Node Localization using Naive Bayes Classifier and Trilateral Algorithm

K. Madhumathi, T. Suresh ., R. Maruti

Abstract: *Node localization is an important problem considered among the researchers in the area of Wireless Sensor Networks (WSN). The WSN is formed by a group of sensor nodes having limited energy and other resources that transfers data among each other or to a base station in an ad-hoc fashion. The estimation of the geo location (co-ordinates in the two-dimensional space) of the sensor nodes is essential for ensuring the QoS within the network. The different applications of WSN require varied level of accuracy in the estimation of the location of the sensor nodes. Different localization schemes are adopted in the literature for better estimation of the node location and each of them has both merits and demerits. This paper focuses on analyzing the different node localization mechanism used in the WSN and to identify various issues and challenges in the estimation of the node location. This paper also proposes an optimal approach with less computational effort and high accuracy in prediction based on trilateration algorithm and the RSSI (Received Signal Strength Indicator) values extracted from the target nodes antennas. The network is segmented in to different blocks of unequal size and the block number in which the node is present will be identified using the naive bayes classifier.*

Keywords: RSSI, naive bayes classifier, trilateration, node localization, wireless sensor network

I. INTRODUCTION

In the future applications of wireless communication network, the accurate and less computational, and reliable real time estimation of the location of sensor node and location-based services are essential [1, 2]. At present the WSN are employed for variety of applications including monitoring of natural resources and environment, object tracking, and others. In all the applications the role of sensor nodes is to sense the respective environment, collect the required data, and forward it to the base or control station respectively. In geographic routing protocols and some of the hierarchical clustering approach the estimation of the sensor nodes location is very essential [3, 4]. The problem of node localization has been studied exhaustively in the recent years and majority of the literature follows an anchor node-based approach by which the position of those nodes are estimated using sophisticated GPS device or mathematical estimation [5-7]. The normal sensor nodes calculate their position based on their distance with respect to the anchor nodes using triangulation or multi-lateration schemes. A small proportion of anchor nodes in the network are used to estimate the position of the other neighboring nodes with satisfactory level of accuracy [8-9].

The localization methods proposed in the literature can be broadly categorized in to range-free and range-based approaches. The categorization is done based on whether the distance between the nodes has to be measured before estimating the location of the nodes. The range based approach includes trilateral, multilateral, DEP, and other algorithms. The multilateral algorithm gives more accurate estimation of location of the nodes and it is most widely used range based estimation algorithm [10]. Based on the distance between the anchor and unknown nodes and the least square method the position is estimated. The selection of anchor node will have a great impact on the accuracy of the estimated position. The range-free approach includes centroid localization, distance vector hop, and approximate point-in triangulation test algorithms. The centroid algorithm [11] depends on the topology of the network which is influenced by the communication range of the nodes present in the network. In practice various factors disturb the range of the nodes in the network and eventually the accuracy of the estimation will become poor. The APIT algorithm [12] iteratively explores the unknown area and determines the relationship between the sensor nodes and the neighboring anchor node. When the network is sparsely connected the APIT algorithm cannot be employed to detect the location of the nodes. The information of each anchor node is collected through distance vector routing protocol in another popular localization algorithm called as DV-Hop algorithm [13]. The work in [14] compares the results of the APIT and the DV-HOP algorithm based on various parameters and the results showed that the DV-HOP algorithm is performing better when compared to its counterpart irrespective of the count of the anchor nodes. When the hop count between the anchor node and the sensor nodes are increased then the accuracy of the estimated location starts declining.

The architecture of WSN can be considered as a weighted undirected graph where the distance between the adjacent nodes is considered as the weights of the edge connecting the nodes. The MDS algorithm accepts this distance matrix as an input and utilizes Floyd algorithm to estimate the shortest path between any two given nodes. The accuracy of the algorithm

Ms. K. Madhumathi, Assistant Professor, Department of BCA, Anna Adarsh College for Women, Chennai,

Dr.T.Suresh,, Associate Professor, Department of Computer Science and Engineering, Annamalai University, TN, India.

Dr. R. Maruti, Professor, Ponnaiyah Ramajayam Institute of Science and Technology, Chennai.

