffusion the most likely position of the node is at the centroid of its neighboring known nodes.

APIT: In [8] the authors describe a novel area-based range free localization scheme, called APIT which requires a heterogeneous network of sensing devices where some devices are equipped with high-powered transmitters and location information. These devices are known as anchors. In this approach the location information is performed by isolating the environment into triangular regions between beaconing nodes. An unknown node chooses three anchors from all audible anchors and tests whether it is inside the triangle formed by connecting these three anchors. APIT repeats this tests with different audible anchor combinations until all combinations are exhausted or the required accuracy is achieved. At this point, APIT calculates the centre of gravity of the intersection of all triangles in which the unknown node resides to determine the estimated position.

The advantage of APIT lies in its simplicity and ease of implementation. But APIT requires a high ratio of beacons to nodes and longer range beacons to get a good position estimate. For low beacon density this scheme will not give accurate results.

4.2.1.2. Bounding Box

Bounding box forms a bounding region for each node and then tries to refine their positions.

4.2.1.2.1. Collaborative Multilateration

In [9] the authors present a collaborative multilateration approach that consists of a set of mechanisms that enables nodes found several hops away from location aware beacon nodes collaborate with each other to estimate their locations with high accuracy. Collaborative multilateration consists of three phases:

- *Forming Collaborative sub trees*: A computation sub tree constitutes a configuration of unknowns and beacons for which the solution of the position estimates of the unknown can be uniquely determined. The requirement of one-hop multilateration for an unknown node is that it is within the range of at least three beacons (see Fig 4(a)). A two hop multilateration represents the case where the beacons are not always directly connected to the nodes but they are within a two hop radius from the unknown node (see Fig. 4(b)).
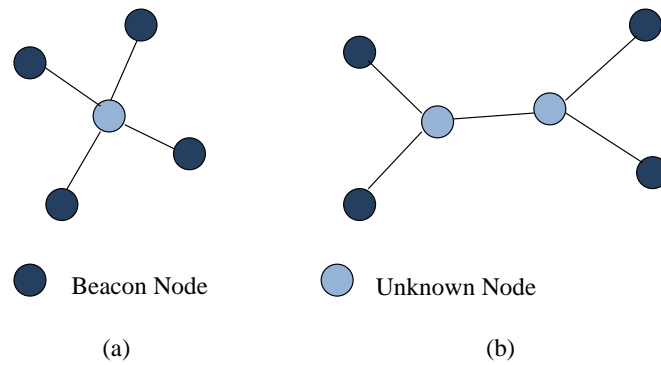
Fig.4. (a) One-hop Multilateration (b) Two-hop Multilaretation

- *Obtaining initial estimates*: This phase is explained by the help of Fig. 5. In Fig. 5 A and B are beacons where C is the unknown node. If the distance between C and A is a then the x coordinate of C are bounded by a to the left and to the right of the x coordinate of A, $x_A - a$ and $x_A + a$. Similarly beacon B which is two hops away from C, bounds the coordinate of C within $x_B - (b + c)$ and $x_B + (b + c)$. by knowing the information, C can determine that its x coordinate bounds with respect to beacons A and B are $x_B + (b + c)$ and $x_A - a$. The same operation is applied on the y coordinates. C then combines its bounds on x and y coordinates, to obtain a bounding box of the region where it lies.
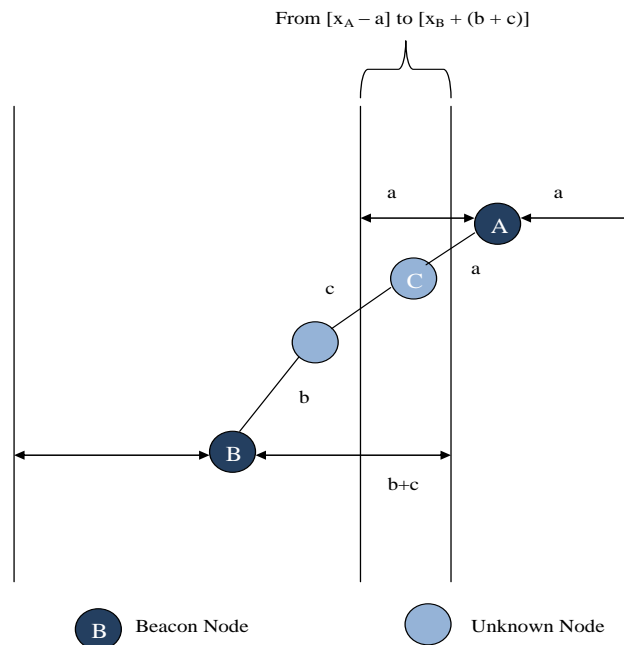

Fig.5. X coordinates bounds for C using initial estimates [9]

- *Position refinement:* In the third phase, the initial node positions are refined using Kalman Filter implementation (mentioned in the Appendix B). Now as most unknown nodes are not directly connected to beacons, they use the initial estimates of their neighbours as the reference points for estimating their locations. As soon as an unknown node computes a new estimate, it broadcasts this estimate to its neighbours, and the neighbour use it to update their own position estimates. As shown in Fig. 6 first node 4 computes its location estimate using beacons 1 and 5 and node 3 as

reference. Once node 4 broadcasts its update, node 3 recomputes its own estimate received from node 4. Next node 3 broadcasts the new estimate and node 4 uses this to compute a new estimate that is more accurate than its previous estimate.

The collaborative multilateration enables sensor nodes to accurately estimate their locations by using known beacon locations that are several hops away and distance measurements to neighboring nodes. At the same time it increases the computational cost also.
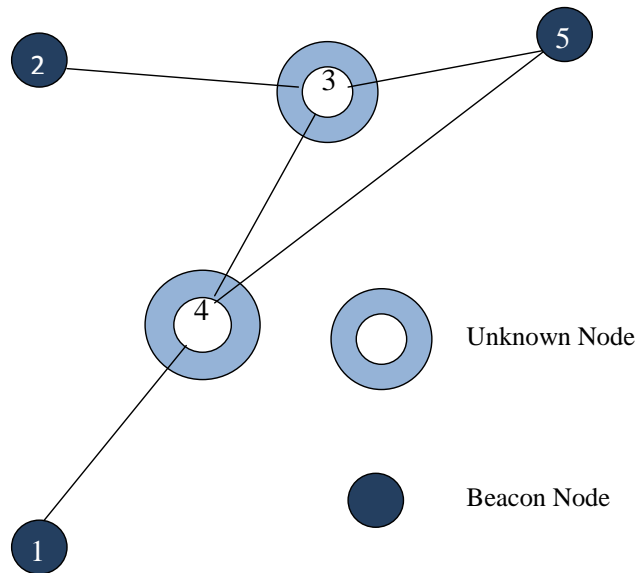

Fig.6. Initial estimates over multiple hops

4.2.1.2.2. *Node localization assuming the region as square box*

In [10] the authors frame the localization problem as follows. They have assumed that in a square region $Q = [0, s] \times [0,s]$, called region of operations, N nodes $S_1, \ldots\ldots, S_N$ have been scattered and each of which is equipped with an RF transceiver with communication range $r > 0$. In other words a node S_i can communicate with every node which lies in its communication region, which is the disk with radius r centered at S_i. The nodes form an ad hoc network Ŋ in which there is an edge between S_i and S_j if their distance is less than r. They scheme assume that there are certain positive number of beacons nodes in Q and other are unknown nodes. Now for any integer $n > 0$, partition Q into n^2 congruent squares called cells of area $(s/n)^2$ and for every known node S, we know the cell which contains S. To make the problem tractable the authors assume that communication range is ρ calls where ρ = [nr/s√2], where [x] denotes the integer part of x, which means that each node S can communicate with every node lying in the square centered at S and containing $(2\rho + 1)^2$ cells. Usually n is large and r is much smaller than n. In particular $2\rho + 1 < n$. Then for an arbitrary unknown S in Ŋ the localization algorithm at S can be written as:

Step 1: Initialize the estimate: $L_s = Q$.

