Localization through Anchor and non-Anchor Nodes in Wireless Sensor Network using Ant Colony Optimization

MS Research Thesis

By

Waqas Ahmad

557-FBAS/MSCS/F09

Supervised by

Dr. Muhammad Zubair

Department of Computer Science and Software Engineering

Faculty of Basic and Applied Sciences

International Islamic University Islamabad

2012

بسم الله الرحمن الرحيم

Faculty of Basic & Applied Sciences, Department of Computer Science & Software Engineering

International Islamic University, Islamabad

FINAL APPROVAL

Dated: 19th/Oct/2012

It is certified that we have read the thesis, entitled **"Localization through Anchor and non-Anchor Nodes in Wireless Sensor Network using Ant Colony Optimization"**, submitted by Mr. Waqas Ahmad Reg. No. 557-FBAS/MSCS/F09. It is our judgment that this thesis is of sufficient standard to warrant its acceptance by the International Islamic University Islamabad for MS Degree in Computer Science.

THESIS EVALUATION COMMITTEE

External Examiner:

Dr. Aqeel Abbas

Assistant Professor

Department of Electronics,

Quaid-i-Azam University, Islamabad

Internal Examiner:

Prof. Dr. Muhammad Sher

Dean, Department of CS & SE

International Islamic University, Islamabad

Supervisor:

Dr. Muhammad Zubair

Assistant Professor

Chairman, FET

International Islamic University, Islamabad

Dedication

I dedicate this thesis to my beloved Ami and Abu. They sacrificed all their life for me, my brothers and sisters to see us successful in all the fields of life.

I could not find words to present them reward of all they have done for me. Any way I thank them for all their honest efforts and prayers. May Almighty ALLAH bless them beautiful and long life.
AMEEN

A dissertation submitted to the

Department of Computer Science,

International Islamic University, Islamabad

As a partial fulfillment of the requirements

For the award of the degree of

MS in Computer Science

Declaration

I hereby declare that this dissertation, neither as a whole nor a part thereof has been copied out from any source. It is further declared that no portion of the work presented in this thesis has been submitted in support of any application for any other degree or qualification of this or any other university or institute of learning.

Waqas Ahmad

557/FBAS/MSCS/F09

Acknowledgement

This dissertation and all my efforts are fruitful only due to the blessings of Allah Almighty, the Most Merciful and Beneficent Who gave me strength to complete this task to the best of abilities and knowledge.

I would like to take this opportunity of thanking my supervisor **Dr. Muhammad Zubair** and to all my teachers in the whole educational carrier who shared their knowledge, guidance and support to boost my confidence and learning. He was always there to solve problems and helped me through this project. I am very much thankful to him for all this sincere efforts.

Last but not the least; I would like to acknowledge Mr. Usman Ghani who helped me a lot in this research thesis. I am thankful to Mr. Muhammad Yasir Shabir, Mr. Zahid Mahmood, Zaffar akbar chaudhary, Mr. Waqas Nasar, Mr. Rashid Abbasi and all my fellows who helped me in this thesis. I am also thankful to all the member of research committee for their precious guideline. All of them encouraged and provided me logistic and technical support. I would like to admit that I owe all my achievements to my truly, sincere and most loving parents, especially to my father, to my younger brother Usama, to my beloved wife, to all my family and friends who mean the most to me whose prayers have always been a source of determination for me.

Waqas Ahmad

557/FBAS/MSCS/F09

Abstract

This dissertation presents some contribution in WSN (Wireless Sensor Network) regarding the problem of sensor localization. Sensors are deployed in some environment for monitoring of any event; in real scenarios, usually this deployment is random. There are applications, which only demand for the transmission of the information without knowing the source location, but there are applications in which location of source should be known to take some necessary action in response. Sensor Localization means finding of deployed position of a senor. We propose a distributed approach for WSN localization. It works as: At first step, we measure distance between neighboring nodes, as our proposed approach is range based then we find initial estimate of sensor's position with the help of trilateration. Finally, we refined the initially estimated position using Ant Colony Optimization to get optimal solution. The final optimum solution will be the optimum coordinates of a node. We got better localization accuracy and comparatively less average error.

Contents

List of Figures

Chapter 1 Introduction

1.1 Introduction

In this thesis, we tackled the problem of localization. The process of finding the spatial location of nodes in a wireless network has been called localization, positioning, geolocation, and self-organizing in the literature [32]. Different approaches have been proposed in literature for localization.

In some applications it is very important to know the source of information. In typical scenarios the sensor deployment is uncontrolled or random. WSN localization has its importance such as for military purposes and many other applications where detection of sensor location is of much importance. It can be used for routing, enhance security [32] etc. In the Fig. 1.1 a scenario is presented relating to sensor positions. In it (Xi, Yi) denotes the positions that are already known and are called anchor or beacon nodes, (X? , Y?) are the coordinates of sensor nodes that are to be found yet. There are many type of techniques that have been used to find the position of unknown sensors. We have discussed many of them in the chapter of literature survey.

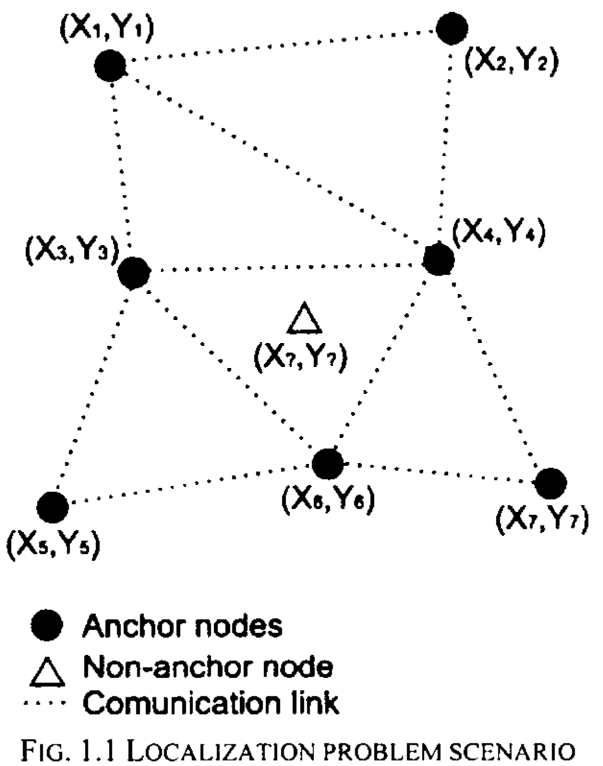


FIG. 1.1 LOCALIZATION PROBLEM SCENARIO

A **wireless sensor network (WSN)** consists of spatially scattered self-directed sensors to supervise some phenomena, which transmits gathered information to some main location (Base station or Sink etc). Due to the enhanced technology now, established networks provide bi-directional communication and controls sensor activities as well. Mobile sensor is one example of controlled activity. Initially WSN were designed for military purposes for battlefield observations. Afterword it moved towards various industrial and consumer applications.

Wireless Networks have now become the need of this age and provide very good data rates too. Wireless sensor networks (WSN) are multi-hop and self-organized networks, established with large number of micro sensor nodes, which are deployed in field for the monitoring of occurrence of some event [1].

1.2 WSN Architecture

The sensor nodes have capabilities to collect data and route that data back to sink by a multi-hop structure less architecture through the sink as shown in **Fig.1.2**. The sink may communicate with the task manager node via Internet or satellite for further processing of data that is received from deployed sensors. The design of the Wireless Sensor Network is influenced by many factors like false localization, fault tolerance, scalability, production costs, operating environment, sensor network topology, hardware constraints, transmission media, noise, power consumption etc.

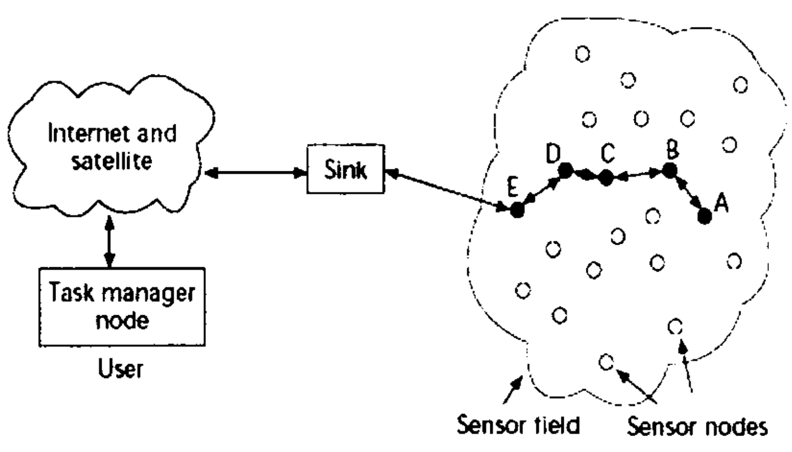


FIG. 1.2 SENSOR NODES SCATTERED IN A SENSOR FIELD [1]

Characteristics of WSN

- Broadcast in nature
- More security hazards than wired networks

- Less data rates than wired networks
- Easy remote access
- Easily effected by environmental factors
- WSN configure and organize themselves by their own
- Scalability of large scale deployment

Basic Architecture of a Sensor

A sensor node consists of four basic components as shown in **Fig.1.3** a sensing unit, a processing unit, a transceiver unit and a power unit. They may also have application dependent additional components such as a location finding system, a power generator and mobilizer etc. Sensing units are usually composed of two subunits: sensors and analog to digital converters (ADCs). The analog signals produced by the sensors based on the observed phenomenon are converted to digital signals by the ADC and then fed into the processing unit. The processing unit which is generally associated with a small storage unit, this processing unit manages the procedures that make the sensor node to collaborate with other nodes to carry out the assigned sensing tasks. A transceiver unit connects the node to the network. One of the most important components of a sensor node is the power unit. Power units may be supported by a power scavenging unit such as solar cells. There are also many other components which are application dependent.

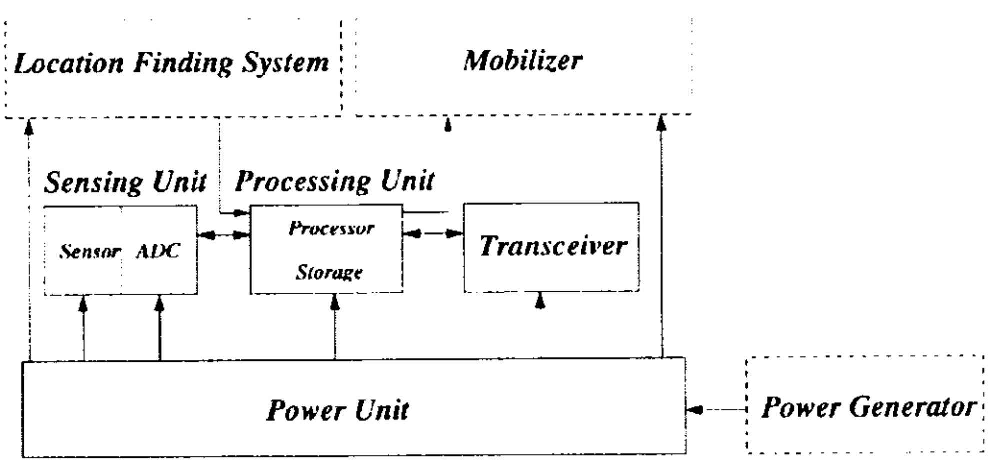


FIG. 1.3 COMPONENTS OF SENSOR NODE [1]

Basic Characteristics of a Sensor

- Usually battery powered which cannot be recharged

- Broadcast in nature
- Prone of failure etc.
- Easily effected by environmental conditions
- Limited in communication range
- Limited computation power
- Small in size
- Limited or no storage capacity

The sensor nodes are battery powered so they have issues of limited power, computational capacity and memory. Because of limited power capacity sensor nodes cannot transmit signals up to long distance. So to gain much longer distance coverage sensor nodes use multi-hops architecture to save energy.