Retrieval Number: C5604029320/2020©BEIESP
DOI: 10.35940/ijeat.C5604.029320

2467

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

is linearly proportional to the hop count between the nodes. The location estimation accuracy will be poor when the hop count is larger. The range-based location estimation approaches requires sophisticated hardware and exhibits high accuracy. Majority of real time applications employs range-based algorithms.

This paper focuses on estimating the location of the target nodes in two levels. Initially the network is segmented in to blocks and the nodes are labeled according to their blocks. A naive bayes classifier is trained with the labeled data to identify the block number of the target nodes based on the RSSI values extracted from the antennas and based on the measured distance from the RSSI values the exact location of the nodes are determined using the trilateral algorithm.

II. LITERATURE REVIEW

Some of the literatures have combined the range based and the range free approaches along with the heuristic optimization algorithm for estimating the location of the nodes in the network. The objective of these combined approaches is to build an optimal location estimation model using the hop count and distance information between the nodes. Then using the optimization algorithm (heuristic) an optimal solution will be calculated by iteratively redefining the parameters of the location estimation model. Few of the optimization algorithms used in such literatures includes Artificial Fish Swarm Optimization (AFSA) [12], Particle Swarm Optimization (PSO) [14] and Flower Pollination Algorithm [13]. For a real time application these heuristic algorithms are not suitable since they are computationally expensive and consume more time in a centralized location estimation scenario. The heuristic algorithms greatly rely on the initial parameters and they may adversely affect the performance of the algorithm.

Ultra-wide Band (UWB) is a technology for transferring the data between transmitter and receiver through an extremely short nanosecond pulses and does not uses any carrier signal. UWB is characterized by low transmission power, and resistance to antifading. The ranging techniques can be of following types Time of Arrival (TOA), Time Difference of Arrival (TDOA), Received Signal Strength Indication (RSSI), and Angle of Arrival (AOA). The TOA based approach requires the clock signals to be synchronized and the hardware requirements are expensive. TDOA works on the principle of using transmission speed difference for measuring the distance between the sensors. The ranging accuracy is higher and requires an ultrasonic launcher. AOA works based on the angle information extracted from the antennas. RSSI measures the distance between the nodes based on the loss of power of wireless signal during transmission. RSSI approach does not require any additional hardware equipment's but rather the embedded chip in the sensor has the potential to estimate and transmit the signals. Also, when compared to the range-free algorithms the RSSI based estimation has merits.

Consider a large-scale network with anchor nodes distributed and lying on the edges and there is limitation in the transmission range. The location of all unknown nodes cannot be estimated by 1-D multilateral algorithm. Due to this the

nodes must be re-arranged in a hierarchical order and then the multilateral algorithms can be used to estimate the position of the nodes. When the position of the node is estimated then they can be considered as another anchor node. Until all the geographical position of the nodes is estimated, the multilateral algorithm works in an iterative procedure and generates a cumulative error. When the level of the nodes is higher, then the cumulative error magnitude will also be higher.

III. PROBLEM STATEMENT

Let there are N nodes in a WSN and they are labeled as *1,2,....,N*in a 2-D space and the location of *M* nodes out of *N* total sensor nodes is known where (M<<N). Then the location of N-M nodes has to be estimated. Using the RSSI signal the distance between a node and its neighboring nodes can be estimated. The distance between the nodes is termed as $\hat{d}_{ij}$. Let $X = \{x_i\}_{i=1}^{N}; x_i = [x_i, y_i] \in R$ denotes the real coordinates of all the sensor nodes. $A = \{x_i\}_{i=1}^{M}$ denote the set of anchor nodes. The position of the anchor nodes will be in the edge of the network area. The average localization error is measured using the mathematical equation given below where R is the transmission range of the node in the network. The aim is to reduce the average location error and estimate the position of the individual nodes more accurately.

$$Error_A = \frac{\sum_{i=1}^{N-M} \sqrt{(x_i - \tilde{x}_i)^2 + (y_i - \tilde{y}_i)^2}}{(R * (N - M))} \qquad \text{EQ. 1}$$

One of the most important characteristics of wireless transmission is that when the transmission distance is increased then the signal starts attenuating. The log-normal shadowing model is used to generate RSSI sample as a function of distance measured in RF channels [15].

$$P_R(d) = P_T - PL(d_0) - 10\eta \lg \left(\frac{d}{d_0 + X_\sigma}\right) \qquad \text{EQ. 2}$$

Where P_R _received signal power

P_T _ transmission Power

$PL(d_0)$ – Path Loss

d_0 _ reference distance

η – Path loss exponent

X_σ - Gaussian noise whose mean is zero and standard deviation is σ.