Step 2: Send Hello packets to the neighbours. Each known neighbor sends back (1, a, b), where (a, b) is its grid position, while each unknown neighbor sends (0, 0, 0).

Step 3: For each response (1, a, b), update the estimate as shown in equation (3).

$$L_s = L_s \cap [a - \rho, a + \rho] \times [b - \rho, b + \rho] \qquad (3)$$

Step 4: Stop when all responses are received. The position estimate is L_s.
In this approach an unknown node could query some of its neighbours which reduce communication cost but increases computations.

4.2.1.3. Gradient

In [11] the authors describe an algorithm for organizing a global coordinate system from local information. In this approach ad-hoc sensor nodes are randomly distributed on a two dimensional plane and each sensor communicates with nearby sensors within a fixed distance r, where r is much smaller than the dimension of the plane. In their algorithm they assume some set of sensors as "seed" sensors which are identical to other sensors in capabilities except that they are programmed with their global position. The algorithm consists of two parts:

- *Gradient Algorithm*: - Each seed sensor produces a locally propagating gradient that allows other sensors to estimate the distance from the seed sensors. A seed sensor initiates a gradient by sending its neighbors a message with its location and a count set to one. Each recipient remembers the value of the count and forwards the message to its neighbors with the count incremented by one. Hence a wave of messages propagates outwards from the seed. Each sensor maintains the minimum counter value received and ignores messages containing larger values, which prevents the wave from traveling backwards. If two sensors can communicate with each other directly then they are considered to be within one communication hop of each other. The minimum hop count value, h_i, that a sensor i maintains will eventually be the length of the shortest path to the seed in communication hops. In the proposed ad hoc sensor network, a communication hop has a maximum physical distance of r associated with it. This implies that a sensor i is at most distance $h_i r$ from the seed. However as the average density of sensors increases, sensors with the same hop count tend to form concentric circular rings, of width approximately r, around the seed sensor.

- *Multilateration Algorithm*: - Each sensor uses a multilateral procedure to combine the distance estimates from all the seed sensors to produce their own positions. After receiving at least three gradient values, sensors combine the distances from the seeds to estimate their position relative to the positions of the seed sensors. In particular, each sensor estimates its coordinates by finding coordinates that minimize the total squared error between calculated distances and estimated distances. Sensor j's calculated distance to seed i is:

$$d_{ji} = \sqrt{[(x_i - x_j)^2 + (y_i - y_j)^2}\qquad\qquad(4)$$

and sensor j's total error is:

$$E_j = \sum_{i=1\text{ to }n} (d_{ji} - \hat{d}_{ji})^2 \qquad\qquad(5)$$

In equation (4) and equation (5), n is the number of seed sensors and $\hat{d}_{ji}$ is the estimated

distance computed through gradient propagation. The coordinates are then incrementally updated in proportion to the gradient of the total error with respect to that coordinate.

The advantage of this algorithm is that it can be easily adapted to the addition of sensors, addition of seeds and also death of sensors and seeds. But it requires substantial node density before its accuracy reaches an acceptable level. Besides this hop count is not reliable in measurement because environmental obstacles can prevent edges from appearing in the connectivity graph that otherwise would be present as shown in Fig 7. In Fig 7 the hop count distance between A and E is four hops due to the obstacle, but the real distance is far lesser than four values.
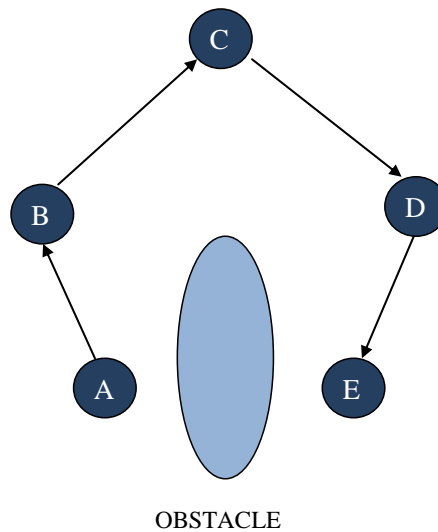

OBSTACLE
Fig.7. Error in hop count distance matrices in the presence of an obstacle

4.2.2. *Relaxation Based Distributed Algorithm*

4.2.2.1. *Spring Model*

In [12] the authors propose an Anchor Free Localization (AFL) algorithm where nodes start from a random initial coordinate assignment and converge to a consistent solution using only local node interactions. The algorithm proceeds in two phases and it assumes the nodes as point masses connected with strings and use force-directed relaxation methods to converge to a minimum-energy configuration.

The first phase is a heuristic that produces a graph embedding which looks similar to the original embedding. The authors assume that each node has a unique identifier and the identifier of node i is denoted by ID_i and the hop-count between nodes i ad j is the number of nodes $h_{i,j}$ along the shortest path between i and j. The algorithm first elects the five reference nodes in which four nodes n_1, n_2, n_3 and n_4 are selected such that they are on the periphery of the graph and the pair (n_1, n_2) is roughly perpendicular to the pair (n_3, n_4). The node n_5 is elected such that it is in the middle of the graph. At first the node with smallest ID is selected. Next the reference node n_1 is selected to maximize $h_{1,2}$. After that n_3 is selected to minimize $| h_{1,3} - h_{2,3} |$ and the tie-breaking rule is to pick the node that minimizes $h_{1,3} + h_{2,3}$. In the next stage n_4 is selected to minimize $| h_{1,4} - h_{2,4} |$ and the ties are broken by picking the node that maximizes $h_{3,4}$. Next n_5 is selected which minimizes $| h_{1,5} - h_{2,5} |$ and from contender nodes

pick the node that minimizes $| h_{3,5} - h_{4,5} |$. So node n_5 is the center of the graph and node n_1, n_2, n_3, n_4 becomes the periphery of the graph. Now for all nodes n_i the heuristics uses the hop-counts $h_{1,i}$, h_{2i}, $h_{3,i}$, $h_{4,i}$, and $h_{5,i}$ from the chosen reference nodes to approximate the polar coordinates (ρ_i, θ_i) where

$$\rho_i = h_{5,i} * R \qquad\qquad (7)$$
$$\theta_i = \tan^{-1}[(h_{1,i} - h_{2,i}) / (h_{3,i} - h_{4,i})] \qquad\qquad (8)$$

and R is the maximum radio range. In the first stage when calculating ρ_k the use of range R to represent one hop-count results in a graph which is physically larger than the original graph and this error can be eliminated in the next stage.