Application of sensors

- Temperature
- Humidity
- Vehicular movement
- Pressure
- Noise Levels
- The presence or absence of different kinds of objects
- Mechanical stress levels on attached objects
- The current characteristics such as speed, direction and size of object
- Emergency response system
- Battle field etc.

1.3 Research Background of WSN Localization

In some types of sensor applications, we don't need to know the physical position of sensor node. In these types of application, the only thing that is important that is information. Whereas there are many applications where information provided by the sensor is not of important or meaningful until we know the physical position of sender of the information. To find the location of the source of information is called Localization problem in Wireless Sensor Network. In other words we can say that Localization refers to the finding of the physical position of deployed sensor.

1.4 Broad Categories of Sensor Deployment

1 Random/Uncontrolled Deployment
2 Controlled Deployment

In real scenarios deployment of sensor nodes is **Random** [27]. In this type of deployment strategy it is difficult to have any proper pre-plan. In such deployment sensors can be deployed by an aeroplane or by any other mean. There are many scenarios where we cannot have pre-planned deployment such as in the situation of war, flood etc.

In **Controlled deployment** [28] sensors are deployed by a pre-planned strategy. This deployment can be manual. In such s kind of sensor deployment, we knew the deployed position of sensor.

1.5 Motivation of Research

Until now there is a lot of work have been done on the topic of WSN localization. Many different techniques have been used to solve this problem. Broadly there are two approaches used for this first one is distanced based and the other which don't depend on distance estimation. The approaches based on distance are better in estimation of positions. Distanced based approaches are costly in terms of cost and time complexity. These approaches have hardware requirement but the real advantage is that they provide more accurate results in distance measurement than the approaches that do not rely on distance. The major problem encountered during localization is the error in position estimation of the nodes. A number of researchers have employed evolutionary techniques such as Q. Zhang et al. [14] used genetic algorithm to solve for localization in the centralized way. This approach have some drawbacks such as: if central node is compromised then whole burden is on one central node, a lot of communication is involved, timely complex etc.

In [15] genetic algorithm is also used for WSN localization. This approach was distributed one but faces a bit large localization error which cannot be tolerated in many typical applications.

There was the need of distributed approach with better localization accuracy and an algorithm with comparatively better performance. We proposed Ant colony Optimization algorithm for WSN Localization problem in distributed scenario. Genetic algorithm is used in our based paper. Genetic algorithm is a evolutionary optimization techniques which is based on Darwin theory whereas ant colony algorithm adopts the behavior of social insects. Ant colony Optimization algorithm is better in convergence [21, 22]. To the best of our knowledge we did not found ACO algorithm for WSN Localization problem especially when the nodes are static and approach is distributed.

1.6 Problem definition

There are many methodologies proposed in literature for localization in WSN as in our base paper genetic algorithm is used. Genetic algorithm took much time to get the optimal solution as the search space increase in exponantional way where as Ant Colony Optimization (ACO) [21, 22] does not. As we knew that the WSN has energy issue, as usually sensors can't be recharged and in this way dependency on any node will exhaust their battery very soon due to a lot of communication.

In **Fig. 1.4** we have shown a central approach problem scenario that is the approach of base paper [14]. Central approach work as: every node sends its collected information (distance) between their neighboring nodes to one central node. Centralized approach involves a lot of communication. Communication is costly as compared to computation that's why there should be an approach, which is distributed, and not dependent on one central node so the communication cost can be reduced. In central approach whole burden is on one central node which can be divided to all the nodes in the network using distributed approach. In distributed scenario all relevant communication is done on the nodes themselves [23]. In centralized approach if central node is compromised or down then remaining nodes in the network cannot be localized. Distributed approach is easily scalable [14] as compared to central approach. In base paper genetic algorithm is used for WSN localization where as we proposed Ant colony optimization as [21] shows that genetic algorithm is timely complex, in it there is easy adjustment of parameters than genetic algorithm [22] and better in accuracy [34]. Below is the problem scenario of centralized approach for localization:

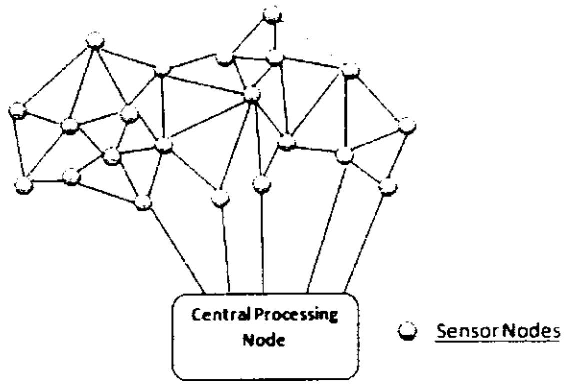


FIG. 1.4 CENTRALIZED APPROACH SCENARIO

1.7 Objective of Proposed Technique

Objective of proposed technique is to enhance the existing mechanism of locating the sensor nodes more efficiently with respect to time, to reduce mean location error, to achieve more accuracy while being aware of the time complexity, to minimize the feasible region that hopefully contain the optimum solution to our localization problem. Time complexity is directly proportional to the feasible region. The larger the feasible region, the more time it requires to find location of sensor.

1.8 Approaches regarding coordinate system

1. Global Approach
2. Local Approach

Global approach uses some external source to calculate position like GPS (Global Positioning System) to be aware of location. These approaches are more reliable than Local.

Local Approach does not use external source. It uses arbitrary rigid transformation like rotation, reflection, translation which is away from global coordinate system. [26]

1.9 Classification of Localization methods

1. Centralized vs Distributed
2. Anchor-free vs Anchor-based
3. Range-free vs Range-based
4. Mobile vs Stationary

In centralized systems all deployed sensors transmit their computations to one main (sink) node, which further performs calculation on the basis of information delivered from the sensors. After computations sink node replies back to the deployed nodes in the field using their node *id*'s. The drawback of this method is that it puts the entire burden on single central node and because the sensors are battery powered their life time will start to decline sharply due to excessive communication to central point. We knew that communication cost is several time of computation cost and it is also not timely efficient [29]. But one good thing of centralized system is that their algorithms are easy as compared to distributed systems and are more reliable too because central node have global view.

In distributed systems, every sensor is capable of performing computation and able to take decision. In distributed systems sensors are stand alone devices. In these systems computation is distributed among all the nodes. The distributed system reduces burden from one central node and they are easily scalable too. Algorithms are more complex in distributed systems than centralized but timely efficient because communication messages are not in queue to wait for processing at the central node [29].

Anchor-based and Anchor-free

Anchor or sink nodes are the nodes who knows their positions. They may know their positions by some external mean like GPS or by hard coded. It means that they have prior knowledge of their position. When these anchor nodes are deployed in some environment other non-anchor nodes uses the known coordinates of anchor nodes to localize themselves. This is known as **anchor-based** [30].

Anchor-free, here in this type of network there are no such nodes that knew about their positions. These nodes use relative coordinate and network connectivity information among other nodes to know their positions [31].

Range-based and Range-free

Classification of WSN Localization on distance estimation can be categorized into Range-based and Range-free.

Range-based approaches use some well known methods for distance measurement. This kind of approaches rely on distance or angle measurement than uses trilateration, triangulation, multilateration or other approaches to find the initial estimate of nodes. These approaches are hardware based so are costly but provide good results. When geographic position is of more importance then we should use range-based approaches to calculate distance [32].

Range-free approaches do not calculate distance or angle for localization of sensor/nodes. Although these approaches do not require additional hardware but the estimated results are not much reliable. Localization error in these methodologies is expected to be high. Range-free approaches use network connectivity information to localize the nodes [33].

Mobile and stationary

Nodes that can move from one place to another are mobile nodes and are expensive too. Such as a MS (Mobile Station) to which we want to localized with the help of BS (Base Station) in GSM. **Mobile** nodes change their position time-to-time that is why it is difficult to find/locate them.

A sensor node that is not capable of movement by itself is called **static** node. Static nodes are easier to find than mobile node because of no frequent displacement.

1.10 Organization of the thesis

In second chapter we discussed literature survey, Chapter 3 discussed proposed solution. Implementation and produced results are presented in chapter 4. Chapter 5 consists of conclusion and future work.

Chapter 2 Literature Survey

Wireless Sensor Network is applied in various fields [1]. Information transmitted by the sensors will be meaningless if we don't know the location of the deployed sensor node. Accurate self-localization capability is highly desirable in wireless sensor network [2]. Distance estimation in WSN Localization divided into two kinds Range based and Range-free [3]. In literature Range-based approaches used different algorithms for distance measurement like ToA, TDoA, AoA, RSSI where as in Range-free approaches dv-hop, Centroid, APIT etc. are used.

2.1 Range-based approaches

Time of Arrival: Sinan G *et al.* in [4] presented a two steps ranging approach based on time resolution of UWB (Ultra Wideband). Minimum Lower bound for ToA can be achieved but in the case of extremely high-rate sampling and large number of multipath component achievement but it was impractical. Algorithms in the literature were traditional correlation-based algorithms which performs serial search which causes delay in estimation of ToA for UWB signal. So there was need of some different approach. Authors in this paper proposed two steps to speed up estimation process which was slow due to serial search approaches. Algorithm firstly estimates the rough ToA on received signal strength (RSS). As the second

of the channel. In this regime, the amplitude and phase change imposed by the channel varies considerably over the period of use) and depending on some other factors. To cop step ToA of first signal was estimated by hypothesis testing approach (method of making decision using data whether from a controlled experiment or an observation study).In Low-rate correlation is used which is able to perform accurate ToA estimation within reasonable time factor. The future work included the optimization of the different estimator parameters.

Received Signal Strength: Approach in [5] computes distance based on receiving signal power at the receiver. If the transmission power is P_{tx}, the path loss model and the path loss coefficient α are known then at the receiver side then signal strength formula can be used to find distance as below.

$$P_{rcvd} = c\frac{P_{tx}}{d^{\alpha}} \leftrightarrow d = \sqrt[\alpha]{\frac{c P_{tx}}{P_{rcvd}}}$$

The advantage of this approach is that no additional hardware is required. There are no additional communicational overhead because the calculation is performed on the received signal strength. RSS lacks at a point when values change rapidly, which can be the result of fast fading (**Fast fading** occurs when the coherence time of the channel is small relative to the delay constraint this problem repeated measurements can be done.