IV. METHODOLOGY

The location estimation problem is solved using a probabilistic approach based on naïve bayes classifier.

The trained naïve bayes classifier classifies the RSSI measurements in to certain block of the network area. The aim is to locate the position of the node within the network area instead of finding the two-dimensional coordinate.

The block information for each corresponding RSSI vectors was saved earlier during offline phase. The classifier maps the given RSSI vector input to a block number in the network area. The block number is treated as class and the RSSI vector forms the input to the classifier.

$$P(R|v) = \frac{P(v|R).P(R)}{P(v)} \qquad \text{EQ. 3}$$

The posterior probability $P(R/v)$ is estimated

Retrieval Number: C5604029320/2020©BEIESP
DOI: 10.35940/ijeat.C5604.029320

2468

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

based on the *P(R)* the prior probability, *P(v/R)* the likelihood function which determines the observing the pattern *v* in case of node being in block *R,* and the *P(V)* termed as the evidence which follows a normal distribution with mean *μ* and the standard deviation *'σ'*. The naïve bayes classifier requires less computational efforts and accepts input that has conditional independent features. The probability of a node being in a block *'R'* of network can be estimated based on the estimated RSSI values as follows:

$$P(R|v_1,\ldots,v_n) = \frac{1}{Z} P(R) \prod_{i=1}^{n} P(v_i|R) \qquad \text{EQ. 4}$$

The naïve bayes does calculates the probability of node being in each of the block and the required output is the block number whose the probability is higher.

$$R = argmax \left\{ P(R_j) \prod_{i=1}^{n} P(v_i|R_j) \right\} \qquad \text{EQ. 5}$$

The mathematical relationship between the RSSI expressed in *dBm* and the distance between the anchor and the sensor nodes. The power of the received signal P_R, can be estimated using the power of transmitted signal P_T, the gain of antenna in the receiver G_R, the gain of antenna in the transmission end, the wavelength of transmitted signal λ, the distance *d* in meters and the signal propagation constant *n* based on the mathematical formula given below.

$$P_R = P_T \frac{G_T G_R \lambda^2}{(4\pi)^2 d^n} \qquad \text{EQ. 6}$$

By converting the terms in watts to dBm then the equation above is rewritten as

$$P[dBm] = 10.\log_{10}(P[W].10^3) \qquad \text{EQ. 7}$$

Based on the above form of the equation the case of a 1-meter reference distance the relationship among the distance between the nodes and the received power can be expressed as

$$RSSI = -(10.n.\log_{10} d - A)$$

where A denotes the power of the received signal in dBm with two antennas separated apart in 1m distant and n denoting the loss parameter. The distance d can be calculated easily as

$$d = 10^{\frac{A-RSSI}{10.n}}$$

The parameter *A* depends upon the physical properties of the communication device and the value of *n* is constant and fixed to be equal to 2. For outdoor environment the value of *A* and *n* need to be found empirically where the channel may suffer due to fading and path-loss. The value of *A* and *n* are estimated by collecting the RSSI values between the two nodes separated at different distance. The plot shown in Fig. 1 below presents logarithmic interpolation of RSSI values estimated as per the above equation.
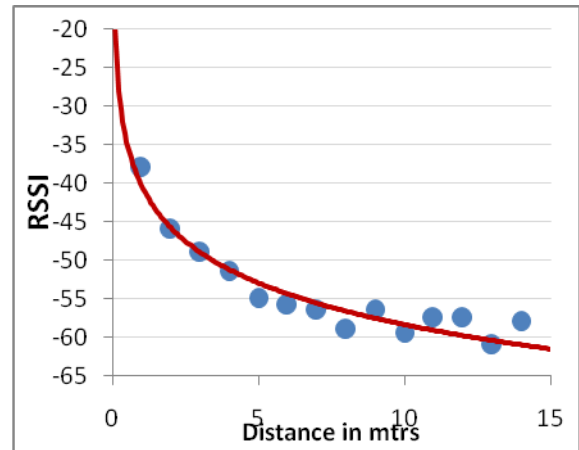