In the second phase, each node n_i calculates the estimated distance $\hat{d}_{i,j}$ to each neighbours n_j and it also knows the measured distance $r_{i,j}$ to neighbour n_j. Now if $\hat{v}_{i,j}$ represent the unit vector in the direction from $\hat{p}_i$ to $\hat{p}_j$ ($\hat{p}_i$ and $\hat{p}_j$ ate the current estimates of i and j respectively) then the force $F_{i,j}$ in the direction $\hat{v}_{i,j}$ is given by

$$F_{i,j} = \hat{v}_{i,j} (\hat{d}_{i,j} - r_{i,j}) \qquad\qquad (9)$$

And the resultant force on node i is given by

$$F_i = \sum_{i,j} F_{i,j} \qquad\qquad (10)$$

The energy $E_{i,j}$ of nodes n_i and n_j due to the difference in measured and estimated distances is the sequence of the magnitude of $F_{i,j}$ and the total energy of node i is equal to

$$E_i = \sum_j E_{i,j} = \sum_j (\hat{d}_{i,j} - r_{i,j})^2 \qquad\qquad (11)$$

And the total energy of the system E is given by

$$E = \sum_i E_i \qquad\qquad (12)$$

Now the energy E_i of each node n_i reduces when it moves by an infinitesimal amount in the direction of force F_i. In the optimization, the magnitude of F_i for each node n_i is zero and the global energy of the system E is also zero and the algorithm converges.

Extensive simulations show that the proposed algorithm outperforms incremental algorithm by both being able to converge to correct positions and by being significantly more robust to errors in local distance estimate [13]. The limitation of this approach is that the algorithm is susceptible to local minima.

4.2.2.2. Cooperative Ranging Approach

In [13] the authors describe a Cooperative ranging approach which uses Assumption Based Coordinate (ABC) as its primitive to solve the localization problem. ABC algorithm determines the location of the unknown nodes by making assumptions when necessary and

compensating the errors through corrections and redundant calculations as more information becomes available. The algorithm starts with the assumption that node n_0 is located at (0, 0, 0). n_1 is the first node to establish communication with n_0 and is assumed to be located at (r_{01}, 0, 0), where r_{01} is the RSSI distance between n_0 and n_1. The location of the next node n_2 (x_2, y_2, z_2) can be obtained on the basis of two assumptions: y_2 is positive and $z_2 = 0$, so

$$x_2 = (r_{01}^2 + r_{02}^2 + r_{12}^2)/\ 2r_{01} \qquad\qquad (13)$$
$$y_2 = \surd\ (\ r_{02}^2 - x_2^2) \qquad\qquad\qquad (14)$$

Next location of n_3 (x_3, y_3, z_3) can be determined by assuming $z_3 = 0$, so

$$x_3 = (r_{01}^2 + r_{03}^2 + r_{13}^2)/\ 2r_{01} \qquad\qquad (15)$$
$$y_3 = (r_{03}^2 - r_{23}^2 + x_2^2 + y_2^2 - 2x_2x_3)/\ 2y_2 \qquad (16)$$
$$z_3 = \surd\ (\ r_{03}^2 - x_3^2 - y_3^2) \qquad\qquad\qquad (17)$$

From this point onwards the system of equations used to solve for further is no longer underdetermined and so the standard algorithm can be applied for each node and its neighbours.

Next the authors propose a cooperative ranging approach that exploits the high connectivity of the network to translate the global positioning challenge into a number of distributed local positioning problems that iteratively converge to a global solution by interacting with each other. In the proposed approach, every single node plays the same role repetitively and concurrently executes the following functions:

- Receive ranging and location information from neighbouring nodes.
- Solve the local localization problem by ABC algorithm.
- Transmit the obtained results to the neighbouring nodes.

After some repetitive iteration the system will converge to a global solution.
The advantage of this approach is that no global resources or communications are needed. The disadvantage is that convergence may take some time and that nodes with high mobility may be hard to cover.

4.2.3. *Coordinate System Stitching*

4.2.3.1. *Cluster based Approach*
In [14] the authors propose a distributed algorithm for locating nodes in a sensor network in which the nodes have the ability to estimate the distance to nearby nodes. Before describing the algorithm we need to know the distinction between non-rigid and rigid graphs. Non-rigid graphs can be continuously deformed to produce on infinite number of different realization, while rigid graphs cannot. However, in rigid graphs there are two types of discontinuous deformations that can prevent a realization from being unique.
- Flip ambiguities occur for a graph in a d-dimensional space when the positions of all neighbours of some vertex span a (d-1) dimensional subspace. In this case, the

neighbours create a mirror through which the vertex can be reflected. As shown in Fig. 8(a) vertex A can be reflected across the line connecting B and C with no change in distance constraints.

- Discontinuous flex ambiguities occur when the removal of one edge allows part of the graph to be flexed to a different configuration and the removal edge reinserted with the same length. As in Fig. 8(b) first AD is removed and then reinserted, the graph can flex in the direction of arrow, taking on a different configuration but preserved all distance constraints.
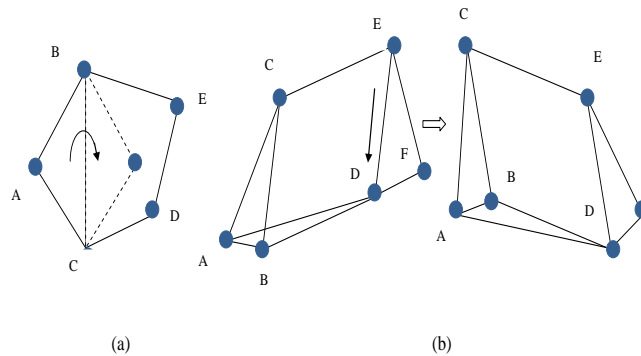

Fig.8. (a) Flip ambiguity (b) Discontinuous Flex ambiguity [14]

The algorithm is basically consists of two phases. Phase 1 is *cluster localization* where each node becomes the centre of the cluster and estimates the relative location of its neighbours which can be unambiguously localized. For each cluster, all the robust quadrilaterals as well as the largest sub graph composed solely of overlapping robust quads are identified. The authors define robust triangles to be a triangle which satisfies

$$B \sin^2 \theta > d_{min} \qquad\qquad (18)$$

In equation (18), b is the length of the shortest side and θ is the smallest angle and d_{min} is the threshold based on the measurement noise. If a quadrilateral has four robust sub-triangles then the quadrilateral is a robust quadrilateral. The algorithm starts with a robust quadrilateral and when two quads have three nodes in common and the first quad is fully localized, the second quad can be localized by trilaterating from the three known positions.

In the second phase i.e. *cluster transformation*, the position of each node in each local coordinate system are shared. As long as there are at least three non-collinear nodes in common between the two localizations, the transformation can be computed by rotation, translation, reflection.

The advantage of this scheme is that cluster based localization supports dynamic node insertion and mobility. The limitation is that under condition of low node connectivity or high measurement noise, the algorithm may be unable to localize a useful number of nodes.

