Abstract
Growing convergence among mobile computing devices and embedded technology promotes the development of “context-aware” applications, where location is the most essential topic. In present time, several systems are available in case a laptop, PC with a WLAN card and WLAN outdoor localization, such as GPS, assisted GPS and other access points (APs). An important characteristic of this system systems working on cellular networks. However, there is no proper way to find exact location to use an existing proper location system for indoor scenarios. A novel indoor positioning system based on received signal strength (RSS) in wireless networks with high accuracy is presented in this paper. The three improvement mechanisms, called signal strength filtration, user location filter and path tracking assistance, are employed to improve the positioning accuracy of the system. The comprehensive performance of the proposed system is analyzed in detail and compared with the Radar system.

Keywords: Indoor location, probabilistic Method, reference tag, RFID.

1. INTRODUCTION
There are two methods to get the location estimation, distance related method and distance unrelated method. The first method calculate the distance from the access points (AP) to the mobile terminal (MT), and then estimate the position by Trilateration or Triangulation algorithm. TOA(Time of Arrival), AOA(Arrival of Angle), TDOA(Time Difference of Arrival) and RSSI-based techniques are the four principal techniques to get the distance estimation. The distance unrelated method doesn’t need distance information and it mainly includes proximity based algorithm and distance in hops based algorithm. The RSSI-based technique estimates the distance by the RF propagation loss model which is a simple mathematical expression representing the relationship between the RSSI and the distance between the sender and the receiver. However, RSSI is influenced by so many parameters and establishing an appropriate RF propagation loss model for off the shelf APs is not possible. One another method is called Fingerprinting.

Fingerprinting
Fingerprinting is the most widespread positioning technique used for WLANs. It is based on a set of RSS measurements taken of a target from multiple access points and comparison of the results with a previously compiled database. The method has several advantages. No time synchronization is required, as for TDOA. The reading of RSS is inherent to the IEEE 802.11 protocol, and no special hardware is needed. Tags or devices based on the IEEE 802.11 standard can be tracked. The method is particularly applicable to indoor networks as the vagaries of multipath propagation are automatically accounted for in the reference database. On the negative side, the method implies creating a database for the area to be covered, and changes in AP deployment and physical features of the environment require updating the database. The method is computationally intensive and a special location server is required for position outputs. The fingerprint database is created from a grid mapped to a floor plan of the coverage area that includes physical characteristics—partitions, wall materials, furnishings and the position of access points. Ray trace software creates vectors of signal strengths at grid positions throughout the area, and actual measurements are added as needed for increased accuracy. RF signal strength prediction is based on reflection, attenuation and multiple transmission paths between grid points and each AP. Grid points can represent an area as small as 15 cm square. Realtime signal strengths from a target to all access points in range are compared to the database to estimate the target location. Targets can be tracked to an accuracy of a few meters.

2. TRIANGULATION ALGORITHMS OF INDOOR POSITIONING SYSTEM
2.1. Positioning Algorithms
There are at least four location fingerprinting-based positioning algorithms using pattern recognition technique: probabilistic methods, k-nearest-neighbor (kNN), neural networks, and support vector machine (SVM). In wifi tracking system, the Probabilistic Methods and kNN Method are used to forecast and calculate the user location.
1) Probabilistic Methods: One method considers positioning as a classification problem. Assuming that there are n location candidates L1; L2; L3; ∙ ∙ ∙ ; Ln and s
is the observed signal strength vector during the online stage, the following decision rule can be obtained: Choose Li if \( P(L_i) > P(L_j) \) for \( i; j = 1; 2; 3; \ldots; n; j = 6; \) Here, \( P(L_i) \) denotes the probability that the mobile node is in location \( L_i \), given that the received signal vector is \( s \).

2) kNN Methods: The kNN averaging uses the online RSS to search for \( k \) closest matches of known locations in signal space from the previously-built database according to root mean square errors principle. By averaging these \( k \) location candidates with or without adopting the distances in signal space as weights, an estimated location is obtained via weighted kNN or unweighted kNN. In this approach, \( k \) is the parameter adapted for better performance. KNN method is based on estimating the position \( i \) depending on the average (in physical space) of the coordinates of the \( k \) closest reference points to the received RSSI vector \( x \).

3) Bayesian Learning Method: Bayesian learning method employs the Bayesian theorem to estimate the unknown position, given the signal measurements \( x \) in position \( i \), then the position \( i \) is calculated by the posteriori \( i = \text{argmax}_i P(i|x) = \text{argmax}_i P(x|i)P(i)/P(x) \) where \( p(x|i) \) is the probability of receiving a sample \( x \) from position \( i \) and \( p(i) \) is the probability of a MT being at this position \( i \) which initially can be considered as uniform in the location area, \( p(x|i) \) can be calculated from the calibration information table (CIT) that is built ahead of time. Therefore, the location estimation problem becomes a standard maximization problem. The main drawback of the above method is the large number of calibration samples necessary to construct the distribution \( p(x|i) \). One possible approach to reduce the number of calibration samples is clustering.

3. THREE IMPORTANT MECHANISMS BASED ON INDOOR POSITINING SYSTEM

3.1. Signal Strength Filtration: The received signal strength is influenced by various types of noise, such as band interference among access points, sudden person walking and opening or closing of windows and doors, etc. These disturbances of received signal strength occurs decrease the accuracy of indoor positioning systems. Therefore, three different filtering methods named Max Filter, Limit Filter and MA Filter are studied in detail in this paper in order to improve the positioning accuracy of system because of harsh multi-path environment and all kinds of disturbance in indoor areas.

(i) Max Filter: The received signal strength is interfered by noise, such as sudden person walking and opening or closing of windows and doors. However, the noise always decreases the value of RSS, and does not increase the value of it. Therefore, we can calculate the maximum RSS value of \( N \) continuous samples in order to filter the lower RSS which is interfered by noise.

(ii) Limit Filter: The initial value is set to the average value of samples. If the changes between the next sample and the current sample exceed the maximum Limit, then the next sample is invalid. Otherwise, the next sample is valid.

(iii) MA Filter: In Move Average Filter, called MA Filter, the size of time sliding window is set to Window Size.

3.2. User Location Filter:
The number of received signal strength samples scanned by wireless network card is less and the standard deviation of forecasted user location is larger, which occurs seriously decrease the performance and stability of indoor positioning systems. Therefore, we can use Kalman filter algorithm to filter the forecasted user location in order to further improve the performance and stability of indoor positioning system.

3.3. Path Tracking Assistance:
In real life, people always move according to a certain paths, for example, walking from one side of corridor to the other side of it, not pass through the wall, etc. Therefore, all kinds of possible paths could be predefined, then the indoor positioning systems can make full use of these paths to further improve the positioning accuracy.

4. PROPOSED METHOD
We use the Euclidean metric (L2) for computing distances. It is easy to modify our kNN algorithm to accommodate other distance metrics. Our implementation makes extensive use of the two distance estimates MinDist and MaxDist (Fig. 1). Given two blocks \( q \) and \( s \), the procedure MinDistance \( q; s \) computes the minimum possible distance between a point in \( q \) to a point in \( s \). When a list of blocks is ordered by their MinDist value with respect to a reference block or a point, the ordering is called a MinDist ordering. Given two blocks \( q \) and \( s \), the procedure MaxDistance \( q; s \) computes the maximum possible distance between a point in \( q \) to a point in \( s \). When a list of blocks is ordered by their An ordering based on MaxDist is called a MaxDist ordering. The kNN algorithm identifies the \( k \) nearest neighbors for each point in the data set. We refer to the set of kNNs of a point \( p \) as the neighborhood of \( p \). While the neighbourhood is used in the context of points, locality defines a neighborhood of blocks. Intuitively, the locality of a block \( b \) is the region in space that contains all the kNNs of all points in \( b \). We make one other distinction between the concepts of neighborhood and locality. In particular, while neighborhoods contain no other points other than the kNNs locality is more of an approximation and thus the locality of a block \( b \) may contain points that do not belong to the neighborhood of any of the points contained within \( b \). Itrst builds the locality for a block and later searches the locality to construct a neighborhood for each point contained within the block. The pseudo-code
The labeling scheme assigns each block a label concatenated with the number of points that it contains. q is the query block. Blocks x and y are selected based on the value of MaxDist, while blocks b, e, f, i, d, p, q, k, m, and o are also selected as their MinDist value from q. p PruneDist.

The mechanics of the algorithm are described in Fig. 2. The figure shows q in the vicinity of several other blocks. Each block is labeled with a letter and the number of points that it contains. For example, suppose that k=3, and let Q={A,B,C,D,E,F,I,J,K,L,M,O,P,Q,X,Y } be a decomposition of the underlying space into a set of blocks. The algorithm rst visits blocks in a MaxDist ordering from q, until 3 points are found. That is, the algorithm adds blocks X and Y to S and PruneDist is set to MaxDist(Q,Y). We now choose all blocks whose MinDist from q is less than PruneDist resulting in blocks B,E,F,I,D,P,Q,K,M and O being added to S.

MLR Algorithm Description

We represent an improved algorithm-machine learning reference (MLR) which takes advantage of the Bayesian and WKNN idea and improves the overall location accuracy by introducing the reference tag concept. The final position estimation is obtained by followings:

(i) Input the objective vector to the Bayesian model, then the output which is two-dimensional coordinate is the rough position estimation of the target.
(ii) Find the nearest reference tags in accordance with the WKNN algorithm and compute the amending position.
(iii) Regard the average of above two positions as the final position of the target.

In the first phase, we take the following steps to derive the primary position:
(1) Build a Bayesian model of which the input is an observation vector of which the dimension is the reader number and the vector elements is the RSSI for each reader and the output is coordinate based on the calibration data.
(2) Input the objective vector and get the final result. The critical issues are collecting of the calibration data and the training procedure.

In accordance with the Bayesian rule, we only need to compute the product of \( p(l) \) and \( p(o | l) \) . \( p(l) \) can be considered as the occurrence frequency of samples which is obtained in \( l \) of all training samples. We adopt a probabilistic method Kernel Method to compute

\[
p(o|l) = \frac{1}{ni} \sum_{i=1}^{\infty} k(o,oi)
\]

where \( ni \) is the number of training vectors in \( l \) and \( k(o|oi) \) denotes the kernel function.

In the second phase, we define the signal strength vector of a tracking tag as \( S={S_1,S_2,S_3,...,S_n} \) where \( S_i \) denotes the signal strength received on reader \( i \), where \( i \in \{1...n\} \) , and for reference tags, the corresponding signal strength vector as \( \Theta=(\Theta_1,\Theta_2,...,\Theta_n) \)

Fig. 2. Illustration of the workings of the BUILDLOCALITY algorithm.
When there are m reference tags, a tracking tag has the vector \( E=(E_1,E_2,\ldots,E_m) \). We select k reference tags which have lower E value.

5. EXPERIMENTAL ANALYSIS
The experiments were based on actual rectangular scene (30x20m) and each corner of the area was deployed with a reader and the area was divided by 2-meter grids on average. The training data is collected systematically on the center of each grid, which we call calibration points, 40 observations are recorded, each consisting of RSSI from all readers. They placed reference tags on the four acmes which we call reference points of each grid. Similarly, the reference data is also consisting of RSSI value from all readers and they are collected on reference points. At each of the reference point, 20 observations are gathered. We select the junctions of each grid as the test points, at each of which we collect 20 observations.

![](image)

Fig.3. Relation between length of history (test points) and accuracy

From the above graph, we can see for the same test data, our algorithm has much higher location accuracy than the machine learning method in the whole level, and it appears higher precision with the increase of the number of test observations, so we can affirm that the location accuracy is related with the history data closely. Our observations are obtained in certain time interval, so if many observations are detected on certain location, it means that the target object is moving at low speed or it doesn’t move at all. Similarly, if few observations are detected, we can think the target travels in so high speed that we can’t record its observation value. However, the precision is related with other factors also. Such as the number of readers and training samples, the deploy way of readers and tags and the parameters of MLR method and so on. In view of the paper space, we only present the experiment conclusion which is when there are four readers and they are deployed in the four corners, 15 observations at each calibration point, we get much higher accuracy.

6. CONCLUSION
In this paper, we research the indoor location method based on RFID technology and RSSI. The MLR algorithm and its periodicity of RFID make us measure RSSI quickly and accurately. But the location accuracy of machine learning usually cannot meet the user’s need. So we bring in the reference tag concept to amend the rough accuracy. In the kNN based build_locality algorithm, we can only find out the max and min distance and it is very complicated in comparison to MLR algorithm. The result of experiments showed that the improved idea is feasible in the actual location but amount of effort must be used in the procedure of collecting training data. The main idea is algorithm has much higher accuracy. It should be emphasized that this algorithm is fit for the objects which move in low speed. The achieved positioning accuracy with the trials is sufficient to provide basic location Based Services (LBS) such as, providing advertisement messages or providing warning messages if weather or traffic conditions at a particular area change and tracking people with special needs within buildings.

The future work is the Research on the self-learning process RTT (Round trip-time) measurement using RTS/CTS frames estimate the coefficient depending on the environment (multipath radio channel) and other factors related to the traffic load of the network, such as the APs’ processing time. This might be a way to avoid having to carry out the calibration process and the empirical and manual calculation of the readings.

7. ACKNOWLEDGMENT
The authors are grateful to Galgotias college of engineering and technology Gr. Noida India, for their support of this work.

8. REFERENCES
2. Binghao Li, Yufei Wang, Hyung Keun Lee, Andrew Dempster, and Chris Rizos, “A New Method of Yielding a Database of Location Fingerprints in WLAN”, Communications, IEE Proceedings, 2005 vol 4-5

IJCEM
www.ijcem.org
7. Tayebi, A., J. Gomez, F. Saez de Adana, and O. Gutierrez “The application of ray-tracing to mobile localization using the direction of arrival and received signal strength in multipath indoor environments," vol-91
8. Veljo Otsason, Alex Varshavsky, Anthony LaMarca and Eyal de Lara, “Accurate GSM indoor location”, 2005