Fig. 1 Plot of logarithmic Interpolation of RSSI values

V. V. EXPERIMENTS & RESULTS

The sensor nodes are Based on the RSSI values the distance d_i is estimated and as per the Trilateration algorithm circles centered around the anchoring nodes are computed. The radius r is assumed to be equivalent to d. The target nodes lies exactly at the intersection of the circles as presented in Fig. .

$$(x - x_i)^2 + (y - y_i)^2 = r_i^2$$

After estimating the centre of the circles and its respective radius then the equation of the circle is mathematically represented as in Eq.
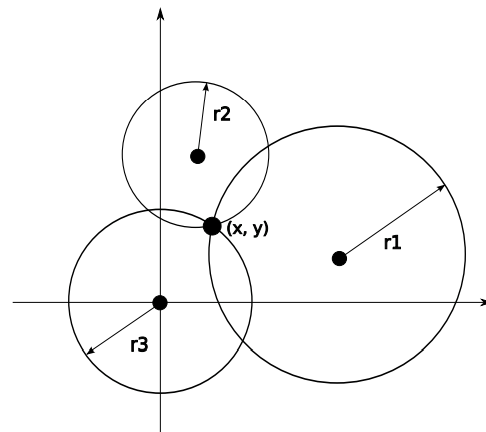

Fig. 2 Distribution of Anchor node and sensor node

The intersection of the three circles can be calculated by solving

$$\begin{cases} ((x - x_1)^2 + (y - y_1)^2 = r_1^2) \\ (x - x_2)^2 + (y - y_2)^2 = r_2^2 \\ (x - x_3)^2 + (y - y_3)^2 = r_3^2 \end{cases}$$

If the radius r_i is not estimated accurately instead of an intersection point an area will be extracted or they may not intersect at all. Then two more circles are drawn centered around the anchor nodes, radius are increased till the circles intersect. For accurate estimation of the sensor nodes location, multiple antennas are used on the target nodes. The antennas are separated by the 10cm and the RSSI values are read from the antennas alternatively.

The average of the two RSSI measurements is given as input to the localization.

VI. CONCLUSION

This study reviewed different ranging algorithms and proposed a localization scheme using probabilistic approach and trilateral algorithm. The RSSI values are accurately measured by using multiple antennas in the sensor nodes. The RSSI values are later used to compute the distance between the anchor nodes and the target nodes. For measuring the exact location of the target nodes trilateral algorithm is adopted and whose accuracy greatly relies on the accuracy of the estimated radius of the circles around the anchor nodes. The naïve bayes classifier predicts the block number with less error and the training the classifier does not require much labeled data i.e. with less number of training samples model learns to classify the location of the target nodes in the network area. In overall the accuracy of the results are found to be accurate when compared to the results in the literature.