4.2.3.2. Construction of Global Coordinate System in a network of Static Computational Nodes from Inter Node Distance

In [15] the authors propose an algorithm which is based on coordinate system stitching which constructs a spatial map and a distance matrix and then tries to minimize the discrepancies between them by translation, rotation and reflection. The distance matrix is explained with the help of Fig. 9. and Fig. 10. In Fig. 9 a collection of nodes and estimates of distances between some pairs of these nodes has been shown. A distance matrix of an individual node may acquire some subset of the distance estimates. So the distance matrix for node 2 is shown in Fig. 10(a). The distance matrix of two different nodes may overlap as shown in Fig. 10(b). Now to construct the spatial map from a distance matrix we need to construct an initial map containing a triangle of three non-collinear pair-wise neighbouring nodes. Then more nodes are inserted into the map, one at a time, based on distances to nodes already in the map, in an iterative process so that the node must have at least three non-collinear neighbour nodes. The process terminates when all nodes are inserted into the map or when no uninserted node can be inserted.
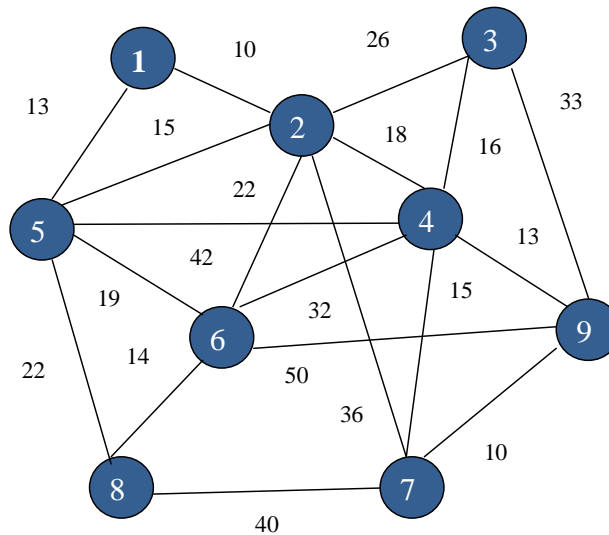

Fig.9. Network Topology with inter node distances between nodes [15]
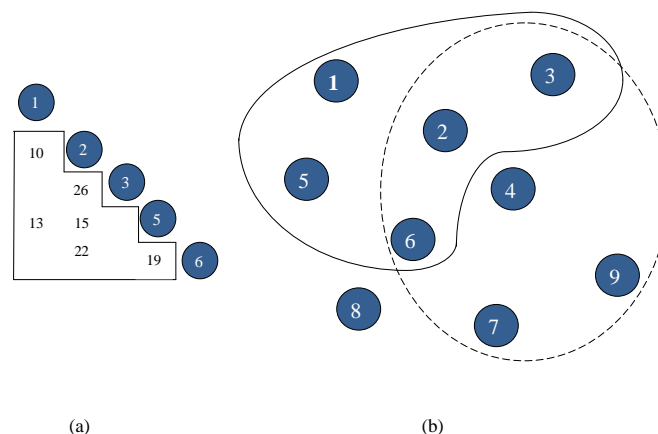

(a)   (b)
Fig.10. (a) Distance Matrix of Node 2 (b) Overlapping local maps [15]

Now to compute the initial map we need to find the longest side and denote it's end node as p and r and then allign this side with x axis by setting p's position to (0,0) and r's position to $(D_{pr}, 0)$. Then choose any third node q whose position is (x, y) where x=(D_{pq}^2 + D_{pr}^2 +

D_{qr}^2)/2 D_{pr} and y=$\surd$(D_{pr}^2-x^2) (as shown in Fig. 11). Next at each iteration, a node with highest number of neighbours already in the maps is chosen for insertion and the process will stop when no remaining unmapped node can be found with at least three mapped neighbours that are non-collinear.
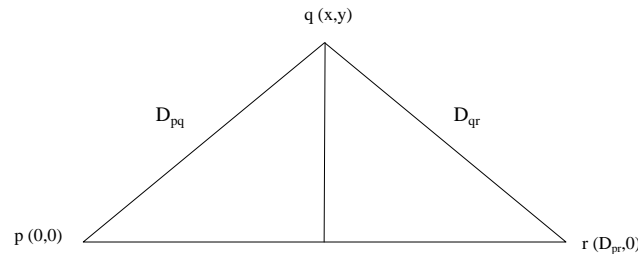
$$x^2 + y^2 = D_{pq}^2, \quad (D_{pr} - x)^2 + y^2 = D_{qr}^2$$

Fig.11. Initial map [15]

Next the authors discusses the process of reconciling two maps that have at least some nodes in common but that differ on the position of those common nodes by rotation, translation and reflection. When a node has sufficient distance estimates, it locally broadcasts a map of its neighbourhood. When a node receives a map from a neighbour, it reconciles its own map with its neighbour's and broadcasts its own map. In this way, each node should quickly acquire a map of its neighbourhood. Eventually, this agreement should spread throughout the network so that a common coordinate system is formed.

The advantage of this scheme is that it does not need anchor or beacon nodes for localization. But in traditional communication model, where nodes can communicate only with neighbors, this algorithm may converge quite slowly since a single coordinate system must propagate from its source across the entire network.

4.2.4. *Hybrid Localization*

4.2.4.1. *Localization scheme composed MDS and PDS*

In [16] the authors present a localization scheme composed of two localization techniques: multidimensional scaling (MDS) and proximity based map (PDM). At first some anchors are deployed denoted by primary anchors. In the first phase, some sensors are selected as secondary anchors which are localized through multidimensional scaling. Nodes which are neither primary anchors nor secondary one are called normal sensors. In the second phase, the normal sensors are localized through proximity distance mapping.

In the first stage each primary anchor sends an invitation packet containing its unique ID, a counter initialized to zero and a value k_s controlling the number of secondary anchors, to *one* of its neighbors. Normal sensor receiving this packet will perform a Bernoulli trial with a success rate of p. If the outcome is true, the normal sensor increments the counter by one and becomes a secondary anchor. The packet will be forwarded to another neighbor until the counter equals to k_s. After sending the invitation packet, each primary anchor sends packets containing its unique ID and coordinates to *all* of its neighbors. The packet also bears a field marking the proximity, i.e. the distance or hop count the packet has travelled. The value is

initialized to be zero. Secondary anchors will also do what primary anchors do, sending out packets with its unique ID but leaving the coordinate field blank. Every node (including anchors) receiving a proximity packet from an anchor (either primary or secondary) will store its ID and the proximity value. If a packet from a particular anchor has been received before, the node examines the proximity and check whether it is larger than the stored proximity. If it is larger than the stored value, the packet will be discarded. Otherwise, the stored value and the proximity field of the packet will be updated and the packet will be forwarded to other neighbors. Thus the stored proximity always reflects the shortest path distance or hop count from a particular anchor. After an anchor has discovered its proximities to all anchors, it will send the proximities it has collected to other anchors and wait for other anchors to do the same thing. When all anchors distribute the proximities to their counterparts, each anchor knows the proximity information between every pair of anchors. Each secondary anchor can now determine its location through classical MDS.

After the first phase, each secondary anchor also knows the position estimates of other secondary anchors as MDS provides a configuration about the primary and secondary anchors and calculates the proximity distance mapping. The mapping and the position estimates of secondary anchors obtained from the first phase are distributed to the normal nodes nearby. Normal sensor node uses the mapping to process the proximity vector it has stored when it aided anchors exchanging proximity information. Finally, the node position is calculated by multilateration with the processed proximity vector and the position information of primary and secondary anchors.

The main advantage of this scheme is to minimize the computation cost. For classical MDS, the complexity is $O(n^3)$ where n is number of nodes. The complexity for PDM is $O(m^3)$ where m is the number of anchors. But the scheme composed of MDS and PDM has a complexity of $O(m_x^3)$ where m_x is the total number of primary and secondary anchors. So by keeping m_x as a reasonable number, the complexity can be made similar to the complexity of PDM. The limitation of this scheme is that it does not perform well when there are only a few anchors.

4.2.4.2. *Simple Hybrid Absolute-Relative Positioning (SHARP)*

In [17] the authors present a localization scheme refers to as: Simple Hybrid Absolute-Relative Positioning (SHARP) which uses multidimensional scaling (MDS) and Ad-hoc Positioning System (APS) for localization. The localization scheme consists of three phases. In the first phase a set of reference nodes are selected randomly or along the outer perimeter of the network. In the second phase a relative localization method MDS is used to relatively localize the reference nodes selected in first phase. At first shortest-path distance between each pair of reference nodes are computed and then MDS is applied to construct the relative map. The result of first and second phases is a set of nodes with known coordinates according to some coordinate system. In third phase, an absolute localization method APS is used to localize the rest of the nodes in the network using the reference nodes as anchors. Each node uses the shortest-path distance information to estimate its distances to anchors. Then, it performs multilateration to estimate its position.

SHARP outperforms MDS if both the localization error and the cost are considered. The

limitation of this scheme is that for anisotropic networks SHARP gives poor performance.

4.2.4.3. *Localization scheme composed inductive and deductive approach*

In [27] the authors present a localization scheme for indoor environment. There are two main methods to estimate the position in indoor environments. On the one hand, there are the so-called deductive methods. These take into account the physical properties of signal propagation. They require a propagation model, topological information about the environment, and the exact position of the base stations. On the other hand, there are the so-called inductive methods. These require a previous training phase, where the system learns the signal strength in each location. The main shortcoming of this approach is that the training phase can be very expensive. The complex indoor environment makes the propagation model task very hard. It is difficult to improve deductive methods when there are many walls and obstacles because deductive methods work estimating the position mathematically with the real measures taken directly from environment in the training phase. In [27] the authors present a hybrid location system using a new stochastic approach which is based on a combination of deductive and inductive methods.

The advantage of this method covers a hard indoor environment without many base stations. Besides that, this technique reduces the training phase without losing precision.

4.2.5. *Interferometric Ranging Based Localization*

The idea behind the Radio Interferometric Positioning System (RIPS) proposed in [18], [19], [20] is to utilize two transmitters to create the interference signal directly. If the frequencies of the two emitters are almost the same then the composite signal will have a low frequency envelope that can be measured by cheap and simple hardware readily available on a WSN node. But due to the lack of synchronization of the nodes there will be a relative phase offset of the signal at two receivers which is a function of the relative positions of the four nodes involved and the carrier frequency. By making multiple measurements it is possible to reconstruct the relative location of the nodes in 3D. But localization using interferometric ranging is an NP-Complete problem [20]. To optimize the solution globally [18] uses genetic algorithm approach whereas [19] reduces the search space with additional RSSI readings.

Compared to the more common techniques such as received signal strength, time of arrival, and angle of arrival ranging, interferometric ranging has the advantage that the measurement could be highly precise. But localization using interferometric ranging requires a considerably larger set of measurements which limits their solutions to smaller networks (16 nodes in [18] and 25 nodes in [19]). To solve this problem an iterative algorithm has been proposed in [20] which calculates node locations from a set of seeding anchors and gradually builds a global localization solution. Compared to [18] and [19], which treat localization as a global optimization problem, the iterative algorithm is a distributed algorithm that is simple to implement in larger networks.

4.2.6. *Error Propagation Aware Localization*

An error propagation aware (EPA) algorithm has been proposed in [21] which integrates

the path loss and distance measurement error model. In the start of the algorithm, anchor nodes broadcast their information which includes their unique ID, global coordinates, and the position error variance σ_p^2. Each node senses the channel and records the TOA information to each anchor. The power of the detected direct path is translated to a ranging variance σ_r^2. After getting σ_r^2 and σ_p^2 the sensor node formulates weighting matrix given by equation (19).

$$W = W_r + W_p \qquad\qquad\qquad (19)$$
$$W_r = \mathrm{diag}(\sigma_{r1}^2,\ldots\ldots,\sigma_{rn}^2) \text{ and } \quad W_p = \mathrm{diag}(\sigma_{p1}^2,\ldots\ldots,\sigma_{pn}^2)$$

for n range measurement to anchors. In the next stage the node computes its position by incorporating its weighting matrix into Weighted Least Square (WLS) algorithm. After getting its own position the sensor node becomes an anchor and starts broadcasting its ID, global coordinate and σ_p^2. This process is repeated until all the nodes obtain their position and transformed into anchors.

The algorithm takes advantage of the ranging and position information obtained from each involved anchor and so it produces precise estimation than other localization schemes.

5. Summary Of Proposals

The performance of any localization algorithm depends on a number of factors, such as anchor density, node density, computation and communication costs, accuracy of the scheme and so on. All approaches have their own merits and drawbacks, making them suitable for different applications.

Some algorithms require beacons (Diffusion, Bounding Box, Gradient, APIT) and some do not (MDS-MAP, Relaxation based localization scheme, Coordinate system stitching). Beaconless algorithms produce relative coordinate system which can optionally be registered to a global coordinate system. Sometimes sensor networks do not require a global coordinate system. In these situations beaconless algorithms suffice.

Certain algorithms are centralized while some are distributed. Centralized algorithms generally compute more accurate positions and can be applicable to situations where accuracy is important. Distributed algorithms on the other hand do not depend on large centralized system and potentially have better scalability.

Beside these factors battery life and communication costs are also important for sensor networks. Generally centralized algorithms the communication costs are high to move data back to the base station. But the accuracy is also high in centralized schemes than the distributed approaches. Moreover some schemes perform well in high anchor density while some need only few anchors. As shown in [25], multilateration has low computation and communication cost and performs well when there are many anchors. On the other hand MDS-MAP has higher computation and communication cost and performs well when there are few anchor nodes.

The different schemes reviewed in this article are summarized in Table 1.

Table 1. Summary of proposals for Localization in WSN

PROPOSALS	OBJECTIVE	CENTRALIZED/ DISTRIBUTED	DESCRIPTION (IN BRIEF)	ACCURACY	COST (MESSEGING & COMPUTATION)
SHANG, RUML, ZHANG, FROMHERZ [5]	Present an MDS-MAP algorithm that uses connectivity information to derive the locations of the nodes in the network.	Centralized	Presented an algorithm MDS-MAP that uses an all-pairs shortest-paths algorithm to roughly estimate the distance between each possible pair of nodes. Then they have used MDS to derive node locations that fit those estimated distances and finally they have normalized the resulting coordinates to take into account any nodes whose positions are known.	High	High
ANUSHIYA A KANNAN, GUOQIANG MAO BRANKA VUCETIC [6]	Implement a two phase localization method based on the simulated annealing technique which also takes account the error due to flip ambiguity.	Centralized	Proposed a two-phase simulated annealing based localization algorithm, where an initial location estimate is obtained in the first phase using the simulated annealing technique and the large error due to flip ambiguity is mitigated in the refinement phase using neighborhood information of nodes.	High	High
CESARE ALIPPI, GIOVANNI VANINI [7]	Present a multi-hop localization technique for WSNs exploiting acquired received signal strength indications.	Centralized	The RSSI values of the packets exchanged among nodes at different power levels are collected (RF mapping phase) and processed both to build the ranging model to be fed into a centralized Minimum Least Square (MLS) algorithm.	High	High
T. HE, C. HUANG, B. BLUM, J. STANKOVIC, T. ABDELZAHER [8]	Describe a range free algorithm to make the scheme cost effective than range based approaches.	Distributed	Test whether the anchor node is inside or outside the triangles made by three anchors and then by utilizing combination of anchor positions, the diameter of the estimated area in which the node resides can be reduced to provide good localization estimate.	Low	Low
A. SAVVIDES, H. PARK, M. SRIVASTAVA [9]	Present a collaborative multilateration approach which enables ad hoc deployed sensor nodes to accurately estimate their locations by using known beacon locations that are several hopes away and distance measurements to neighbouring nodes.	Distributed	Presented a collaborative multilateration approach in which all nodes obtain a bounding box of region where the nodes lie. Next the error estimates are minimized by Kalman Filter implementation.	Low	Low

PROPOSALS	OBJECTIVE	CENTRALIZED/ DISTRIBUTED	DESCRIPTION (IN BRIEF)	ACCURACY	COST (MESSEGING & COMPUTATION)
S. SIMIC AND S. SASTRY [10]	Present a distributed algorithm for localization of nodes in a discrete model of random ad hoc network.	Distributed	Divide the whole region in a number of square cells. Each unknown node will send *hello* packets to its neighbours and based on the response the beacons, the unknown nodes can update their position estimates.	Low	Low
R. NAGPAL, H. SHROBE, AND J. BACHRACH [11]	Discover the position information of ad hoc wireless networks even when the elements have literally been sprinkled over the terrain.	Distributed	Each seeds produces propagating gradient that allows the sensors to estimate their distance from the seed. Next the sensors combine all the distance estimates to produce their own positions.	Low	Low
N. PRIYANTHA, H. BALAKRISHNAN, E. DEMAINE, S. TELLER [12]	Propose an anchor free solution to the problem of node localization where node starts from a random initial coordinate assignment and converge to a consistent solution using only local node interactions.	Distributed Algorithm	The algorithm models the nodes as point masses connected with springs and use-force directed relaxation method to converge towards a minimum energy configuration.	Low	Low
C. SAVARESE, J. RABAEY, J. BEUTEL [13]	Implement a cooperative ranging approach to get rid of the burden of using beacons in wireless sensor networks.	Distributed Algorithm	Proposed a cooperative ranging scheme which transmits the local information of each node to the neighbouring nodes and by iteratively doing this results in a convergence to global solution.	Low	Low
DAVID MOORE, JOHN LEONARD, DANIELA RUS [14]	Localize sensor nodes in a region by the use of robust quadrilaterals.	Distributed Algorithm	Each node measures the distances of the neighbouring nodes and form a cluster in some local coordinate system. In the next phase coordinate transformation can be computed to stitch the clusters in a global coordinate system.	Low	Low
LAMBERT MEERTENS AND STEPHEN FITZPATRICK [15]	Construct a Global Coordinate System in a network of static computational nodes from inter-node distances.	Distributed Algorithm	Each node in the network first construct a spatial map of its own position based on inter-node distance information and then each node reconcile its own map with those of nearby nodes and forms a global coordinate system.	Low	Low

PROPOSALS	OBJECTIVE	CENTRALIZED/ DISTRIBUTED	DESCRIPTION (IN BRIEF)	ACCURACY	COST (MESSEGING & COMPUTATION)
KING-YIP CHENG, KING-SHAN LUI AND VINCENT TAM [16]	Present a distributed algorithm which composed of MDS and PDM.	Distributed	In the first phase, some sensors are selected as secondary anchors which are localized through MDS. In the second phase, the normal sensors are localized through proximity distance mapping.	High	Low
A. A. AHMED, H. SHI, AND Y. SHANG [17]	Proposed a localization scheme by using MDS and APS.	Distributed	First MDS is used to relatively localize reference nodes. Then, APS uses these nodes as anchors to localize the rest of the nodes.	High	Low
J.LIORET, J.THOMAS, M. GARCIA, A. CANOVAS [27]	Propose a localization scheme by using inductive and deductive approach.	Distributed Algorithm	The algorithm uses a probabilistic model using both deductive and inductive approach.	High	Low
M. MAROTI, B. KUSY, G. BALOGH, P. V OLGYESI, A. NADAS, K. MOLNAR, S. DORA, AND A. LEDECZI [18]	Propose a localization scheme based on Interferometric ranging.	Distributed Algorithm	The algorithm uses a genetic algorithm approach to optimize the localization problem using interferometric ranging.	High	High
N. PATWARI AND A. O. HERO [19]	Propose a localization scheme based on interferometric ranging.	Distributed Algorithm	Proposed an interferometric ranging based localization which first estimates pairwise distances and then uses them to estimate coordinates via distributed weighted multidimensional scaling.	High	High
RUI HUANG, GERGELY V. ZARUBA, AND MANFRED HUBER [20]	Propose a localization scheme based on interferometric ranging for larger networks.	Distributed Algorithm	Proposed an iterative algorithm to solve the localization problem using interferometric ranging.	High	Low
N. A. ALSINDI, K. PAHLAVAN, AND B. ALAVI [21]	Present an error propagation aware routing for localization in WSNs.	Distributed Algorithm	Each node gets position error variance and ranging variance from some anchors and formulates a weighting matrix. Next each node calculates its position by incorporating the weighting matrix into WLS.	High	Low

6. Open Problems

There are considerable amount research activities to improve localization in wireless sensor networks. But there are also some interesting open problems that need further attention.

Interferomatric ranging based localization that takes error propagation into account: Interferometric ranging technique has been recently proposed as a possible way to localize sensor networks as it gives precise measurements than other common techniques. But

simulation results from [20] indicate that error propagation can be a potentially significant problem in interferometric ranging. In order to localize large networks using interferometric ranging from a small set of anchors, future localization algorithms need to find a way to effectively limit the error propagation.

Robust algorithm for mobile sensor networks: Recently there has been a great deal of research on using mobility in sensor networks to assist in the initial deployment of nodes. Mobile sensors are useful in this environment because they can move to locations that meet sensing coverage requirements. New localization algorithms will need to be developed to accommodate these moving nodes. So, devising a robust localization algorithm for next generation mobile sensor networks is an open problem in future.

Attack the challenges of Information Asymmetry: WSNs are often used for military applications like landmine detection, battlefield surveillance, or target tracking. In such unique operational environments, an adversary can capture and compromise one or more sensors physically. The adversary can now tamper with the sensor node by injecting malicious code, forcing the node to malfunction, extracting the cryptographic information held by the node to bypass security hurdles like authentication and verification, so on and so forth. In a beacon-based localization model, since sensor nodes are not capable of determining their own location, they have no way of determining which beacon nodes are being truthful in providing accurate location information. There could be malicious beacon nodes that give false location information to sensor nodes compelling them to compute incorrect location. This situation, in which one entity has more information than the other, is referred to as *information asymmetry*. To solve this problem, in [22] the authors propose a Distributed Reputation-based Beacon Trust System (DRBTS), which aimed to provide a method by which beacon nodes can monitor each other and provide information so that unknown nodes can choose who to trust, but future research work is needed in this field.

Finding the minimum number of Beacon locations: Beacon based approaches requires of a set of beacon nodes, with known locations. So, an optimal as well as robust scheme will be to have a minimum number of beacons in a region. Further work is needed to find the minimum number of locations where beacons must be placed so the whole network can be localized with a certain level of accuracy.

Finding localization algorithms in three dimensional space: WSNs are physical impossible to be deployed into the area of absolute plane in the context of real-world applications. For all kinds of applications in WSNs accurate location information is crucial. So, a good localization schemes for accurate localization of sensors in three dimensional space can be a good area of future work.

So these are the few problems for future research work to improve localization in wireless sensor technology.

References

[1] J.Li, J. Jannotti, D. S. J. DeCouto, D. R. Karger and R. Morris, "A Scalable Location Service for Geographic Ad-Hoc Routing", in *Proceedings of Sixth Annual International Conference on Mobile Computing and Networking*, August 2000, Boston, Massachusetts, USA, pp. 120-130.

http://dx.doi.org/10.1145/345910.345931

[2] K. Amouris, S. Papavassiliou, M. Li, "A Position-Based Multi-Zone Routing Protocol for Wide Area Mobile Ad-Hoc Networks", in *Proceedings of IEEE Vehicular Technology Conference (VTC '99),* May 1999, Houston, Texas, USA, Vol. 2, pp.1365-1369. http://dx.doi.org/10.1109/VETEC.1999.780570

[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, November 2001. http://dx.doi.org/10.1109/65.967595

[4] M. B. Srivastava , R. Muntz and M. Potkonjak, "Smart Kindergarten: Sensor-based Wireless Networks for Smart Developmental Problem-solving Environments", In *Proceedings of Seventh Annual International Conference on Mobile Computing and Networking*, July 2001, Rome, Italy, pp. 132-138. http://dx.doi.org/10.1145/381677.381690

[5] Y. Shang, W. Ruml, Y. Zhang, and M. Fromherz, "Localization from mere connectivity", In *Proceedings of ACM Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc'03)*, June 2003, Annapolis, Maryland, USA, pp. http://dx.doi.org/201-212. 10.1145/778415.778439

[6] Anushiya A Kannan, Guoqiang Mao and Branka Vucetic, "Simulated Annealing based Wireless Sensor Network Localization", *Journal of Computers,* Vol. 1, No. 2, pp 15-22, May 2006.

[7] Cesare Alippi, Giovanni Vanini, "A RSSI-based and calibrated centralized localization technique for Wireless Sensor Networks", in *Proceedings of Fourth IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW'06),* Pisa, Italy, March 2006, pp. 301-305. http://dx.doi.org/10.1109/PERCOMW.2006.13

[8] T. He, C. Huang, B. Blum, J. Stankovic, and T. Abdelzaher, "Range-free localization schemes in large scale sensor networks", In *Proceedings of the Ninth Annual International Conference on Mobile Computing and Networking (MobiCom'03),* September 2003, San Diego, CA, USA, pp. 81-95. http://dx.doi.org/10.1145/938985.938995

[9] A. Savvides, H. Park, and M. Srivastava, "The bits and flops of the n-hop multilateration primitive for node localization problems", In *Proceedings of the 1st ACM international Workshop on Wireless Sensor Networks and Applications (WSNA'02),* September 2002, Atlanta, Georgia, USA, pp. 112-121. http://dx.doi.org/10.1145/570738.570755

[10] S. Simic and S. Sastry, "Distributed localization in wireless ad hoc networks", Technical Report UCB/ERL M02/26, UC Berkeley, 2002, Available HTTP: http://citeseer.ist.psu.edu/simic01distributed.html.

[11] J. Bachrach, R. Nagpal, M. Salib and H. Shrobe, "Experimental Results for and Theoritical Analysis of a Self-Organizing a Global Coordinate System from Ad Hoc Sensor Networks", *Telecommunications System Journal,* Vol. 26, No. 2-4, pp. 213-233, June 2004. http://dx.doi.org/10.1023/B%3ATELS.0000029040.85449.7b

[12] N. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller, "Anchor-free distributed localization in sensor networks", MIT Laboratory for Computer Science, Technical Report TR-892, April 2003, Available HTTP: http://citeseer.ist.psu.edu/681068.html.

[13] C. Savarese, J. Rabaey, and J. Beutel, "Locationing in distributed ad-hoc wireless sensor networks", in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'01)*, May 2001, Salt Lake City, Utah, USA, vol. 4, pp. 2037-2040. http://dx.doi.org/10.1109/ICASSP.2001.940391

[14] David Moore, John Leonard, Daniela Rus, and Seth Teller, "Robust distributed network localization with noisy range measurements", in *Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems (SenSys'04)*, November 2004, Baltimore, MD, pp. 50-61. http://dx.doi.org/10.1145/1031495.1031502

[15] Lambert Meertens and Stephen Fitzpatrick, "The Distributed Construction of a Global Coordinate System in a Network of Static Computational Nodes from Inter-Node Distances", Kestrel Institute Technical Report KES.U.04.04, Kestrel Institute, Palo Alto, 2004, Available FTP: ftp://ftp.kestrel.edu/pub/papers/fitzpatrick/LocalizationReport.pdf.

[16] King-Yip Cheng, King-Shan Lui and Vincent Tam, "Localization in Sensor Networks with Limited Number of Anchors and Clustered Placement", in *Proceedings of Wireless Communications and Networking Conference, 2007 (IEEE WCNC 2007)*, March 2007, pp. 4425 – 4429. http://dx.doi.org/10.1109/WCNC.2007.806

[17] A. A. Ahmed, H. Shi, and Y. Shang, "Sharp: A new approach to relative localization in wireless sensor networks," in *Proceedings of IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2005, Columbus, Ohio, USA. June 6-10, 2005. http://dx.doi.org/10.1109/ICDCSW.2005.125

[18] M. Maroti, B. Kusy, G. Balogh, P. V olgyesi, A. Nadas, K. Molnar, S. Dora, and A. Ledeczi, "Radio Interferometric Geolocation," in *Proceedings of 3rd International Conference on Embedded Networked Sensor Systems (SenSys)*, pp. 1-12, San Diego, California, USA, Nov. 2005. http://dx.doi.org/10.1145/1098918.1098920

[19] N. Patwari and A. O. Hero, "Indirect Radio Interferometric Localization via Pairwise Distances," *in Proceedings of 3rd IEEE Workshop on Embedded Networked Sensors (EmNets 2006)*, pp. 26-30, Boston, MA, May 30-31, 2006.

[20] Rui Huang, Gergely V. Zaruba, and Manfred Huber, "Complexity and Error Propagation of Localization Using Interferometric Ranging", *in Proceedings of IEEE International Conference on Communications ICC 2007*, pp. 3063-3069, Glasgow, Scotland, June 2007. http://dx.doi.org/10.1109/ICC.2007.509

[21] N. A. Alsindi, K. Pahlavan, and B. Alavi, "An Error Propagation Aware Algorithm for Precise Cooperative Indoor Localization", *in Proceedings of IEEE Military Communications Conference MILCOM 2006*, pp. 1-7, Washington, DC, USA, October 2006. http://dx.doi.org/10.1109/MILCOM.2006.302311

[22] A. Srinivasan, J. Teitelbaum, J. Wu. "DRBTS: Distributed Reputation-based Beacon Trust System", *2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC '06)*, September 2006, Indianapolis, USA, pp. 277-283. http://dx.doi.org/10.1109/DASC.2006.28

[23] J. Bachrach and C. Taylor, "Localization in Sensor Networks," in *Handbook of Sensor Networks: Algorithms and Architectures*, I. Stojmenovic, Ed., 2005.

[24] Yunhao Liu, Zheng Yang, Xiaoping Wang, and Lirong Jian, Location, localization, and localizability, Journal of Computer Science And Technology 25(2): 274-297 Mar. 2010. http://dx.doi.org/10.1007/s11390-010-9324-2

[25] Yi Shang, Hongchi Shi and A. Ahmed, "Performance study of localization methods for ad-hoc sensor networks", *in proceedings of IEEE International Conference of Mobile Ad-hoc and Sensor Systems,* Fort Lauderdale, FL, October 2004. http://dx.doi.org/10.1109/MAHSS.2004.1392106

[26] Eiman Elnahrawy, Xiaoyan Li and Richard P. Martin, "The Limits of Localization Using Signal Strength: A Comparative Study", *in Proceedings of IEEE SECON*, pp. 406-414, Santa Clara, California, USA, October 2004. http://dx.doi.org/10.1109/SAHCN.2004.1381942

[27] Jaime Lloret, Jesus Tomas, Miguel Garcia, Alejandro Canovas, "A Hybrid Stochastic Approach for Self-Location of Wireless Sensors in Indoor Environments", *Sensors* 9, no. 5: 3695-3712. http://dx.doi.org/10.3390/s90503695

Appendix: A. Multidimensional Scaling

In Multidimensional Scaling (MDS) a set of points whose position is unknown and measured distances between each pair of points are given. It can be used to localize sensor nodes in a network. Without anchors or GPS, MDS can solve for the relative coordinates of a group of sensor nodes with resilience to measurement error and rather high accuracy.

Let there be n sensors in a network, with positions X_i, i = 1 . . . n, and let $X = [X_1, X_2, \ldots, X_n]^T$. X is nxm, where m is the dimensionality of X.

Let $D = [d_{ij}]$ be the nxn matrix of pairwise distance measurements, where d_{ij} is the measured distance between X_i and X_j for $i \neq j$, and $d_{ii} = 0$ for all i. The distance measurements d_{ij} must obey the triangular inequality: $d_{ij} + d_{ik} \geq d_{jk}$ for all (i, j, k).

Classical metric multidimensional scaling is derived from the Law of Cosines, which states that given two sides of a triangle d_{ij}, d_{ik} and the angle between them θ_{jik}, the third side can be computed using the formula:

$$d^2_{jk} = d^2_{ij} + d^2_{ik} - 2d_{ij}d_{ik} \cos \theta_{jik} \qquad (A.1)$$
$$\rightarrow d_{ij}d_{ik} \cos \theta_{jik} = 1/2(d^2_{ij} + d^2_{ik} - d^2_{jk}) \qquad (A.2)$$
$$\rightarrow (X_j - X_i) \cdot (X_k - X_i) = 1/2(d^2_{ij} + d^2_{ik} - d^2_{jk}) \qquad (A.3)$$

Next choose some X_0 from X to be the origin of a coordinate system, and construct a matrix $B_{(n-1)x(n-1)}$ as follows:

$$b_{ij} = 1/2(d^2_{0i} + d^2_{0j} - d^2_{ij}) \qquad (A.4)$$

Let $X'_{(n-1)xm}$ the matrix X where each of the X_i's is shifted to have its origin at X_0: $X'_i = X_i - X_0$. Then, using equations (A.3) and (A.4):

$$X'X'^T = B \qquad (A.5)$$

We can solve for X' by taking an eigen decomposition of B into an orthonormal matrix of eigenvectors and a diagonal matrix of matching eigenvalues:

$$B = X'X'^T = UVU^T \qquad (A.6)$$
$$X' = UV^{1/2} \qquad (A.7)$$

The problem is that X' has too many columns and we need to find X in 2-space or

3-space. To do this, we throw away all but the two or three largest eigenvalues from V , leaving a 2x2 or 3x3 diagonal matrix, and throw away the matching eigenvectors (columns) of U, leaving $U_{(n-1)x2}$ or $U_{(n-1)x3}$. Then $X^{/}$ has the proper dimensionality.

Appendix: B. Kalman Filter

The Kalman filter is a recursive estimator. This means that only the estimated state from the previous time step and the current measurement are needed to compute the estimate for the current state. In what follows, the notation $\hat{x}_{n|m}$ represents the estimate of x at time n given observations up to and including time m.

The state of the filter is represented by two variables:

- $\hat{x}_{k|k}$, the estimate of the state at time k given observations up to and including time k;
- $P_{k|k}$, the error covarience matrix (a measure of the estimated accuracy of the state estimate).

The Kalman filter has two distinct phases: *Predict* and *Update*. The predict phase uses the state estimate from the previous timestep to produce an estimate of the state at the current timestep. In the update phase, measurement information at the current timestep is used to refine this prediction to arrive at a new more accurate state estimate, again for the current timestep.

Predict:

Predicted state is given by $\hat{x}_{k|k} = F_k \, \hat{x}_{k-1|k-1} + B_k u_k$ and Predicted estimate covarience is given by $P_{k|k-1} = F_k P_{k-1|k-1} F^T_k + Q_{k-1}$ where

- F_k is the state transition model which is applied to the previous state x_{k-1};
- B_k is the control-input model which is applied to the control vector u_k;
- W_k is the process noise which is assumed to be drawn from a zero mean multivariate normal distribution with covariance Q_k i.e. $w_k \sim N(0, Q_k)$.

Update:

Innovation or measurement residual $\tilde{y}_k = z_k - H_k \, \hat{x}_{k|k-1}$ where at time k an observation (or measurement) z_k of the true state x_k is made according to

$z_k = H_k x_k + v_k$

where H_k is the observation model which maps the true state space into the observed space and v_k is the observation noise which is assumed to be zero mean Gaussian white noise with covariance R_k i.e. $v_k \sim N(0, R_k)$.

Innovation (or residual) covariance $S_k = H_k P_{k|k-1} H^T_k + R_k$

Kalman Filter gain $K_k = P_{k|k-1} H^T_k S_k^{-1}$

Updated state estimate $\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k$

Updated estimate covariance $P_{k|k} = (I - K_k H_k) P_{k|k-1}$