TDoA (Time Difference of Arrival): Ali Malekitabar *et al.* in [6] presented a technique for localization of a mobile (MS) with the help of BS (Base Station). Estimation of any node's position becomes difficult especially in urban areas due to obstacles like buildings etc. These hurdles cause the delay which leads to error in localization. Authors also tells that we can reduce localization errors caused by NLS (non line-of-sight) by making comparisons to predefined nodes in the database but at the same time it will increase the delay as a result of increase in computations. In multipath environment sometimes, this approach uses four antennas to reduce multi path fading effect. This technique requires three synchronized BS. All three BS monitor the time of arrival of packets and the difference in times is calculated. Hyperbolas formed from relative time of arrival and the point where they intercepts each other that point is considered as the required position. The below figure shows three BS and a MS and their relative distances. [6]

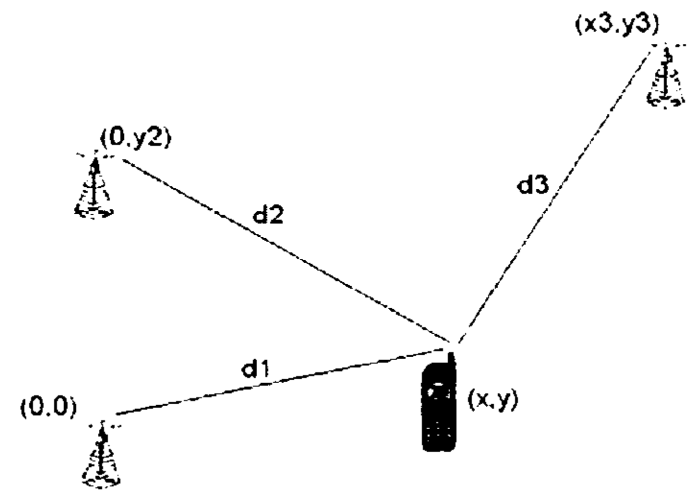


FIG. 2.1 SCENARIO OF TIME DIFFERENCE OF ARRIVAL

Angel-of-Arrival (AoA): The approach presented by Rong Peng *et al.* in [7] works with the angle between propagation direction and some reference direction which is known as "Orientation". Orientation is a fixed direction against AoA which is measured and represented in degree in clockwise direction from North. At minimum two non-collinear neighbor beacons are required to find the location and at least three to discover both location and orientation. Some assumptions are: All angels in the range of $[0,2\pi]$, transmission range is bounded by d_{max} (maximum transmission range) of any packet received outside of transmission is consider as too long, line of sight is assumed which is a major limitation of this research paper. Basically there are two approaches used in this paper for localization: First one was probabilistic localization with orientation information and other was probabilistic localization without orientation information.

2.2 Range-free approaches

DV-hop: Shuang Tian *et al.* in [8] presented DV-hop classical distance vector-hop exchange so that all nodes in the network get distance in hops to the beacon nodes. It works as: anchor nodes broadcast their positions and distances from other beacon nodes. Initially when each anchor node broadcast a packet it sets the flag to zero. Each time this packet is received by any node the value of flag is increased by one. When this packet reaches to other anchor node, the destined anchor node can find the number of hops between other anchor nodes and can calculate the average distance per-hop. One drawback of this approach was that the unknown node in the network considers the first received distance as an average distance value and calculates the distances to beacons and other nodes in the network in their range using this average distance. If the first distance received by any node is more or less than the actual then it definitely leads to wrong measurement of distance. This problem is handled in this paper using RSS (received signal strength) but the proposed algorithm has its own limitations such as "only the neighbors of anchor can update". So the future work is how to precisely localize more nodes and increase the localization accuracy of the whole network.

Centroid: Bulusu and Heidemann *et al.* in [9] had proposed the centroid localization algorithm which is proximity-based (numerical description of how far apart objects are), coarse-grained (which are not absolute) localization algorithm. There are three basic steps in this algorithm: in step one all anchor nodes transmit their position in their range area, which was listened by every node during their listen mode for some time period. As the second step unknown nodes calculate their position with respect to the anchor nodes which are within its range area. In centroid algorithm (X_i, Y_i) is location of anchor or reference nodes which are used to estimate the positions of non-anchor nodes. When a node receives a beacon it calculates its position with the help of following formula.

$$(Xest, Yest) = \left(\frac{X_1 + \dots X_N}{N}, \frac{Y_1 + \dots Y_N}{N}\right)$$

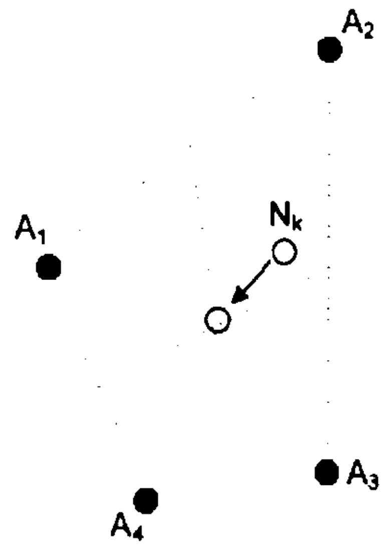


FIG. 2.2 WORKING OF CENTROID APPROACH

Working of centroid localization is presented in the above diagram. Although centroid localization algorithm was simple but a big disadvantage is that it faces huge localization errors.

APIT (Approximate point in Triangulation): In [10] T. He *et al.* divided the sensor deployment area into triangular regions among the anchor nodes. Area of those triangles is narrowed down to the audible anchors. It is to make sure for better performance that unknown node draws a triangle among the nearest possible anchor nodes. The aim is to reduce the triangular area as small as possible in which the desired node is located this technique is known as Point-In-Triangulation Test (PIT). APIT repeats the process of PIT again and again until the desired accuracy or all possible triangulation combinations of anchor nodes are tested and aggregation is performed. As a last step nodes determines its position by calculating the center of gravity (COG) which is the intersection of all of triangles in which a node is lying. APIT is best suited for irregular radio patterns and random placement. As all of the above mentioned steps are performed by the nodes themselves so it is a distributed algorithm. The limitations of this approach are: it requires complex computations, assumes large radio transmission range. This approach is tested in 2*d* the future work is to test this approach in three-dimension.

Trilateration for initial estimate of sensor position

Othman, Shaifull Nizam *et al.* in [11] discussed the theme of trilateration in a very good manner. Trilateration is a geometric based approach to find the relative positions of objects it works like triangulation. The only difference is that triangulation calculates angles and at least one known distance, which trilateration does not do. In **Fig. 2.3** trilateration uses three known positions of reference points and measures the distance between reference points and subject. To achieve accuracy in 2*d* plane it is recommended that at least three known points should be used and four reference points for 3d plane. The point where three spheres intersect is considered as the position of a point we want to know.

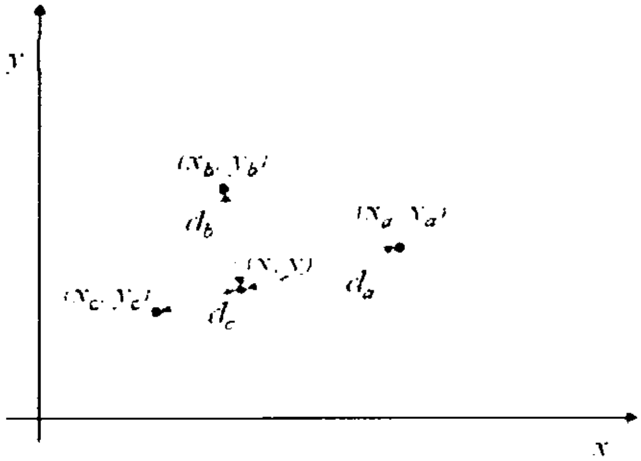


FIG. 2.3 TRILATERATION BASED APPROACH

Koen Langendoen *et al.* in [12] discussed and compared three distributed algorithms, which are most, fitted in distributed environment. Those algorithms are Ad-hoc positioning, Robust positioning, and N-hop multilateration. These algorithms have three common procedures firstly, they calculate distance, find positions and then refine the calculated positions. The algorithms are tested in error ranging, network connectivity and anchor fraction. Three methods are used to measure distance: dv-hop, Sum-dist, Euclidean. For initial estimate of position lateration or Min-max any one of these could be used. At third phase refinement is made. Based on the results the authors stated that every approach has its drawback and advantages so only a single approach does not perform better they could be combined.

2.3 Genetic algorithm for Localization

Ehsan Heidari *et al.* in [13] discussed genetic algorithm from the class of evolutionary algorithms. It has inspiration from the biological evolutions as it works on the theme of genetics or genes. GA encodes feasible solutions in the form of chromosomes and selects these chromosomes on different basis such as Roulette wheel selection (Random), Selection based on fitness (as fitness is calculated for every chromosomes) etc. GA works on encoded solutions in the form of binary or real numbers.

Basic steps in Genetic Algorithm are:

1. Works by selecting initial population of possible solutions
2. Possible available solutions are enhanced for better fitness values
3. Every new individual solution is evaluated for fitness values
4. Stopping criteria can be some predefined number of generation production, till required fitness achievement or it can be time dependent.
    1. Selection of chromosomes for the production of new chromosomes is on the base of best-fitted chromosomes
    2. GA uses two operators' mutation and crossover for the production of new (off springs) for better fitted chromosomes
    3. Fitness of every new off spring is tested
    4. Replace better fitted chromosomes with new individuals

Qingguo Zhang *et al.* in [14] introduced genetic algorithm with two new genetic operators: single-vertex-neighborhood mutation and descend based arithmetic crossover. First operator works by choosing a random vertex and by moving it around the actual position. The second operator sorts the chromosomes in descending order according to their objective values and then crossover is performed in better fitted chromosomes. The basic function of these operators was to get the better fitness value. The fitness value is a constant number, which tells about the fitness of a particular chromosome. There are two assumptions in this paper first one is that unlocalized sensor nodes has at least 3 anchor nodes in their range and second one is that after the network is deployed every node measures its distance to its neighboring node. Central node construct graph on the base of information provided by sensors using multi-hop in the network then run the proposed solution to find out the positions of unlocalized sensors and finally inform the sensors about their coordinates. Although the proposed approach in this paper is centralized however the authors are in the favour of distributed because it is easily scalable. Authors' claims: high position accuracy, best results for uniform topology better work for irregular networks too, no excessive communication required and gaussian noise is bearable. New generations were produced by above mentioned operators to get comparatively better objective values. Performance of genetic algorithm depends on the choice of crossover and mutation operators. Future work of this paper was to investigate the implementation of the proposed algorithm in a distributed architecture as it has easy scalability.

Gao Yang *et al.* in [15] proposed an algorithm for WSN localization known as FRGA, which is basically some change in GA. Author criticize genetic algorithm to have "premature convergence" so globally optimum solution is difficult to obtain. Premature convergence means that the solution at which we reached as a global optimum solution, in reality that is not global optimum solution. Authors' motivation for his work was to reduce localization error which is too high in some nodes. To cop these problems **FRGA** (filter and replenishment strategy) is proposed. The proposed approach worked by limiting the population feasible region. According to the authors premature convergence is because of lack of population diversity. FRGA proposed filtration strategy, which deletes the undesirable chromosomes that are produced by mutation and crossover operators, replenishment strategy maintains population diversity. Population diversity means that there should be diversity of chromosomes to each other. Future work is to implement in FRGA in real environment and applications.

Steven Lanzisera *et al.* in article [16] presented a ranging technique for distance measurement between two identical wireless sensor nodes. This approach has less hardware requirement. RF TOF (Radio Frequency Time of Flight) has good wall penetration and provides better results as for as considering accuracy. There are two types of RF TOF approaches in literature; first one which requires high degree of clock synchronization and second which demands loose absolute time synchronization as it uses two-way ranging approach. It works by sending a range packet and wait for reply the biases are subtracted away .This approach can use short bursts in a frequency hopping manner so it can avoid interference and undesired detection.

In [17] Zenon Chaczkol *et al.* presented ways of WSN localization of senor nodes as it is crucial to know that from where the information is received. It represents methods of localization such as:

(i) Distance and direction from a point (x,y)

(ii) If we knew direction from two known points (Angle of Arrival)

(iii) By knowing distance from three known points (ToA, TDoA, RSS)

If radius of three known points intersects on one point this is the location of unknown node. This paper discussed two exceptions where an unknown node can't be localized.

(i) When at least two known points have the same deployed location

(ii) When all three known points are on the same line or collinear then we cannot find position *(x, y)* of a node.

This paper presented a way to find the position of a deployed sensor when three anchors cannot intersect on a single point. If the position of unknown node is not at the intersecting point then position of unknown node can be determined by the three points where the radius of three known points (reference points) intersect each other. Author recommends that if position of BS which is actually a sensor is erroneous then it should not be propagated until a bit surety of position accuracy. Because other unlocalized nodes will use their propagated positions to localize themselves. As the proposed approach is simulated, the future work is to implement it in real life experiments.

2.4 Ant Colony Algorithm

Marco Dorigo in [18] presented a novel approach that works on the ideology of ants functioning. It was a new approach towards stochastic combinatorial optimization problems. The beauty of this approach was that it considered positive feedbacks, distributed computation and used greedy heuristic method to get most optimum solution. He implemented this approach to classical Travelling Salesman Problem (TSP). This approach consisted of different parameters and also discussed them. This approach works as: ants starts their search at a random point and then move further on the base of pheromone (a substance that has some kind of smell) is laid on the edges and then those edges are traversed using a probabilistic rule (state transition rule). Ants choose the edge/path which has more amount of pheromone on them. The pheromone evaporates with the passage of time, so that all ants do not follow the same path and each path is not used in all iterations for the optimum solution. When an ant completes its tour it updates the pheromone amount on the edges it has crossed/passed through during its current tour. The paths which has less pheromone are no more visited and at last all the ants considers one path that is most optimum (short) to food.

Frank Neumann *et al.* in [19] discussed the combination optimization problem in detail. It consists on triplet (S, Ω, *f*).

S: an area of search space which is defined over discrete variables

Ω: set of limitations and conditions among the variable

$f: S \rightarrow R$ objective function to be minimized/maximized

When search space is searched under Ω (limitation and conditions) the resultant region is the feasible region. This feasible region consists of the solutions, these solutions are improved on meta-heuristic information in ACO and one of these solutions could be the optimum solution.

In [20] Farhad Nematy *et al.* proposed a deployment strategy for robots using ant colony. When robots travel they randomly draw sensors during travel. The energy level of the sensors is taken as pheromone. Probabilistic approach of ant colony is used to decide whether to visit a node or not. The assumptions in this paper are that the deployed sensors are has high energy and are rechargeable. When a sensor is recharged its pheromone value increases. The sensors, which are within the path of robots are recharged by the robots who visit that path which is equal to pheromone deposition. Pheromone evaporation is resembled to the reduction of battery power. If some specific path is not visited by any robot or visited after a long time the life time of that sensor node will expire or at least the battery power will be reduced which is considered as the pheromone is evaporating. Selection of path is decided on roulette wheel mechanism. Future work mentioned by the authors is to extend the proposed work for obstacles avoidance.

Rahul Putha *et al.* in [21] prove that ant colony optimization is consistent in results than genetic algorithm especially when there is large number of trial. ACO reduced the execution time which allows implementing it in real time signal control system.

N. Binti Sariff *et al.* in [22] implemented both ant colony ant genetic algorithm fort Robot paths planning in global static environment. Performance is compared in terms of speed and number of iterations taken by the both algorithms to find optimal path. Implementation showed that ant colony is much faster in terms of speed and took less iteration to complete the task in different node complexity. ACO parameters are easily adjustable to different complexities.

Chapter 3 Proposed Approach

3.1 Ant colony optimization (ACO)

Ant colony optimization (ACO) is a "population-based metaheuristic that can be used to find approximate solutions to difficult optimization problems" [24].

Heuristic

Many applications face problems that are believed not to be solved to optimality within polynomial time. To solve such problems we can use approximated approaches which could provide us optimum solution relatively in short time span. This kind of algorithm is known as heuristic algorithm. In these approaches experience-based knowledge is used to solve the optimization problem [24].

Metaheuristic

The beauty of metaheuristic algorithms is that they define some essential heuristics that can be used to solve different optimization problems, to avoid local solutions [35]. The metaheuristic algorithmic concepts significantly increased the ability to find hard solutions of optimization problems in reasonable time. Metaheuristic approaches lead to the optimum solution either by iteratively enhancing the candidate solution with respect to some quality such as error measurement to know accuracy or by starting from a null solution and move to an optimum solution. [24]

Proposed approach

In contrast to [14] our proposed approach is distributed as it may increase scalability and reduce complexity [14]. We have implemented ant colony optimization for WSN Localization problem instead of genetic algorithm as in literature survey we have studied that ant colony optimization is better in term of time complexity (less iterations) [21], easy adjustment of parameters than genetic algorithm [22] and better in accuracy [34]. We used anchor (beacon or reference) nodes as well as non-anchor nodes to localize the unlocalized nodes in the network.

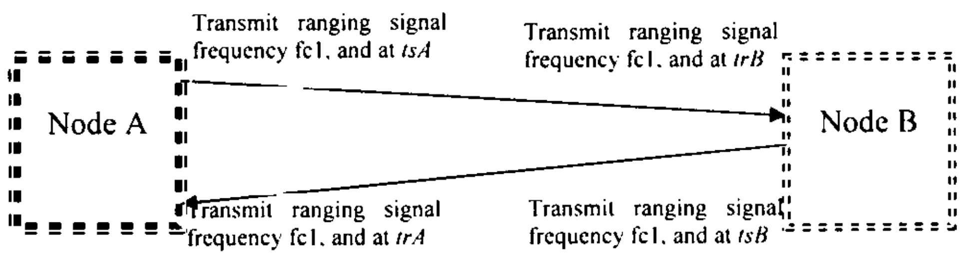
3.2 Working of our proposed approach

Our proposed approach works in three steps:

1. Every node computes distance to anchors and to other nodes within the network in their range. Every node that is directly connected to a node is considered as the neighbor of that node.
2. Find initial position (by trilateration)
3. Refine the initial computed positions to get optimum position

1. Distance between nodes is calculated using Radio Frequency Time of Flight (RFToF) [17] as it requires less synchronization. We preferred this technique on many other methods because it does not require directional antennas such as Angle of Arrival (AoA) [7] demands. Time of Arrival (ToA) [4] is also a ranging technique but it requires perfect synchronization among nodes. ToA requires only one message to calculate the ranging between nodes because the nodes are perfectly synchronized. Perfect synchronization is a difficult task to achieve in dense environment.

RFToF woks as follows: We measured time taken by one side and use this time to measure distance to other nodes in range using speed of light and propagation speed. The working is as follows: Node A transmits hello message at time T*r* with time stamp, Node B receives and respond by sending acknowledgment message after some response delay that is a constant time for response. Node A receive acknowledgment packet at time T*a* and subtracts response delay form signal propagation time then divides it by 2 to get one side time taken to transmit.



Time of flight is measured by the equation given below.

$$\text{Time of flight} = \frac{t_{rA} - t_{sA} - (t_{sB} - t_{rB})}{2}$$

Where $t_{sA} < t_{rB} < t_{sB} < t_{rA}$

By measuring the delay in two signals the round trip time of flight is estimated and range estimation is done by $c.\widehat{T}_{RT}/2$

We assumed that if two nodes *i,j* are located at zero distance then both nodes *i,j* are at the same position (overlapping). If the distance between nodes is so far that they can't communicate with each other then distance is assumed to be -1, if distance between any two nodes is > 0 and are within range of each other than they can communicate. The desirability of any two nodes is higher if they are at shorter distance as compared to others nodes. It means that the probability of selecting the nearer and 1-hop neighboring nodes is higher in the process of localization. We considered 10% Gaussian noise factor as in our base paper.

2. Each anchor node broadcast its *id*, position to initialize the localization process. Each unlocalized node constructs a graph locally as oppose to the approach used in our base paper [15] in which the graph is constructed only at central node. The feasible region is produced by the trilateration using distance measurement. The process of trilateration is repeated for some time to get the pool of candidate solutions (coordinates) in the feasible region this region is then searched for optimum solution using our proposed algorithm.

Candidate solutions are the solutions that fulfill some predefined criteria such as *Residue* given below. Residue measures the quality of produced solutions and enables us to accept or reject the estimated positions to construct the pool of feasible solutions. Trilateration requires at least three anchor nodes in 2d plane to get the estimates of position of a node. As in our based paper, we too assumed that there are at least three anchor nodes in our proposed methodology.

$(x_i - X_{un})^2 + (y_i - X_{un})^2 = r_i^2$

The simplification of trilateration technique is:

$(x_1 - x_{un})^2 - (x_3 - x_{un})^2 + (y_1 - x_{un})^2 - (y_3 - x_{un})^2 = r_1^2 - r_3^2$

$(x_2 - x_{un})^2 - (x_3 - x_{un})^2 + (y_2 - x_{un})^2 - (y_3 - x_{un})^2 = r_2^2 - r_3^2$

By eliminating r2

$2(x_3 - x_1)\, x_{un} + 2(y_3 - y_1)\, y_{un} = (r_1^2 - r_3^2) - (x_1^2 - x_3^2) - (y_1^2 - y_3^2)$ ..... (3.1)

By eliminating r1

$2(x_3 - x_2)\, x_{un} + 2(y_3 - y_2)\, y_{un} = (r_2^2 - r_3^2) - (x_2^2 - x_3^2) - (y_2^2 - y_3^2)$ ..... (3.2)

Writing (3.1) and (3.2) in matrix form

$$2\begin{bmatrix} x_3 - & x_1 & y_3 - & y_1 \\ x_3 - & x_2 & y_3 - & y_2 \end{bmatrix}\begin{bmatrix} x_{un} \\ y_{un} \end{bmatrix} = \begin{bmatrix} (r_1^2 - r_3^2) - & (x_1^2 - x_3^2) - & (y_1^2 - y_3^2) \\ (r_2^2 - r_3^2) - & (x_2^2 - x_3^2) - & (y_2^2 - y_3^2) \end{bmatrix} \text{.....(3.3)}$$

In **equ. (3.3)** the matrix x_{un}, y_{un} is unknown but the parameters used in other matrices are known. So by solving the above equations, estimate of coordinates can be achieved. After getting the initial estimate of coordinates, we take a decision regarding adoption or leaving of the initially produced coordinates with the help of *residue* given below. The least square solution should be:

$$\bar{X} = (A^t A)^{-1} A^t b$$

Residue is a type of check that is used to know the difference between measured (d_i) internodes range and estimated range ($\tilde{d}_i$). Initially measured coordinates are acceptable if the residue difference is small.

$$Residue = \frac{\sum_{i=1}^{n}\sqrt{(x_i - \hat{x})^2 + (y_i - \hat{y})^2} - d_i}{n} \text{ .....(3.4)}$$

(X_i,Y_i) are the positions of anchors which are assumed to be pre-known by the GPS or by any other mean (pre-coded), (X_{un}, Y_{un}) represents the coordinates of unknown node. r_i is considered as distance. The **Fig. 3.1** represents the working of trilateration, in this figure AN_i are Anchor nodes, d_i are distances, UN is representing the unknown position of a node.

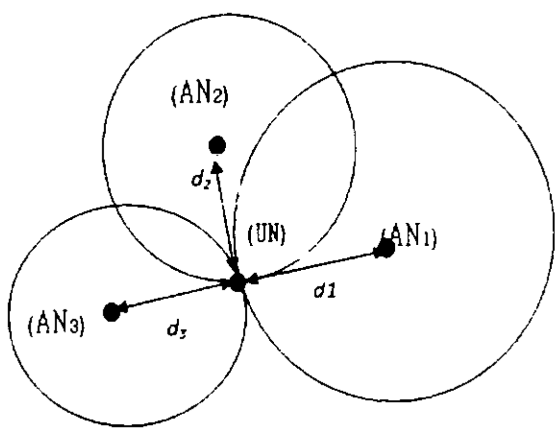


FIG. 3.1 METHODOLOGY OF TRILATERATION

The spheres in the above figure are dependent on the distance between any known and unknown nodes it means that if an unknown node is within the range of a known node the spheres can be decreased within the radius to that range so that the spheres can intersect at a point, which is the position of unknown node. If a node is lying outside the range of three known nodes then that node cannot be localized. There are two exceptions when the position of unknown points can't be founded.

(i) When all there known points (A, B, C) are collinear (lies on one line). As in the **Fig. 3.2** below all three known point are on the same line and they intersect at two points *p* which creates ambiguity in selection that which point should be considered as actual position.

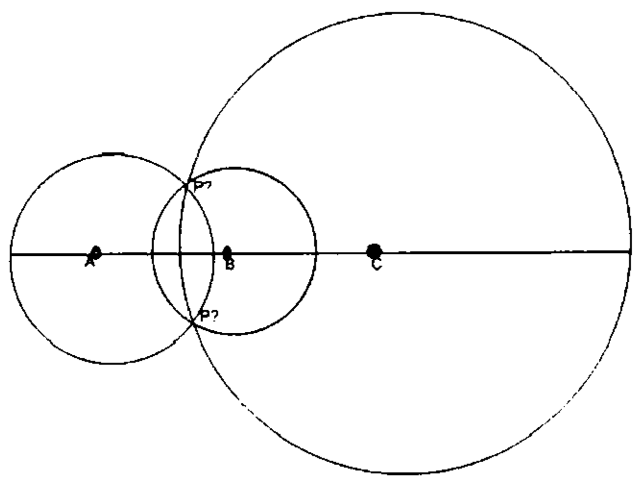


FIG. 3.2 COLLINEAR ANCHOR/REFERENCE NODES

(ii) When at least two known points are on same position

If the spheres are not intersecting at one point due to error in distance measurement then the unknown point can be calculated by knowing the position of the points where spheres are intersecting each other [16].

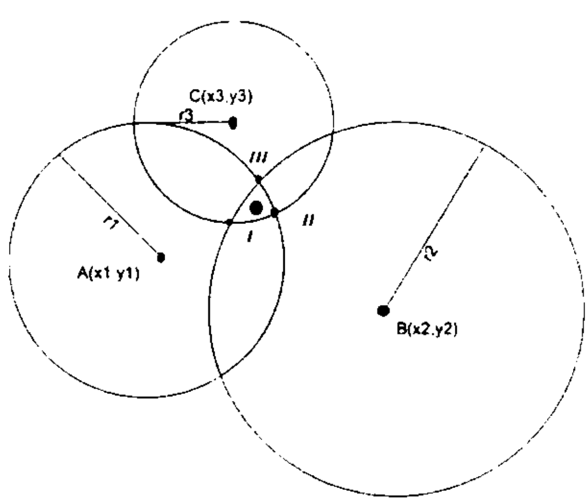


FIG. 3.3 INTERSECTING POINTS OF SPHERES

In such a situation where spheres are not intersecting at one point we can calculate the position of intersecting points then use those calculated positions to find the position of unknown node, such as we calculated the third point III(xIII, yIII) of intersection below.

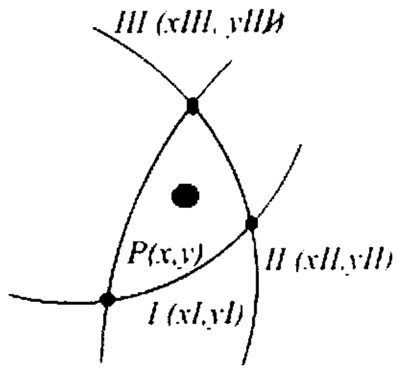


FIG. 3.4 POINTS OF INTERSECTION OF ANCHOR NODES

$$x_{III} = x_1 + r_1 Cos(\alpha \pm \beta) \; ..... (3.5)$$
$$y_{III} = y_1 + r_1 Sin(\alpha \pm \beta) \; ..... (3.6)$$

Where

$$\alpha = arctan\left[\frac{y_2 - y_1}{x_2 - x_1}\right] \; ..... (3.7)$$
$$\beta = arccos\left[\frac{r_1^2 - r_2^2 + d_{12}^2}{2r_1 d_{12}}\right] \; ..... (3.8)$$
$$d_{12} = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2} \; ..... (3.9)$$

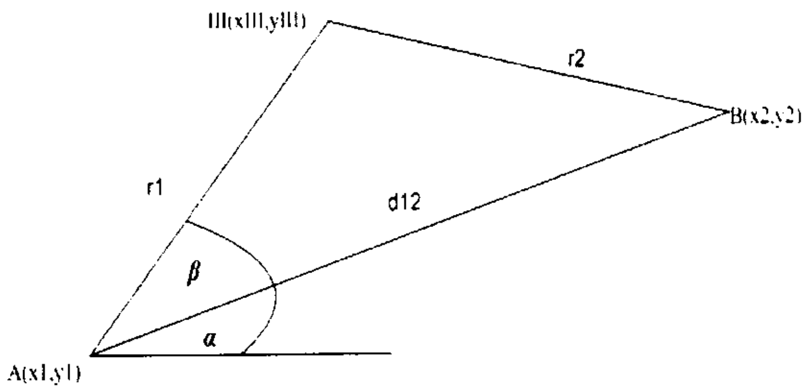


FIG. 3.5 TRIANGLE FORMED BY INTERSECTING POINTS

Two nodes with the known positions (x_1, x_2) and (y_1, y_2) and one unknown point (x_{III}, y_{III}) forms a triangle in which all sides are known now by using *cosine* theorem.

$$r_2^2 = r_1^2 + d_{12}^2 - 2r_1 d_{12} cos\beta$$

Now the position of unknown point can be calculated by:

$$X = \frac{x_I + x_{II} + x_{III}}{3}$$
$$Y = \frac{y_I + y_{II} + y_{III}}{3}$$

This process is repeated for some iteration to get a pool of solution. That poll of solution will be the input to the ant colony algorithm.

3. ACO works with artificial ants who searches for optimal solution from a given search space. As it works by selecting the shortest paths (edges) between two points (source, food) so to implement any problem in ACO we have converted our problem in graph form. Then the software agents (artificial ants) build solutions by moving on paths of the weighted graph. The resultant solution is produced on the base of pheromone (substance that has smell and is evaporatable with respect to time) updating model and the produced solutions accepted as the optimal solution. Solutions can be constructed in two ways:

Parallel: all ant moves together to next state then update the pheromone

Sequential: an ant complete the tour before next ant starts

Both of these approaches construct solutions. In AS (ant system) these approaches do not have significant influence on the performance of algorithm so any one of the two mentioned approaches can be used.

Selection of the parametric value is important as pheromone trail value should not be very low so that the selection of a path at first tour do not become biased (emphasize on one path or poor solution are searched). If it happens it means that we did not explore the search space properly which will lead us to local optimum solution. At the same time pheromone, trail value should not be very large so that it took long time for evaporation and the limit of iterations or of time approaches and the search space remained unexplored. This leads to lack of diversity, which is the basic need for best optimum solution and is the basic building block of evolutionary algorithms.

As in [14] we take (m) number of anchors as 10% of the total (n) number of sensor nodes. Connected graph is constructed between the feasible solutions (possible coordinates) which were produced by the iteration of trilateration, parameters for ant colony algorithm are initialized. The feasible region works as allowed set of coordinates to be visited by ants or in other words the set of possible coordinates where a sensor node could be then one of them is selected as optimum coordinate for a node. All the coordinates within the feasible region are candidate solutions. Initially all the edges are assigned the same amount of initial pheromone except those who are anchor nodes which has more amount of pheromone. The pheromone amount on the anchor nodes will be the criteria for the selection of a candidate solution as an optimum solution. A candidate solution is selected if it has the pheromone amount equal or more nearest to pheromone amount on anchor nodes.

The movement from one state to another is through the state transition rule. This rule is used in decision making that whether to visit a state or not and if visit then which one of the next available states (positions). Ants remember visited coordinates from a feasible set so that once a state is visited it should not be visited again in that tour. The state transition rule is:

$$P_{ij}^{k} = \begin{cases} \dfrac{[\tau_{ij}]^{\alpha}[\eta_{ij}]^{\beta}}{\sum l \varepsilon N_i^k [\tau_{il}]^{\alpha}[\eta_{il}]^{\beta}} & if\ j \varepsilon N_i^k \\ otherwise\ no\ move \end{cases} \quad ..... (3.10)$$

As ant moves from a node i to node j using the above mentioned **equ. (3.10).** If an ant has visited all the states (possible coordinates) in one tour then there will be no more moves in that particular tour. The parameters α and β are adjustable that controls the relative influence of pheromone trail for all edges (i, j) τ_{ij} and heuristic desirability η_{ij}. The values of $\alpha = 1$ and $\beta = 2 - 5$ suggested in [24]. We take the parameters $\alpha = 1$ and $\beta = 5$ and found the best results. Initial pheromone τ_0 is calculated as m/C^{nn} where m is number of ants and C^{nn} is the length of tour (in our problem it is the sum of the difference of coordinates from pool of solution) [24]. Heuristic desirability is dependent on distance and is computed as inverse of distance of any specific edge like ij as: $\eta_{ij} = \frac{1}{d_{ij}}$., N_i^k is a set of states that an ant k not visited yet. This decision making ability reflects that ants maintain the memory. Later on this memory also help ants to calculate the shorter distance of a tour. Once a tour is completed this memory is cleared for the next iteration.

Transition probability is used to keep balance between pheromone intensity (which is based on the history of good moves) and heuristic information (which expresses how much desirable a move is). Usually the values of α, β are limited as $0 < \alpha, \beta \geq 1$. But these values could be different depending upon the problem.

The effect of α, β parameters is discussed below:

If α= 0 then value of pheromone trail will become 1 which means that the decision of the selection of next move will be totally biased to shortest distanced paths.

If β=0 then the selection of next state will be on high pheromone intensity.

α >1 will lead us to rapid convergence (all ant will follow the same path that is already used by the ants and the search space will remain unexplored).

If any of the above condition occurs then algorithm will be biased depending upon the pheromone or heuristic information. Which means lack of diversity so local solution will occur where as global solution is desirable.

Algorithm evaporates the initial pheromone lying on the edges so that exploration can be done for better solutions. Pheromone evaporation plays very important role in avoiding premature convergence and at the end global optimum solution is achieved. Evaporation is done as:

$$\tau_{ij}\ (1-\rho).\tau_{ij}, \forall (i,j) \varepsilon\ L \ldots\ldots (3.11)$$

The parameter $0 \leq \rho \geq 1$ is a coefficient that represents the evaporation of pheromone, we take ρ= 0.5. Pheromone evaporation plays very important role because if pheromone on an edge is not evaporated we cannot avoid bad moves taken in previous tours, cannot explore the feasible solution search space. If ρ=0 the evaporation will be totally eliminated and if ρ=1 then the evaporation process end very soon and will not be able to get optimum solution due to lack of convergence. We knew that the class of evolutionary algorithms works on the base of two things first one is exploration and the second is exploitation. Exploration means that the search space of feasible solutions should be explored where as exploitation means that taking the necessary steps for better optimum results.

When ants move on the edges they modify the initial pheromone trail so that the ants that will be visiting after them should visit these states and help in the selection of a better solution. Pheromone updating is one of the daemon actions. This pheromone updation is based on the following rule.

$$\tau_{ij} = (1-\rho)\tau_{ij}^{k} + \Delta\tau_{ij}^{k} \ldots\ldots (3.12)$$

Where

$$\Delta\tau_{ij}^{k} = \begin{cases} \frac{1}{C^k} & if\ arc\ i, j \varepsilon T^k \\ otherwise\ Zero \end{cases} \quad \ldots..(3.13)$$

where C^k , the length of the tour T^k travelled by ant k , is computed as the sum of the length of the arcs belonging to T^k

τ_{ij}^{k} is pheromone deposited amount for the transition from i to j node. Ant k is the ant that has completed its tour and will now update the pheromone on the tracks that it has visited/crossed during its tour. Pheromone amount deposited by ant is represented as $\Delta\tau_{ij}^{k}$.

To avoid local optimum some specific number of iterations or CPU time can be used. At the end we come to an optimum solution. At the end we calculated mean location error by using the formula same as in our base paper. This formula tells us error in average of our proposed approach.

$$\text{Mean location error} = \frac{\sum_{i=m+1}^{n} \left[\left(x_i^t - x_i\right)^2 + \left(y_i^t - y_i\right)^2 \right]}{(n-m)R^2} \quad \ldots..(3.14)$$

Where (x_i^t, y_i^t) are the true positions of sensor node (x_i, y_i) are representing the estimated positions of our approach, R is radius. Average connectivity of nodes can be increase/decrease by the increase/decrease in radius. After the localization of one node we further used this node as anchor node to localize the remaining unlocalized nodes. This approach allows us to use less number of anchor nodes as they are expensive.

For example pool of solution is given by the above procedure. We implemented ant colony optimization on this pool of candidate solutions. We assumed the connected edges between the candidate solutions as shown in the **Fig. 3.6**.

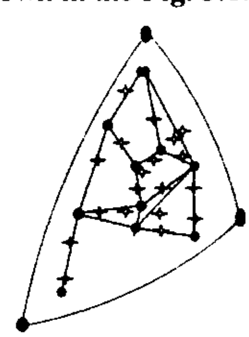


FIG. 3.6 POOL OF SOLUTIONS GENERATED BY PREVIOUS MENTIONED PROCEDURE

Initially pheromone amount is equal on the edges of candidate solutions. A careful selection of initial pheromone is m/C_{nn} [24]. Whereas the amount of pheromone on the anchor nodes is (pre-defined number of iteration * m/C_{nn}). This amount of pheromone on anchor nodes is the criteria of selection for an optimum coordinate for a node. Ants are placed in this region of feasible solutions. Ants move to next state depending on pheromone value and heuristic desirability at the same time by checking feasibility of that move. A move is a feasible move in a tour if the destined state is not visited yet in that tour/iteration. Ant moves in this feasible region as shown in the **fig. 3.7** where plus sign on the lines shows the movement of ants. When all ants complete a tour they update the amount of pheromone on the edge they have utilized in their tour using **equ. (3.12)**.

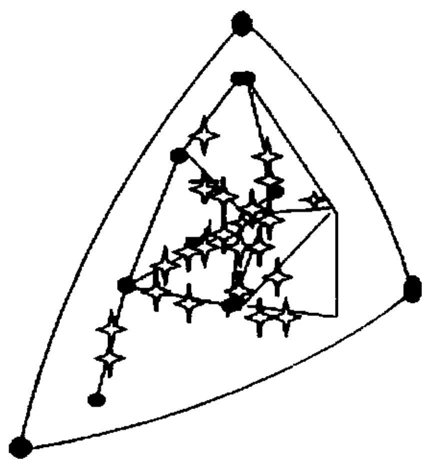


FIG. 3.7 ANTS FUNCTION ON CANDIDATES POOL OF SOLUTION

The decay parameter is used as 0.5 so diversity can be achieved. The maximum utilized edge is having the maximum amount of pheromone. In this way all ant completes the pre-defined number of iteration and updates the pheromone value. The most visited/used edge is having maximum amount of pheromone approximately equal to the amount of pheromone on the anchor nodes that is considered as the optimum coordinate.

As shown below in **fig. 3.8** the triangle is representing initial calculated position after some iteration than white shaded circle is representing the position before and final position is in presented in star after the completion of all predefined iterations.

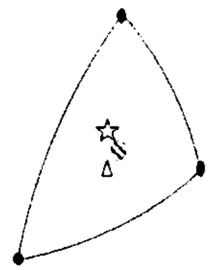


FIG. 3.8 OPTIMIZED POSITION

*3.3 Distributed Approach Scenario (100m*100m)*

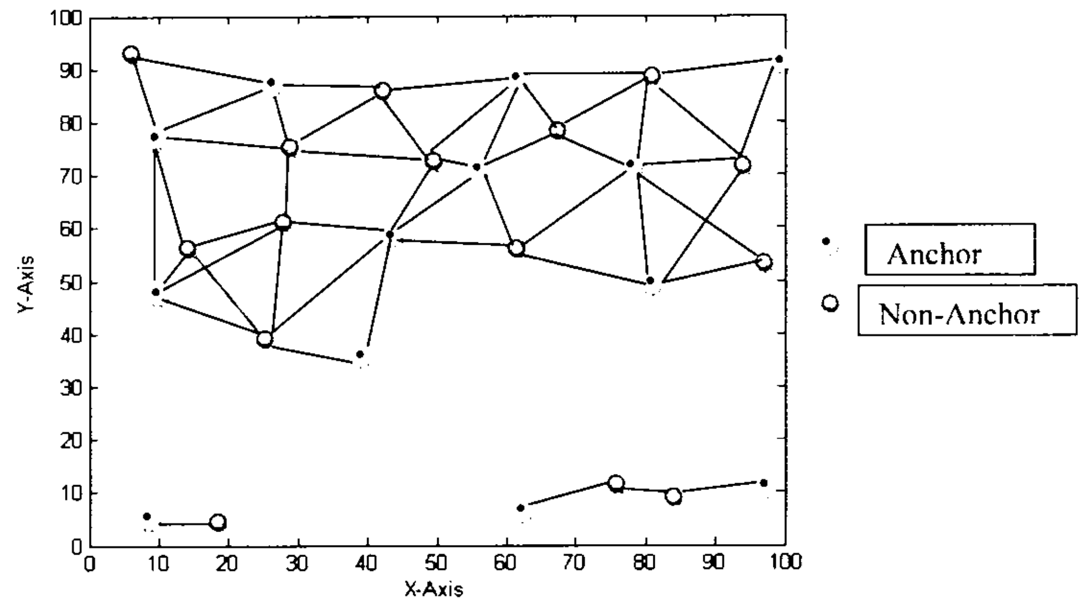


FIG. 3.9 PROPOSED APPROACH SCENARIO

Anchor nodes are aware of their positions and non-anchor nodes don't. In the above scenario any non-anchor node outside the range of at least three anchor nodes cannot be localized. In distributed scenario every node performs its position by its own so no burden and dependency on the central node, no issue of compromising of single central node, no issue of communication cost.

3.4 Flow of Proposed approach

The **fig. 3.9** is a flow chart of the proposed approach, which works as: First of all we generate pool of feasible solutions by trilateration and with some further procedure then we initialize the ant colony parameters.

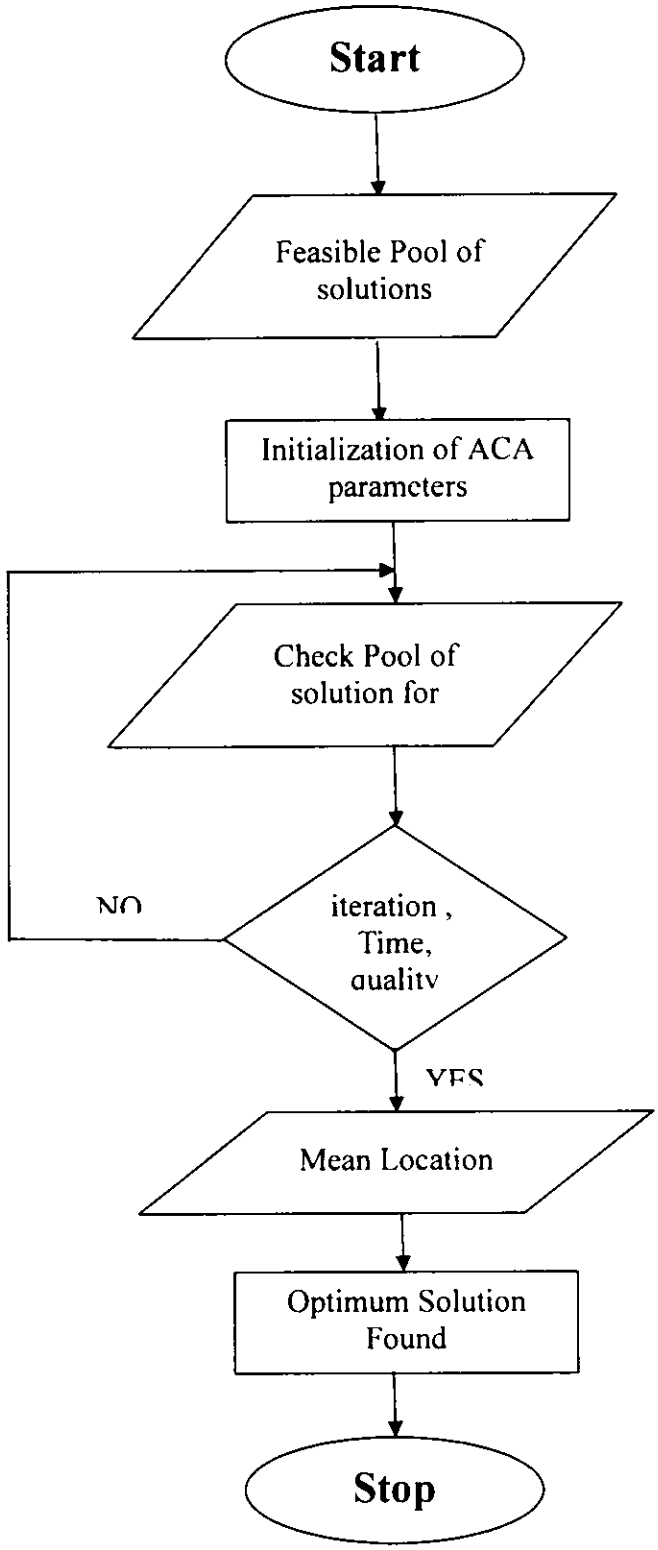


FIG. 3.10 FLOW CHART OF OUR PROPOSED APPROACH

The generated pool of solution is then traversed for some optimum solution. This process continues for some specified number of iterations, till some time limit approaches or until the specified quality criteria satisfies. This gives us an optimum solution which is considered the most appropriate coordinated position for any node. Finally we calculate the average localization error for accuracy measurement.

Chapter 4 Implementation and Results

4.1 Data sets

Data sets are the parametric values we used to obtain our required results. We take n=100 sensor nodes, m=10% anchor nodes. Percentage of anchor nodes is varied and the results are compared with respect to time taken by the nodes to localize. Area of sensor node deployment is 100m*100m. We considered two types of deployment grid and random. As in real scenarios the type of deployment is random so we emphasized on it. We varied values of parameters $0 < \rho < 1$, $0 < \alpha, \beta \geq 1 - 5$. Gaussian noise factor is considered as 10%.

4.2 Network Simulator-2 Tool of Implementation

We used Network Simulator-2 [25] for the evaluation of our proposed approach. NS-2 is widely used tool for the simulation of computer network filed problems. NS-2 required some extra necessities to work for WSN but good thing is that those are available free of cost. In **Fig. 4.1** we discussed the structure of NS-2 and up versions. TCL scripting language is used for simulation in ns-2. It initiates event scheduler, set up network and necessary events for simulation. As the event scheduler schedules the message packet are transmitted. For the execution of a program first we have to compile it, if there are no errors then network animator ("nam") is used to animate. This gives us a GUI network animator where simulation shows the results. Simulation requires some numeric form data to simulate; we can find these numeric data values in a file known as trace file trace_les (text-based). This trace file is generated after the simulation is completed. Trace file information can be extracted by small scripts of AWK or PERL languages. Trace file information enable us to analyze the behavior of proposed approach.

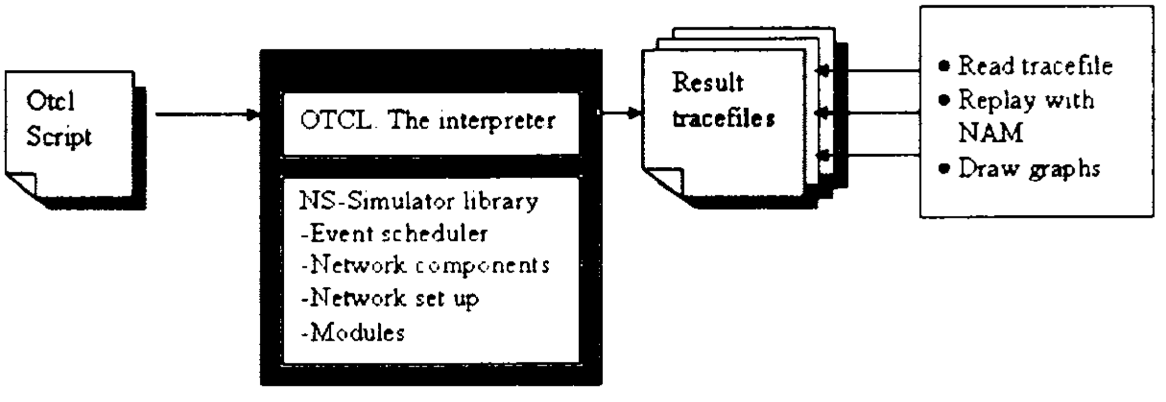


FIG. 4.1 STRUCTURE OF NS-2+

4.3 System Requirements

We have used Windows and Linux operating systems. The system we used has the following specifications:

1. Hardware..............Laptop, 2GB Ram, 2.13 Ghz Processor
2. Linux Operating system......Fedora 14.0
3. Network simulator.....NS-2, version 2.35

4.4 Results Comparison

The graph below **Fig. 4.2** shows the time efficiency of localization problem of base paper and our proposed approach. Results are different due to the factors given below.

- In our base paper centralized approach is used which takes time in communication and if there are large numbers of sensor this approach will take more time to localize the nodes. In contrast, because our approach is distributed it took less time. Whereas proposed approach is distributed. In contrast to centralized approach, it is easily scalable too.
- Search space in genetic algorithm may increase exponantionaly.
- Although scalability is a bit difficult task in centralized approach however it may also time consuming task as well.

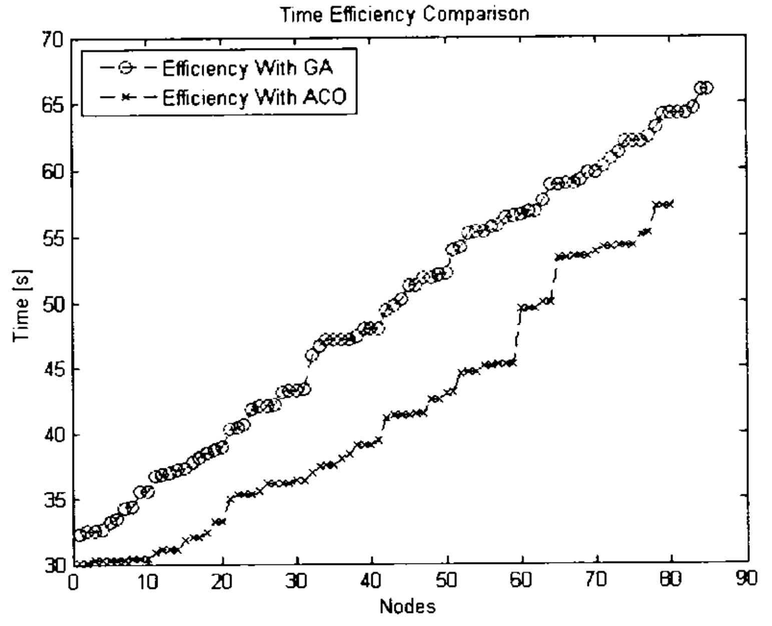


FIG. 4.2 TIME EFFICIENCY COMPARISON OF BASED AND PROPOSED APPROACH

In **Fig. 4.3** the graph below shows the delay comparison when the number of anchor nodes are varied. Rectangle, Circle and Cross signs on the line are representing the three, five and ten number of anchors respectively. We took at least three anchor nodes due to basic condition of trilateration in *2d* plane. We took maximum 10% of anchors as were in base paper. The graph shows that when the number of anchor nodes are increased then task of localization can be achieved in less time, ultimately with more accuracy. Whereas when number of anchor nodes are restricted to only three it took more time to localize the nodes. When there are less number of anchor nodes in the network there will be less number of non anchor nodes that can be localized as at least three anchor nodes are required to localize a node. There will be less chance of fulfillment of above mentioned requirement for each node although we used non-anchor nodes as anchor nodes as well when they become localized but this approach may also not localized network. State change delay in the graph below means that time taken by the simple nodes to localize. In other words we can say that time taken by the unlocalized nodes to change their state from unlocalized to localized nodes after initialization of localization process.

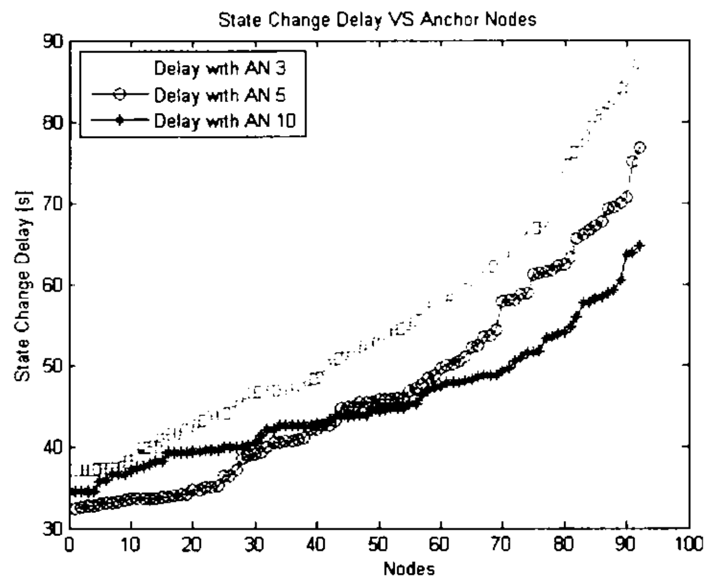


FIG. 4.3 TIME EFFICIENCY WITH DIFFERENT NUMBER OF ANCHORS

In **Fig. 4.4** there is a graph which shows the localization difference of based and proposed approach. Graph line with cross marks is representing base and simple line is of proposed approach. Although there is not very big difference in the based and proposed approach results but there are many other things to be noticed such as how much time is taken and what is rate of accurcy, distributed approach advantages (easyily scalable, less communication cost) over centaralized, implementable in real scenarios etc.

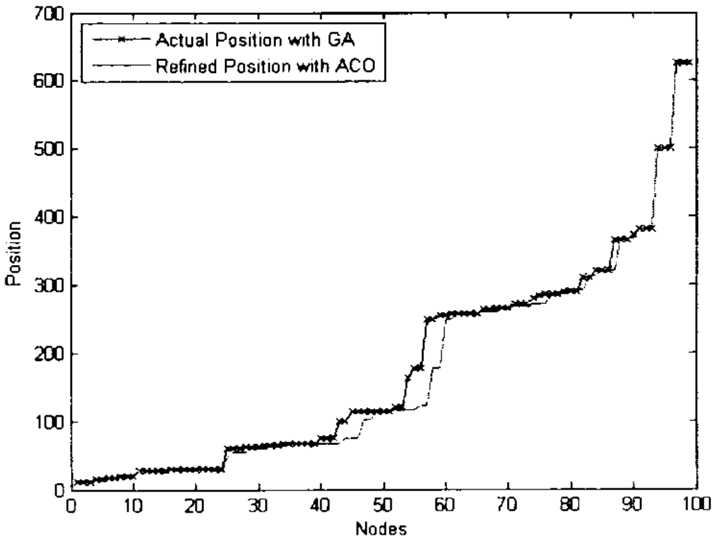


FIG. 4.4 POSITION DIFFERENCE OF BASE AND PROPOSED APPROACH

The **table 4.5** shows the error comparison of bases and proposed approach. The error is calculated with the same data sets on the average base.

[illegible]	[illegible]
0.470937	0.632511
0.500268	0.692854
0.519415	0.750981
0.553039	0.768214
0.570591	0.848556
0.701594	0.865542
0.707295	0.867025
0.735036	0.941663
0.786357	1.070991

TABLE. 4.5 ERROR COMPARISON OF BASED AND PROPOSED APPROACH

The difference in error is due to: in the base paper approach once all nodes estimates the distance they transmits it to a central node. The central node calculates the positions of unknown nodes on the base of estimated distance transmitted by the sensor nodes. If initially estimated distance has some errors then the position calculation will definitely be inaccurate. Whereas in proposed approach nodes compute their distances to other nodes in their range by their own and update the computed distance as well for their position updation to achieve more accuracy.

Error of base approach is shown in **Fig. 4.5** where error is reflected with respect to time. It took 45 seconds and graph is still rising higher as the number of nodes will be increased error will increase too, now error is at 1.4. Where as in graph **Fig. 4.6** which is of proposed approach took approximately 30 seconds and error is about 1.

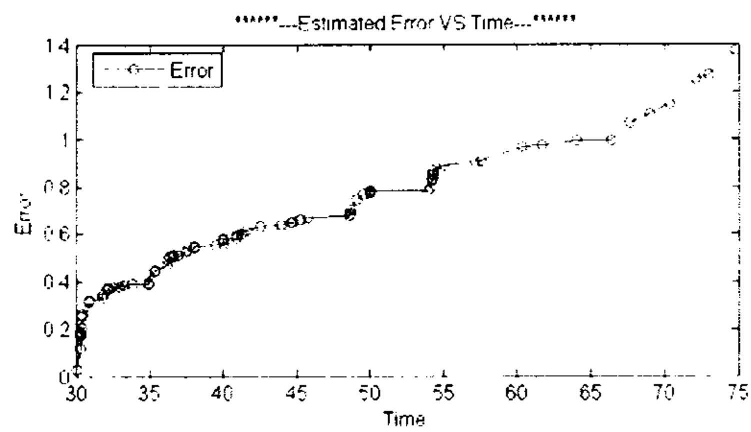


FIG. 4.5 TIME AND ERROR OF BASE PAPER APPROACH

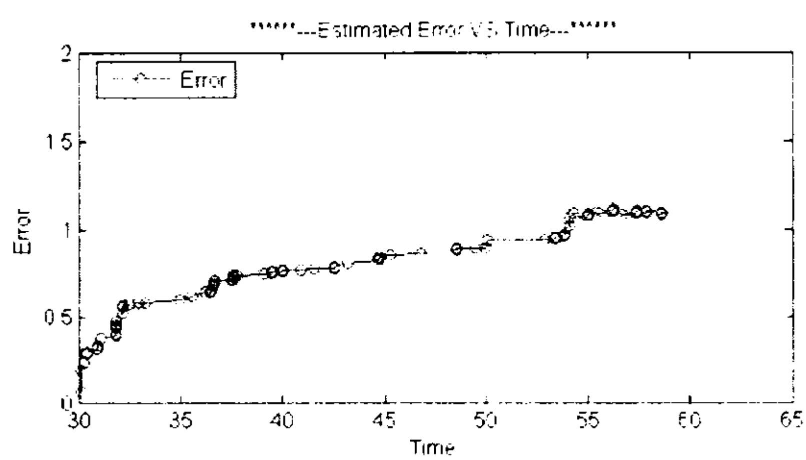


FIG. 4.6 TIME AND ERROR OF PROPOSED APPROACH

4.5 Benefits of Proposed approach

We proposed the AC (Ant Colony) instead of Genetic Algorithm for WSN localization problem. The beauty of this algorithm is that it uses previously gathered knowledge to move further and further for an optimum solution. This information is in the form of pheromone updation. In contrast the genetic algorithm enhanced chromosomes on random bases which may not get better fitness values especially when the selection of chromosome is on random base. We found that proposed algorithm took comparatively less time in position finding, better in accuracy as compared to based approach. Approach in our base paper is centralized which means communication overhead, costly in terms of battery powered sensors in WSN. In contrast to the central node approach as proposed approach is distributed so every sensor/node calculates its own position on its own by relaying on communication to neighboring anchor nodes. Our proposed approach play vital role in increasing the life time of sensor nodes as it reduced communication does not put the entire burden on one central node. In central approach if one central node dies or compromised then whole network could be useless so we proposed distributed approach. Because we used non-anchor nodes as anchor nodes when they become localized, this allowed us to use minimum number of anchor nodes as the anchor nodes were expensive.

Chapter 5 Conclusion and Future Work

5.1 Conclusion and Future Work

We used ant colony optimization for wsn localization problem and the results shows that proposed approach is better in terms of time and accuracy as compared to base approach. Proposed approach is distributed which is an edge over central approach. We used non-anchor nodes as anchor nodes when they become localized.

Future work could be to introduce such techniques that take minimum number of anchor nodes to minimize network cost, localize maximum or almost all nodes in the network with more accuracy. We used one-hop approach; multi-hop (scalable) approach can be used to calculate initial positions. As literature survey tells us that there is not a one single approach that has all the benefits and does not have any drawback at all [33]. It means that we should use hybrid approaches to combine the benefits of different approach. Fuzzy logic is producing very good results it can be used as well.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "A Survey on Sensor Networks," IEEE Communications Magazine, Volume: 40, pp: 102 – 114, 2002

[2] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing," IEEE Computer Society,Volume: 34, pp: 57 – 66, 2001

[3] C. Liu, Kui W and T. He, "Sensor Localization with Ring Overlapping Based on Comparison of Received Signal Strength Indicator", IEEE Conferences on mobile ad-hoc and sensor systems, 2004, pp: 516–518

[4] S. Gezici, Z. Sahinoglu, A. F. Molisch, H. Kobayashi and H. V Poor, "A Two-step Time of Arrival Estimation Algorithm for Impulse Radio Ultra Wideband Systems," Jul 2005

[5] A. E. Moutia, K. Makki and N. Pissinou, "A Time and Power Based Localization Scheme for Indoor WLAN Using Sensor Networks", IEEE Conference on Technologies for Homeland Security, 2007, pp: 117 – 122

[6] A. Malekitabar, H. Aghababa, M. H. Radfar, and B. H.Khala, "Minimizing the Error of Time Difference of Arrival Method in Mobile Networks", IEEE Conference, 2005, pp: 328 – 332

[7] R. Peng and M. L. Sichitiu,"Angle of Arrival Localization for Wireless Sensor Networks", IEEE Communications Society 3rd Annual conference on Sensor and Ad Hoc Communications and Networks, SECON , Volume: 1, 2006, pp: 374 – 382

[8] S. Tian, X. Zhang, P. Liu, P. Sun and X. Wang, "A RSSI- based DV-hop Algorithm for Wireless Sensor Networks", IEEE Conference on Wireless Communications (WiCom), Networking and Mobile Computing , 2007 pp: 2555 – 2558

[9] H. Chen, P. Huang, M. Martins, H. C. So and K. Sezaki, "Novel Centroid Localization Algorithm for Three-Dimensional Wireless Sensor Networks", IEEE Conference on Wireless Communications (WiCOM), Networking and Mobile Computing, WiCOM, 2008 pp: 1 – 4

[10] T. He, C. Huang, B. M. Blum and J. A. Stankovic, T. Abdelzaher, "Range-Free Localization Schemes for Large Scale Sensor Networks", 2003, pp: 81-95

[11] R. Mardeni, Othman and S. Nizam ,"Node Positioning in ZigBee Network Using Trilateration Method Based on the Received Signal Strength Indicator (RSSI)", European Journal of Scientific Research, Vol. 46, No.1, pp: 48-61, 2010

[12] K. Langendoen and N. Reijers, "Distributed localization in wirelesssensor networks: a quantitative comparison", Computer Networks: The International Journal of Computer and Telecommunications Networking, Vol. 43, No. 4, pp: 499–518, 2003

[13] K. Langendoen and N. Reijers, "An Efficient Method Based on Genetic Algorihtms \to Solve Sensor Network Optimization Problem", International journal on

applications of graph theory in wireless ad-hoc networks and sensor networks (GRAPH-HOC) Vol.3, No.1, pp: 18-33, 2011

[14] Q. Zhang, J. Wang, C. Jin, J. Ye, C. Ma, and W. Zhang, "Genetic Algorithm based Wireless Sensor Network Localization", IEE Computer Society Fourth International Conferences on Natural Sciences, Volume: 1, 2008, pp: 608 – 613

[15] G. Yang, Z. YI, N. Mnquan, Y. Keke and X. Tongtong, "An Improved Genetic Algorithm for Wireless Sensor Networks Localization", IEEE Conference Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2010, pp: 439 – 443

[16] S. Lanzisera, D. T. Lin and K. S. J. Pister, "Methods of Sensors Localization in Wireless Sensor Networks", in *Proc. 14th Annual IEEE International Conference and Workshops on the Engineering of Computer- Based Systems*, pp:145-152, 2007

[17] Z. Chaczko, R. Klempous, J, Nikodem and M. Nikodem, "RF Time of Flight Ranging for Wireless Sensor Network Localization", IEE Conference on Workshop on Intelligent Solutions in Embedded Systems (WISES'06), Vienna, 2006, pp: 1-12

[18] M. Dorigo, V. Maniezzo and A. Colorni, "The Ant System: Optimization by a col of cooperating agents", IEEE Transactions on Systems, bernetics–Part B, Vol.26, No.1, pp: 1-13, 1996

[19] F. Neumann and C. Witt, Bioinspired Computation in Combinatorial Optimization – Algorithms and Their Computational Complexity, September 2010

[20] F. Nematy, A. M. Rahmani, M. Teshnelab and N. Rahmani , "Ant Colony Based Node Deployment and search in wireless sensor networks, IEEE conference on computational intelligence and communication networks, 2010, pp: 363-366

[21] R. Putha, L. Quadrifoglio and Emily Zechman"Comparing, "Ant Colony Optimization and Genetic Algorithm Approaches for Solving Traffic Signal Coordination under Oversaturation Conditions, Computer-Aided Civil and Infrastructure Engineering, vol: 27, pp: 14–28, 2012

[22] N. B. Sariff, N. Buniyamin, "Comparative Study of Genetic Algorithm and Ant Colony Optimization Algorithm Performances for Robot Path Planning in Global Static Environments of Different Complexities", in *Proc. 8th IEEE international conference on Computational intelligence in robotics and automation*, pp: 132-137, 2009

[23] A. Pal, "Localization Algorithms in Wireless Sensor Networks: Current Approaches and Future Challenges", Macrothink Institute, Network protocols and algorithms, vol: 2, No. 1, pp: 45-74, 2010

[24] M. Dorigo, T. Stutzle, Ant Colony Optimzation. Bradford: 2004

[25] Y. Xue, H. Lee, M. Yang, Y. Kumarawadu, P. Ghenniwa and W. Shen, "Performance Evaluation of NS-2 Simulator for Wireless Sensor Networks," Canadian Conference On Electrical and Computer Engineering, 2007 pp: 1372 - 1375

[26] D. Niculescu , B. Nath ," Localized positioning in ad hoc networks", Ad Hoc Networks, vol: 1, pp 247–259, 2003

[27] R. Mulligan, "Coverage in Wireless Sensor Networks: A Survey", Network Protocols and Algorithms, Macrothink Institue, Vol. 2, No. 2, pp: 27-53, 2010

[28] A. Howard, M. J. Matari and G. S. Sukhatme," An Incremental Self- Deployment Algorithm for Mobile Sensor Networks", Autonomous Robots, Special Issue on Intelligent Embedded Systems, 13(2), pp: 113-126, 2002

[29] N. Dziengel, G. Wittenburg, and J. Schiller," Towards Distributed Event Detection in Wireless Sensor Networks", 2008, pp: 1-3

[30] P. Liu, X. Zhang, S. Tian, Z. Zhao, P. Sun, "A Novel Virtual Anchor Node-based Localization Algorithm for Wireless Sensor Networks", 2007

[31] N. B. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller, "Anchor-Free Distributed Localization in Sensor Networks", MIT Laboratory for Computer Science, Technical Report #892, April 15, 2003

[32] K. K. Almuzaini, A. Gulliver," Range-Based Localization in Wireless Networks Using Density-Based Outlier Detection", Scientific Research, Wireless Sensor Network vol: 2, pp: 807-814, 2010

[33] K. Li, L. Kang, W. Zhang, B. Li, "Comparative Analysis of Genetic Algorithm and Ant Colony Algorithm on Solving Traveling Salesman Problem", IEEE International Workshop on Semantic Computing and Systems, 2008, pp: 72-75

[34] L. E. A. Santana, L. S. A. M. P Canuto, F. Pintro and K. O. Vale, "A Comparative Analysis of Genetic Algorithm and Ant Colony Optimization to Select Attributes for an Heterogeneous Ensemble of Classifiers", IEEE Conference on Evolutionary Computation (CEC), 2010, pp; 1-8

[35] V. Maniezzo, L. M. Gambardella, F. d. Luigi, "Ant Colony Optimization", pp: 1-21, 2004