REFERENCES

1. Vossiek, M.; Wiebking, L.; Gulden, P.; Wieghardt, J.; Hoffmann, C.; Heide, P. Wireless local positioning. IEEE Microw. Mag. 2003, 4, 77–86.
2. Flora, C.D.; Ficco, M.; Russo, S.; Vecchio, V. Indoor and outdoor location based services for portable wireless devices. In Proceedings of the 25th IEEE International Conference on Distributed Computing Systems Workshops, Columbus, OH, USA, 6–10 June 2005; pp. 244–250.
3. Paul, A.K.; Sato, T. Effective Data Gathering and Energy Efficient Communication Protocol in Wireless Sensor Network. In Proceedings of the Wireless Personal Multimedia Communication (WPMC'11), Brest, France, 3–7 October 2011; pp. 1–5.
4. Al-Karaki, J.N.; Kamal, A.E. Routing techniques in wireless sensor networks: A survey. IEEE Wirel. Commun. 2004, 11, 6–28.
5. Chowdhury, T.; Elkin, C.; Devabhaktuni, V.; Rawat, D.B.; Oluoch, J. Advances on Localization Techniques for Wireless Sensor Networks. Comput. Netw. 2016, 110, 284–305.
6. Halder, S.; Ghosal, A. A survey on mobile anchor assisted localization techniques in wireless sensor networks. Wirel. Netw. 2016, 22, 2317–2336.
7. Paul, A.K.; Sato, T. Detour Path Angular Information Based Range Free Localization in Wireless Sensor Network. J. Sens. Actuator Netw. 2013, 2, 25–45.
8. Yassin, A.; Nasser, Y.; Awad, M.; Al-Dubai, A.; Liu, R.; Yuen, C.; Raulefs, R.; Aboutanios, E. Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications. IEEE Commun. Surv. Tutor. 2017, 19, 1327–1346.
9. Paul, A.K.; Li, Y.; Sato, T. A Distributed Range Free Sensor Localization with Friendly Anchor Selection Strategy in Anisotropic Wireless Sensor Network. Trans. Jpn. Soc. Simul. Technol. 2013, 4, 96–106.
10. R. Gong, J. Zhao, Y. Huang, and X. Meng, "WSN multilateral localization algorithm based on MLE," in Proceedings of the International Conference on Network Computing and Information Security, NCIS '11, pp. 426–429, 2011.
11. Y. Bai, C. M. Li, and Y. Xue, "A Centroid Localization Algorithm for Wireless Sensor Networks Based on RSSI," Applied Mechanics and Materials, vol. 303-306, no. 7, pp. 197–200, 2013.
12. H. B. Tan and F. Liu, "Research and implementation of APIT positioning algorithm in WSN," in Proceedings of the 9th International Conference on Fuzzy Systems and Knowledge Discovery, pp. 2212–2215, IEEE, Sichuan, China, May 2012.
13. D. Niculescu and B. Nath, "DV based positining in ad hoc networks," Kluwer J. Telecommunication Systems, vol. 22, no. 1, pp. 267–280, 2017.
14. S. Anthrayose and A. Payal, "Comparative analysis of approximate point in triangulation (APIT) and DV-HOP algorithms for solving localization problem in wireless sensor networks," in Proceedings of the 7th IEEE International Advanced Computing Conference, IACC '17, pp. 372–378, January 2017.

I. AUTHORS PROFILE

Ms. Madhumathi K is currently working as Assistant Professor in Anna Adarsh College for women, from Chennai, Tamil Nadu. She is pursuing her Ph.D. Degree in Computer Science from Annamalai University. She received M. Phil from Vinayaka Mission University, Salem in the year 2009, and Degree of Master of Computer Applications from SA Engineering College, Anna University in the year 2008. She did her Bachelor's Degree in Computer Science from CTTE College for Women, Chennai in the year 2005. She has published various research papers in reputed International journals..

Dr.T.Suresh received his Bachelor of Engineering Degree in Computer Science and Engineering from Shanmuga College of Engineering, Thanjavur. He did his M.B.A degree from Madurai Kamaraj University. He did his post-graduation and Ph.D. Degree in Computer Science and Engineering from Department of Computer Science and Engineering, Faculty of Engineering and Technology from Annamalai University. He has contributed around 40 technical papers and in various journals & conferences. He is currently working as an Associate Professor in the Department of Computer Science and Engineering, Annamalai University, Annamalai Nagar, Tamil Nadu, India. His area of interest includes MANET, VANET, WSN, Cloud Computing etc.,

Dr.R.Maruthi received her Bachelor'sDegree andMaster of Computer Applications from Bharathidasan University in the year 1997 and 2000 respectively. She completed her M. Phil from Alagappa University in the year 2008 and Ph.D. Degree in Computer Science from Mother Terasa Women University in the year 2011. She has contributed more than 15 research papers for indexed journals. She has more than 17 years of teaching experience and currently working as Professor in PRIST University

Retrieval Number: C5604029320/2020©BEIESP
DOI: 10.35940/ijeat.C5604.029320

2470

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication