# GSM Mobile Positioning for Rescue and Law Enforcement

**Abdellatief Abouali[1], Gamal Othman[2], Tamer Elnady[3], Ahmed Yehia[3], Ahmed Ali[3]**


## Abstract

*This study, proposes algorithms and methodologies for GSM mobile positioning using power reports sent through the network. The intent application for these algorithms and methodologies are rescue and law enforcement. Power reports used by network protocol to enable the communication process however they suffer a high noise factors that include fading, multi-paths, nature of the propagation area and media, and existence of line of site between the antennas. The methodologies proposed offer two stages of noise reduction or filtering. First stage applies to the power reports and the second applies to the candidate location points of the target. The filters included in the study are: mean, mid, clustering, regression, voting, and mean k out of m. The methodologies include a proposed method for estimating cells effective radiated power using simple GSM modem equipped with GPS receiver. The online cells database is used to get cells ID and initial position. The basic steps of the positioning based on power reports are: using the power reports to estimate to cells distances, form a table of possible positions, and filter the possible positions to get the likely position. The results of the study show competitive results in areas of different natures.*


**KEYWORDS**: **GSM Network; Signal Processing, propagation models; filters; Positioning; Cell;**

## I. Introduction

Mobile (cellular) communication system provides wireless communication services to users in geographical regions called *cells*. The established wireless communication links a user device, mobile station (MS), with a fixed infrastructure. The fixed infrastructure

1 ElshrooK Academy aabouali4@gmail.com, 2 BUE gamal_os@hotmail.com
3. EinShames Engineering Collage tamer.elnady@eng.asu.edu.eg

belongs to a specific provider connects the MS to: other MS's within the same provider network, other mobile providers or the existing landlines networks for voice, data services or both.

The mobile network infrastructure includes two basic parts: radio access network RAN, and core network. The radio access network (RAN) consists of Base Stations (BSs) and Base Stations Controller (BSC). The BS provides radio interface to the MS. A BS covers a single cell. The BSC control multiple BSs and allow mobility between adjacent cells. The GSM core network consists of components such as the HLR (Home Location Register) which represents the subscriber's information database, and the MSC (Mobile Switching Center) which interfaces with the landline phone network, the PSTN (Public-Switched Telephone Network). The MSC controls the voice calls setup, release, and routing. Also, it performs functions such as generation of billing information. The MSC has control responsibility on the BSCs attached to it. The MSC is integrated with a temporary database called Visitor Location Register (VLR) that stores subscriber information that are under its area of control. Authentication Center (AuC) is implemented as integral part of HLR. The AuC functionality includes performing the calculations needed for user's authentication and ciphering. A cellular service provider's system also includes services network that interfaces with the core networks to provide special services such as: prepaid calling, ring-tone downloads, and voice and/or text messaging [1][3]. Figure (1) depicts network main architecture.

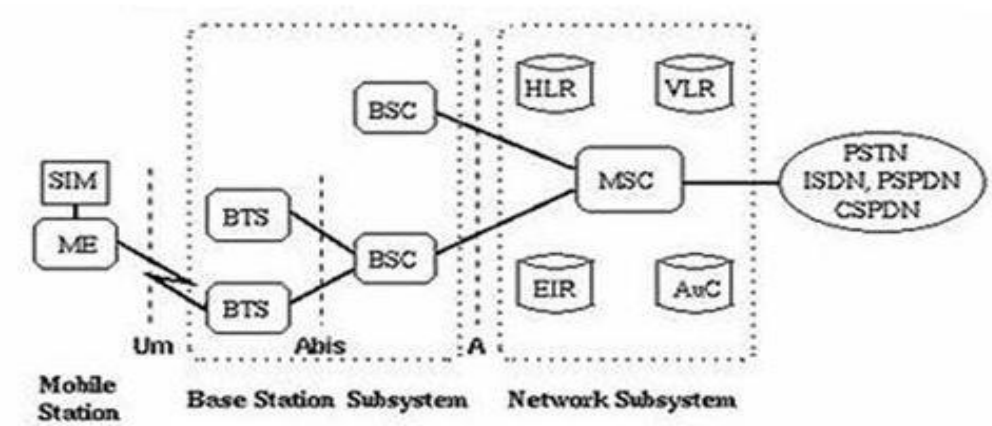


Figure (1) GSM network basic architecture

One of the basic uses of technologies is Law enforcement. Such a use enables law enforcement officers to monitor situations without affecting non-criminal's life. It, also, allows pin pointing to the criminal acts in real time. Another basic use is rescue and medical emergency. In this case, it enables saving more lives and reduce search time for lost and in danger people. One of these technologies is mobile communication. Mobile communication showed, in a very short period of time, explosion in use by almost all people in the globe. Positioning a mobile could be used by law enforcement to face: illegal immigration, drugs smuggling, terrorist's caves, and boarder security. Also, it could be used by recue staff to locate people in danger.

Mobile positioning attracted MS manufacturer, service providers, and researchers [19], [20]. There are two basic approaches to get the position. The first approach depends on the existence of global position networks receivers within the MS set to determine its position and send it over the provider network. This approach requires users enabling and additive cost to have the receiver. The second approach is based on the power reports done and sent by the MS to enable mobility and to have stable communication. The first approach will not be suitable for the cases mentioned above since there is no control over the other side, MS. The second approach based on data already sent over the network which is power measurements and timing advance reports. Such parameters form the base on which the communication is well established [4][5].

Positioning process, in second case, comprises three main steps, distance to cells estimation, computing probable positions, and position refinement. Looking at network information and data interchanged over the network, the position related data are summarized in the following: -

1) Cell Identification or Cell ID: each cell has a specific ID known worldwide and associated with its position (Lat, Lon. , Alt). This data is known and published over the internet [6].
2) Received Signal Strength (RSS): The RSS is a measure of the magnitude of the power of the signal that the MS sees at its terminals. The RSS can give a figure of distance from the cells [7][8][9][10].
3) Timing Advance (TA): TA is a parameter sent to the mobile by its serving BS for timing synchronization purposes. TA is calculated using the round-trip delay time. The TA value is between 0 and 63 with resolution of approximately 3.69 microseconds. The corresponding distance resolution is 550 meters.
4) Angle of Arrival (AoA): This method uses antenna array measurements to calculate the direction of the signal coming from the active MS. Multiple antennas or moving antenna at speed much greater than the target speed could be used to locate the target MS. Also, AoA together with received signal strength could be used to estimate MS position.

5) Other parameters that are in the hand of the providers such as: Time of Arrival (ToA), time Difference of Arrival (TDoA) could be used to enhance and ease the positioning process.

The aim of this study is to provide passive positioning approach that doesn't depend on neither the network operators nor MS user. In section II: an overview of GSM network, network protocols, measurement report format, and challenges. In section III, radio wave propagation models described. In section IV, the proposed methods for BTS transmitted power estimation presented.  In section V, proposed positioning algorithms explained. Finally, in section VI, comparative experiments are shown between different algorithms.

## II. GSM networks

GSM is a standard developed by the ETSI to setup protocols and architectures for the second-generation digital cellular networks. The GSM network main elements are the mobile station (MS), the base transceiver station (BTS), the base station controller (BSC), the mobile switching center (MSC), the home location register (HLR), the visitor location register (VLR), the equipment identity register (EIR). Together, they form a public land mobile network (PLMN). The MS consists of the Mobile Equipment (ME) and the subscriber identity module (SIM) card. The ME is responsible for transmission and reception of voice and data, and the SIM represents the storage for the user identity as well as authentication and encryption parameters [4][5].

GSM air interface is a combination of Frequency Division Multiple Accesses (FDMA) and time division multiple access (TDMA), which results in a two-dimensional channel structure. TDMA frame consists of eight timeslots. Each user is assigned a time slot (TS) on a specific carrier frequency which is called physical channel. A group of 26 traffic TDMA frames is called a traffic multiframe and a group of 51 signaling TDMA frames is called signaling multiframe. Super-frame consists of 51 traffic multiframes or 26 signaling multi-frames. The hyperframe consists of 2048 superframe. In a GSM system, every TDMA frame is assigned a fixed number, which repeats itself in a time period of 3 hours, 28 minutes, 53 seconds, and 760 milliseconds. This duration is the hyperframe duration. The frame hierarchy is used for synchronization between BTS and MS, channel mapping, and ciphering.  Figure (2) shows the logical channels of the GSM network. There are technical challenges in wireless GSM communications like multipath propagation, synchronization, traffic variations over time, mobile mobility at high speed, volume of the traffic, variances in mobile sets, and uplink power limitation.

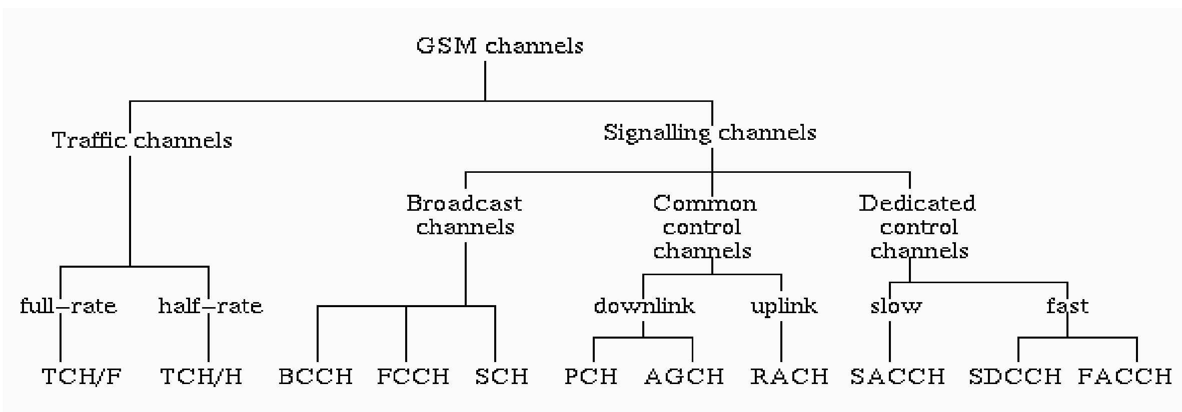


**Figure (2) The GSM logical channels.**

## III. Propagation models and cell power estimation

Modeling the radio channel has historically been one of the most difficult parts of mobile system design, and is typically done in statistical fashion, based on measurements made specifically for an intended communication system or spectrum allocation [7], [10], [21]. Propagation models or path loss models are models that predict the mean signal strength for an arbitrary transmitter-receiver separation distance. They are useful in estimating the radio coverage area of a transmitter and are called large scale propagation models, since they characterize signal strength over large transmitter-separation distances. Propagation models are either deterministic or empirical. Deterministic models based on the theory of electromagnetic wave propagation. These models require great details of the environment. Empirical models are based on extensive experimental data and statistical analysis which enables computing the received signal level in a given propagation medium. Among numerous propagation models, the following are the most popular ones, providing the foundation of today's land-mobile communication services: Okumura-Hata model and Walfisch Ikegami models [9], [10], [13].

***Okumura Hata with the extension COST 231***

The Okumura-Hata model foundation is experimental data collected from various urban environments having approximately 16 % high-rise buildings [9], [10]. The general path loss formula of the model is given by:

$$L_p(dB) = C_o + C_1 + C_2 \log(f) - 13.82 \log(h_b) - a(h_m) + [44.9 - 6.55 \log(h_b)] \log(d) \quad \text{where,}$$

$L_p$ is Path loss in dB, $C_o = 0dB$ for Urban, $C_o = 3dB$ for Dense urban.

$C_1$= 69.55 for 150MHz ≤f≤ 1000MHz and 46.3 for 1500MHz ≤f≤ 2000MHz

$C_2$= 26.16 for 160MHz ≤f≤ 1000MHz and 33.9 for 1600MHz ≤f≤ 2000MHz

f is frequency in MHz, $h_b$ is effective height of the base station in meters.

$h_m$ is the mobile antenna height in meters, and $d$ is the distance between the base station and the mobile in kilometers.

$a(h_m) = [1.1 \log(f) - 0.7]h_m - [1.56 \log(f) - 0.8]$ for urban

$a(h_m) = 3.2[\log(11.75h_m)]^2 - 4.97$ for dense urban

$$L_p(Suburban) = L_p(Urban) - 2\left[\log\left(\frac{f}{28}\right)\right]^2 - 5.4$$

$$L_p(Rural) = L_p(Urban) - 4.78[\log(f)]^2 + 18.33\log(f) - 40.94$$

***COST 231 Walfisch Ikegami model***

The COST 231 Walfisch-Ikegami model or Walfisch-Ikegami model is useful for dense urban environments. The parameters, excess path loss from Walfisch-Bertoni model [11] and final building path loss from Ikegami Model [12] are combined in this model with a few empirical correction parameters. This model uses several urban parameters such as building density, average building height, street widths etc. Antenna height is generally lower than the average building height, so that the signals are guided along the street, simulating an Urban Canyon type environment. The parameters used in Cost 231 Walfisch- Ikegami model are denoted in figure (3).

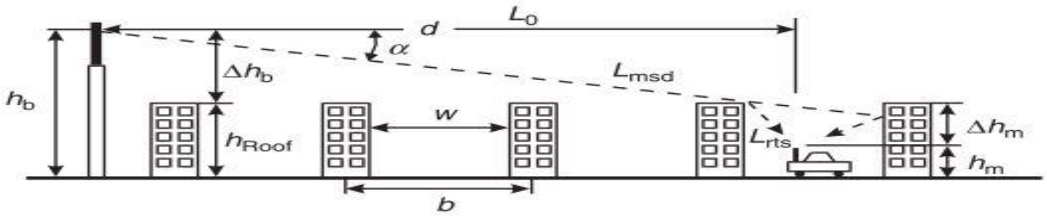


Figure (3) Parameters in COST 231 Walfisch-Ikegami model [16]

For Line of Sight (LOS) propagation, the path loss formula is given by:

$L_p(LOS) = 42.6 + 20\log(f) + 26\log(d)$ Where; $\boldsymbol{f}$ is the frequency in MHz, $\boldsymbol{d}$ the distance in km.

For Non-Line of Sight (NLOS) propagation, the path loss formula is given by,

$L_p(NLOS) = L_p(FreeSpace) + L(diff) + L(mult)$ Where;

$\boldsymbol{L_p}(\boldsymbol{FreeSpace}) = 32.5 + 20\log(f) + 20\log(d)$ and $\boldsymbol{f}$ and $\boldsymbol{d}$ are the frequency and distance respectively.

$\boldsymbol{L}(\boldsymbol{diff.})$ is the rooftop diffraction loss, $\boldsymbol{L}(\boldsymbol{mult})$ is the multiple diffraction loss due to surrounding buildings.

$$L(diff.) = -16.9 - 10\log(\Delta \mathrm{W}) + 10\log(f) + 20\log(\Delta h_m) + L(\phi)$$

The parameters in the above equation are: -

$\Delta\mathbf{W}$ is distance between the street mobile and the building, $\boldsymbol{h_m}$ is the mobile antenna height, $\Delta\boldsymbol{h_m} = \Delta\boldsymbol{h_{roof}} - \boldsymbol{h_m}$, and $\boldsymbol{L}(\boldsymbol{\phi})$ is the loss due to elevation angle.

The above parameters are constants after antenna is installation. Therefore, L(diff.) is constant.

Multiple diffraction and scattering components are characterized by the following equation:

$L(mult) = k_o + k_a + k_d log(d) + k_f log(f) - 9log(W)$ , where:-

$$\boldsymbol{k_o} = -\mathbf{18}\boldsymbol{log}(\mathbf{1} + \Delta\boldsymbol{h_b})$$

$$\boldsymbol{k_a} = \mathbf{54} - \mathbf{0.8}(\Delta\boldsymbol{h_b})$$

$$\boldsymbol{k_d} = \mathbf{18} - \mathbf{16}\left(\frac{\Delta\boldsymbol{h_b}}{\boldsymbol{h_{roof}}}\right)$$

$\boldsymbol{k_f} = -\mathbf{4} + \mathbf{0.7}[\left(\frac{f}{925}\right) - \mathbf{1}]$ for suburban.

$\boldsymbol{k_f} = -\mathbf{4} + \mathbf{1.5}[\left(\frac{f}{925}\right) - \mathbf{1}]$ for urban.

**W** is the street width, $\boldsymbol{h_b}$ is the base station antenna height, $\boldsymbol{h_{roof}}$ is the average height of surrounding small buildings ($h_{roof} < h_b$)

$$\Delta\boldsymbol{h_b} = \boldsymbol{h_b} - \boldsymbol{h_{roof}}$$

The widely used model is Okumura-Hata as it requires only area type. Also, it has a better path loss prediction performance than Walfisch-Ikegami in different areas. Walfisch-Ikegami model is mostly used in very dense urban environments. Also, Walfisch-Ikegami requires more details about the area than Okumura-Hata such as building density and average height as well as street widths [11-13]. Therefore, our choice was Okumura-Hata model. The models' overall requirements are: transmitted effective power, frequency, nature of the propagation area, tower height, and antenna types. The next section explains a simple method to estimate the tower effective transmitted power.

**Estimation of tower powers-based reference measurements**

Our methodology, in tower power estimation, uses real time measurements. These measurements are done using position-known simple GSM modem receiver equipped with global positioning device. This methodology provides quick adapting to the current towers transmission profile. Using the same device for all around towers reduces the overall positioning errors, giving that they are not all coming from the same direction.  This methodology eliminates the need for: transmitted power, antennas gains, EIRP (Effective Isotropic Radiated Power), from the providers. The power estimation, for a base station, is a three-step process: measurement, prediction, and optimum EIRP selection.

In the first step, the received signal strength is measured in different known locations. In the second step, the received signal strength is predicted in these locations through the propagation model using a range of EIRP values. In the third step, the EIRP value that result in minimum overall difference between predicted and measured signal strength is chosen as the EIRP of this base station. That is: Let N be the number of measurement points, and E be the summation of errors (the summation of differences between predicted and measured received power). E value is calculated from the following formula:

$$\mathrm{E} = \sum_{p=1}^{p=N} |Predicted\ Received\ Power - Measured\ Received\ Power|$$

The optimum EIRP selection step is simply selects the EIRP value that have the minimum E value. This process is done simultaneously for all the base stations in the area assuming fixed base station antenna height. Table (1) shows raw measurement position-power table for the base station with Global Cell Identity (GCI) 602-02-1302-274 in the format of MCC-MNC-LAC-CI. Where MCC, MNC, LAC, CI stands for Mobile Country Code, Mobile Network Code, Location Area Code, and Cell Identity, respectively. The cell location is at 30.0157410046544, 31.2311650223917 in Latitude, Longitude. The EIRP estimated for this cell is 34.7 dBm. Table (2) shows the predicted received power for the same cell after power estimation. It is obvious that the estimated power is not the optimum for each point, but for the summation of overall errors. The multipath and shadow fading have significant effect on the power estimation process error.

Table (1) 20 Measurements for cell 602-02-1302-274

| Location (Latitude, Longitude) | Measured |
|---|---|
| 30.0152621249032,31.2287180207324 | -79 |
| 30.0152631249032,31.2287180207324 | -78 |
| 30.0152631249032,31.2287180207324 | -77 |
| 30.0152681249030,31.2287270207322 | -76 |
| 30.0152851249026,31.2287450207320 | -76 |
| 30.0152981249023,31.2287570207318 | -71 |
| 30.0153051249020,31.2287680207316 | -73 |
| 30.0153051249020,31.2287680207316 | -75 |
| 30.0152981249022,31.2287630207317 | -74 |
| 30.0152901249023,31.2287580207318 | -74 |
| 30.0152901249022,31.2287600207317 | -74 |
| 30.0152971249022,31.2287580207318 | -74 |
| 30.0154121249013,31.2288070207311 | -72 |
| 30.0154121249045,31.2286720207331 | -72 |
| 30.0148051249088,31.2284430207364 | -77 |
| 30.0056601249441,31.2262980207662 | -81 |
| 30.0158701247634,31.2348620206385 | -73 |
| 30.0160381247797,31.2341570206493 | -70 |
| 30.0161401248024,31.2331680206645 | -64 |
| 30.0161201248162,31.2325630206738 | -66 |

Table (2) Measured and Predicted for cell 602-02-1302-274.

| Location (Latitude, Longitude) | Measured | Predicted |
|---|---|---|
| 30.0152631249032,31.2287180207324 | -77 | -73.340 |
| 30.0152681249030,31.2287270207322 | -76 | -73.278 |
| 30.0152851249026,31.2287450207320 | -76 | -73.142 |
| 30.0152981249023,31.2287570207318 | -71 | -73.049 |
| 30.0153051249020,31.2287680207316 | -73 | -72.970 |
| 30.0153051249020,31.2287680207316 | -75 | -72.970 |
| 30.0152981249022,31.2287630207317 | -74 | -73.012 |
| 30.0152901249023,31.2287580207318 | -74 | -73.055 |
| 30.0152901249022,31.2287600207317 | -74 | -73.042 |
| 30.0152971249022,31.2287580207318 | -74 | -73.044 |
| 30.0154121249013,31.2288070207311 | -72 | -72.581 |
| 30.0154121249045,31.2286720207331 | -72 | -73.425 |
| 30.0148051249088,31.2284430207364 | -77 | -75.738 |
| 30.0056601249441,31.2262980207662 | -81 | -98.356 |
| 30.0158701247634,31.2348620206385 | -73 | -79.377 |
| 30.0160381247797,31.2341570206493 | -70 | -76.181 |
| 30.0161401248024,31.2331680206645 | -64 | -70.248 |
| 30.0161201248162,31.2325630206738 | -66 | -64.988 |
| 30.0160581248309,31.2319180206836 | -64 | -56.297 |
| 30.0160181248503,31.2310670206966 | -56 | -42.006 |

## IV. <u>Position Estimation</u>

Positioning using power reports includes numerous sources of errors. Due to the propagation problems, nature of the area the path to cell antenna, the probability of being in or out of building, being in very close proximity to cell antenna, saturation of MS receivers, hand overs, Doppler effects, and so on. So, the aim is to get closer to real location as possible. Now, giving a power report and satisfying all the model requirements the models map these measurements to distances. The distances are from the cell position, specifically from center of its transmitting antenna. For simplicity, ignoring the earth curvature the target MS will be, ideally, somewhere in a point of the circumference of a cone center equal to cell position with height equal to the difference of the two heights (mobile and MS). One can easily infer the insignificance of the height difference as the lateral distance from the tower increases. Figure (4) shows real cases examples. Now, for a pair of towers gives us two circles. The two circles, ideally, either intersects in two points, luckily touch in a single point from inside or outside which points to a solution, or being apart.  The intersection points and the requirements of the intersections are equations (1 to 7) based on figure (4).

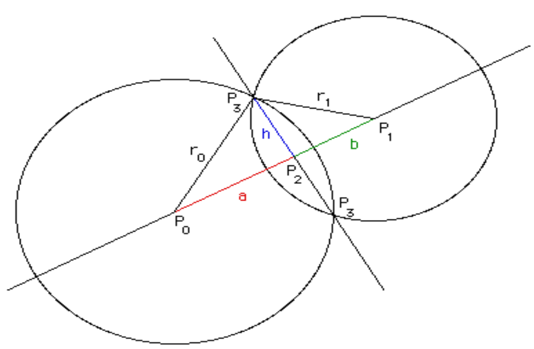


Figure (4) Two Circles Intersection Points

For two circles assuming, $P_0 = (x_0, y_0)$ , $P_1 = (x_1, y_1)$ and $d^2 = (x_0 - x_1)^2 + (y_0 - y_1)^2$, $r_0 \geq r_1$ the intersection condition is:-

$$d < (r_0 + r_1), and\ (d + r_1) > r_0 \qquad (1)$$

From figure case of intersection: -

$$a^2 + h^2 = r_0^2 \quad (2)$$

$$b^2 + h^2 = r_1^2 \quad (3)$$

Using d=a + b, then: $a = (r_0^2 - r_1^2 + d^2) / (2d)$ (4)

$$P_2 = P_0 + a\,(P_1 - P_0) / d \quad (5)$$

Consequently if $P_3 = (x_3, y_3)$, and $P_2 = (x_2, y_2)$, then the intersection points coordinates are:-

$$x_3 = x_2 \pm h\,(y_1 - y_0) / d \quad (6)$$

$$y_3 = y2 \mp h\,(x_1 - x_0) / d \quad (7)$$

Table (3) shows collected at position point at: 30.0161, 31.419. Table (4) shows cells to mobile distance after signal strength to distance mapping using propagation model. Figure (5) shows sample of estimated data mapped to circles. From the real collected data considering a pair of circles all cases are possible: intersecting, tangential outside and inside, apart inside and outside.

Table (3) Measurement report

| Cell ID | 13538 | 13537 | 13536 | 890 | 61183 | 1497 | 3348 |
|---|---|---|---|---|---|---|---|
| LAC | 31131 | 31131 | 31131 | 31131 | 31131 | 31131 | 31131 |
| RSS | -49 | -60 | -62 | -69 | -69 | -71 | -76 |

Table (4) Distance from cells to mobile

| Cell | 13538 | 13537 | 13536 | 890 | 61183 | 1497 | 3348 |
|---|---|---|---|---|---|---|---|
| Dist. MS | 91.288 | 168 | 233.78 | 638.9 | 810 | 159.6 | 1286.68 |

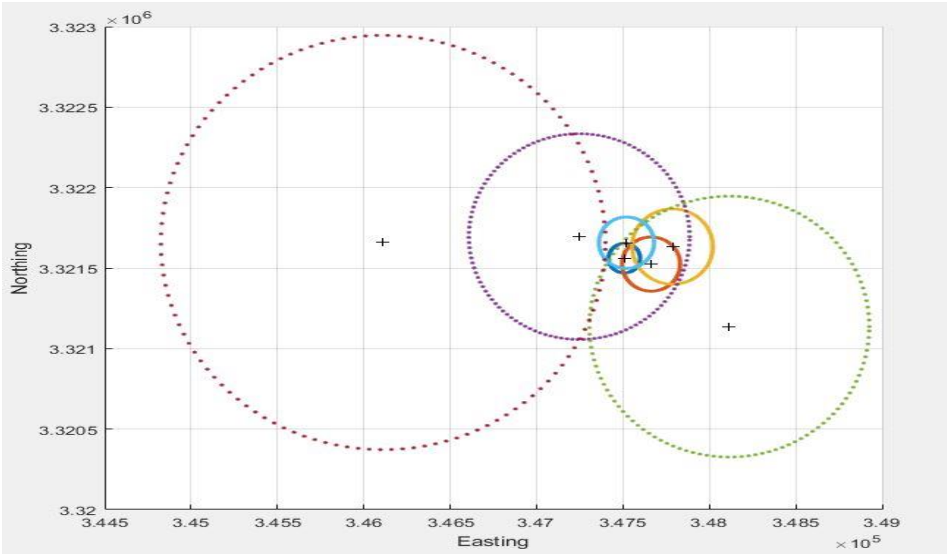


Figure (5) Circles intersection

Measurement reports are up to seven cells. These measurements could lead to between [0, 42] of intersection/tangential points. Handling nonintersecting cases are done as following: -

1) Nonintersecting outside, assuming that both signals are subject to same path loss, therefor coordinates $(x, y)$ of the potential position point is: -

$$x = x_0 + \alpha(x_1 - x_0), \tag{8}$$

$$y = y_0 + \alpha(y_1 - y_0) \tag{9}$$

$$, \ where\ \alpha = \frac{r_0}{r_0 + r_1}$$

2) Case nonintersecting inside, considering one's path is over estimated, and the other is under. So, one circle radius is increased, inner, and the other reduced in ratio of their radiuses. Then assuming $\mathrm{r}_0 > \mathrm{r}_1$,

$$\alpha_0 = \frac{r_0}{d}, \alpha_1 = \frac{1 + r_1}{d}, \tag{10}$$

$$\alpha = \frac{\alpha_1 + \alpha_0}{2}, \tag{11}$$

Then the potential point coordinate is:

$$x = x_0 + \alpha(x_1 - x_0), \tag{12}$$

$$y = y_0 + \alpha(y_1 - y_0) \tag{13}$$

3) Case tangential inside and outside, assuming $r_0 \geq r_1, \alpha = \frac{r_0}{d}$,

Then, the potential point coordinate is:

$$x = x_0 + \alpha(x_1 - x_0), \tag{14}$$

$$y = y_0 + \alpha(y_1 - y_0) \tag{15}$$

Following the special cases handling mentioned above leads to potential $n$ $p$ossible positions where 4<$n$<42. Out of these potential points a method is needed to find out a likely position point. In this research, five methods were considered. These methods namely: Least Squares, Mid, Mean, Voting, Vector quantization, and mean of $k$ out of $n$. Two of these methods, due to their performance, are used in case of having multiple reports, within the same message, at the same position point to find out a likely measurement report.

***1- Least Squares (regression).***

The Position calculation using Least Squares method requires at least 3 BSs with known locations. Let us assume that the distance between the MS, located at position X= [x, y] $^T$, and the possible position point is given by $d_i$, where:

$$d_i^2 = (x_i - x)^2 + (y_i - y)^2 \tag{16}$$

The objective of the method is to find out a position point that minimize the sum of the distances to all positions. That is a typical linear regression problem: -

$$HX = B, \tag{17}$$

Where

$$H = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \\ \vdots & \vdots \\ x_n - x_1 & y_n - y_1 \end{bmatrix},$$

and

$$B = \frac{1}{2}\begin{bmatrix} ({d_1}^2 - {d_2}^2) + ({x_2}^2 - {y_2}^2) - ({x_1}^2 + {y_1}^2) \\ ({d_1}^2 - {d_3}^2) + ({x_3}^2 - {y_3}^2) - ({x_1}^2 + {y_1}^2) \\ \vdots \\ ({d_1}^2 - {d_n}^2) + ({x_n}^2 - {y_n}^2) - ({x_1}^2 + {y_1}^2) \end{bmatrix}$$

Equation (17) solution is [17] [18]: -

$$X = (H^T H)^{-1} H^T B$$

***2- Mid.***

The target position point is mid of Latitude values and the mid of longitude values. That is, assume for $n$ points: $p_1, p_2, p_{3,\ldots\ldots,} p_n$. Where $(x_i, y_i)$ are the coordinates of the point $p_i$. $X = \{x_1, x_2, x_3, \ldots x_n\}$, . $Y = \{y, y_2, y_3, \ldots y_n\}$, $X^s{}_{,}Y^s$ is the sorted version of the two sets consequently? Then, the position point is $p = (x, y)$ where: -

$$from\ the\ set\ X^s\ x = x_m\ , where\ \ m = \frac{n+1}{2} case\ n\ odd\ , x_m\ = (x_{m1}\ + x_{m2}\ )/2\ , where\ m1 = n/2, m2 =\ n/2 + 1.\ simiarly, \quad y = y_m\ , from\ Y^s$$

**3-** ***Mean:*** The position point is $P = \frac{p_1+p_2+\cdots+p_n}{n}$

***4- Voting method.***

Consider each point as a probable position and let neighbor vote for that election. The point with greatest vote is considered the position. The metric used for determining neighboring taken from the standard deviation of inter-points distances. To make it clearer, consider $S = \{\ l_{12}, l_{13}, \ldots\ l_{1n}, l_{23}, l_{24}, \ldots\ldots, l_{n-1n}\}$ the set of distances between the given $n$ points. $l_{ij}$ is the distance between $p_i, p_j$and is the same for $p_j, p_i$ as the distance operator is cumulative. The vote $v_i$ for point $p_i$ is count of $l_{ij}\ or\ l_{ji}$ from set $S\ \forall j\ \in [1, n]\ and\ l_{ij}\ or\ l_{ji} < \frac{\sigma}{\alpha}\ ,\ constant\ \alpha > 1$ , $\sigma$ is the set S standard deviation.

***5- Vector quantization.***

Vector quantization, VQ, is a methodology that is being in use for a long time mainly in: communication, compression and classification. VQ algorithms select a set of representatives from the training set according to some proximity metric. In our case, the set of vectors is the intersection points. The algorithm clusters data and compute a representative for each cluster.

One of these representatives will be the target position point. Specification of target representative could be the set of least variances or the set the highest cardinality. The latter is chosen as the data is subject to significant noise. The VQ methodology used is the well-known LBG [23-25] and Euclidian distance as a proximity metric. As LBG requirements some initial representatives for the C clusters has to be given. The choices are made, in our case, uniform random choices. The LBG is iterative two steps algorithm: decomposition and re-computation. In decomposition steps, the points are distributed over the set of representatives according to the proximity metric. In the re-compute step, a new cluster representative of the set is computed to minimize the sum of the errors. To make it clearer: -

Assume the set of code book/ representatives at iteration i is the set $C^i = \{c_1^i, c_2^i, c_3^i, \ldots\ldots, c_M^i\}$ then the cluster of $c_j^i$ *is set* $S_j^i = \{x / \|x - c_j^i\| < \|x - c_k^i\| \forall c_k^i \in C^i, k \neq j$ , The $c_j^{i+1}$ = mean of set $S_j^i$ out of iteration $i$

**6- *Mean of k out of n***

The mean of k out of n, k<=n, computes the mean of the n candidate positions, iteration 0 the whole set of points is used. Then, drop the farthest point to the mean. The process continues until the set cardinality is k. The target position is the mean of remaining k. To be specific: let us assume for an iteration $i$ set of points are, $P^i = \{p_1^i, p_2^i, p_3^i, \ldots\ldots, p_M^i\}$ ***and*** $m^i = 1/M \sum_{j=1}^{M} p_j^i$

Then, either $M = K$ *then* $m^i$ *is the posion* or a new iteration is considered with set of points $P^{i+1} = \{p_j^{i+1}, j = 1, to, M-1\} = P^i - \{p_x^i / \|p_x^i - m^i\| > \|p_k^i - m^i\|, \forall k, x \neq k\}$ , eased to equality if inequality doesn't exist.

## V. <u>Implementation and Comparative Study</u>

Many field tests are done in different areas of different natures on the positioning methods discussed above. We will consider three experiments out of the tests done. In all evaluation experimentations a GPS used as a reference to the points calculated.

## V.1 Experiments

***Experiment 1:***

The first experiment was in a crowded city area within the school campus, in Cairo. The steps mentioned before is followed and 10 points far apart are chosen that includes the terminals as well as some random middle points. The measurement reports collected at the first two locations are in table 5, 6.

Table (5) First Point Measurement Report

| Location | Latitude | Longitude |
|---|---|---|
| | 30.066217 | 31.27927 |
| Cell ID | LAC | RSS (dbm) |
| 16646 | 1293 | -46 |
| 16645 | 1293 | -55 |
| 21156 | 1293 | -69 |
| 19893 | 1293 | -69 |
| 651 | 1293 | -70 |
| 16647 | 1293 | -70 |
| 653 | 1293 | -71 |

Table (6) Second Point Measurement Report

| Location | Latitude | Longitude |
|---|---|---|
| | 30.064533 | 31.280178 |
| Cell ID | LAC | RSS (dbm) |
| 16646 | 1293 | -54 |
| 21156 | 1293 | -63 |
| 16645 | 1293 | -67 |
| 651 | 1293 | -69 |
| 21155 | 1293 | -72 |
| 655 | 1293 | -72 |
| 653 | 1293 | -76 |

The corresponding circles, the actual and calculated position per method for the first point are shown in figures 6, 7, 8, 9, 10, 11. The summary of the second point results is in table 7.

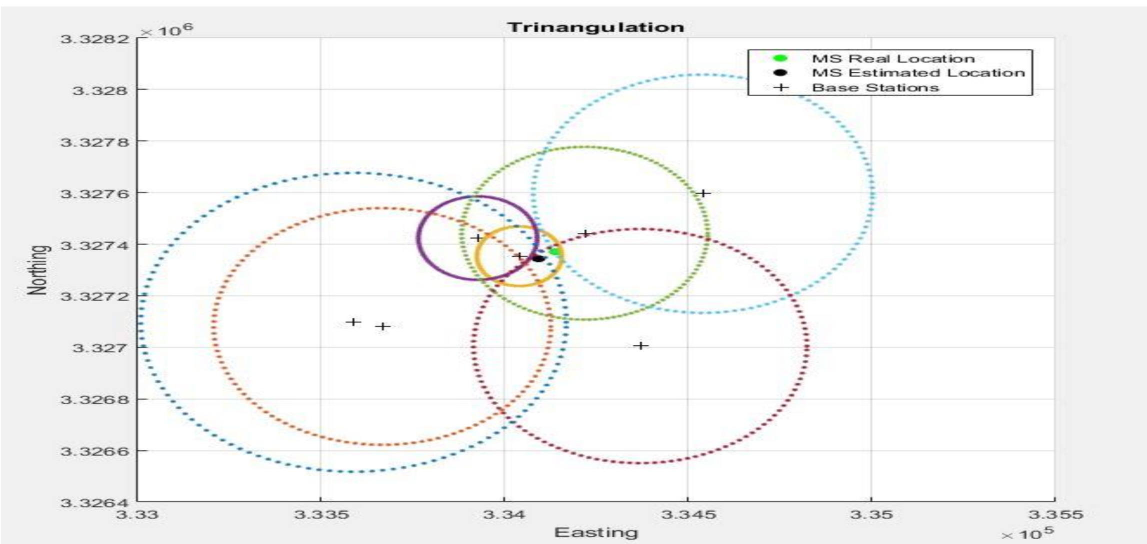

Figure 6 Least Squares is 30.066, 31.2788 error 53.5m

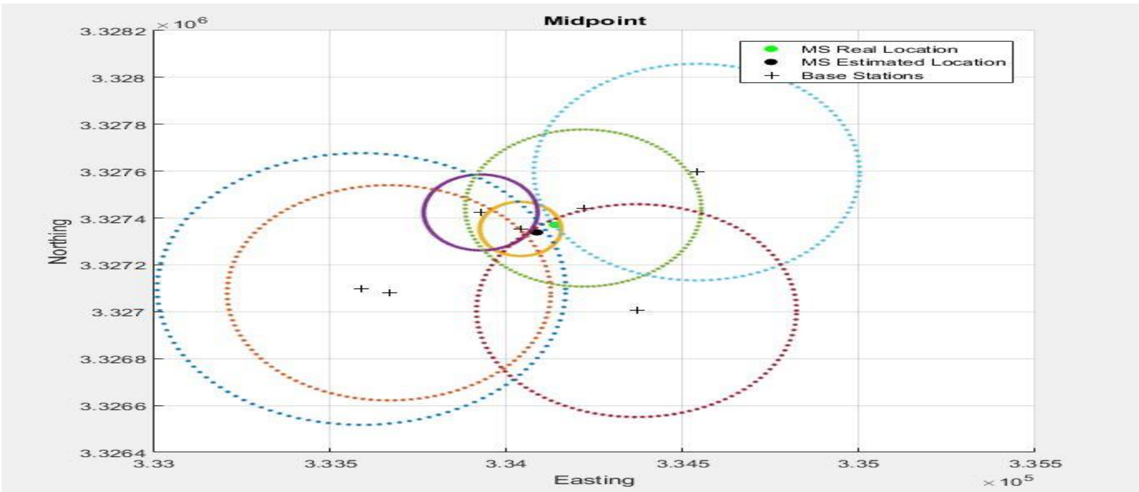


Figure 7 Midpoint is 30.0659, 31.2787 error 61m

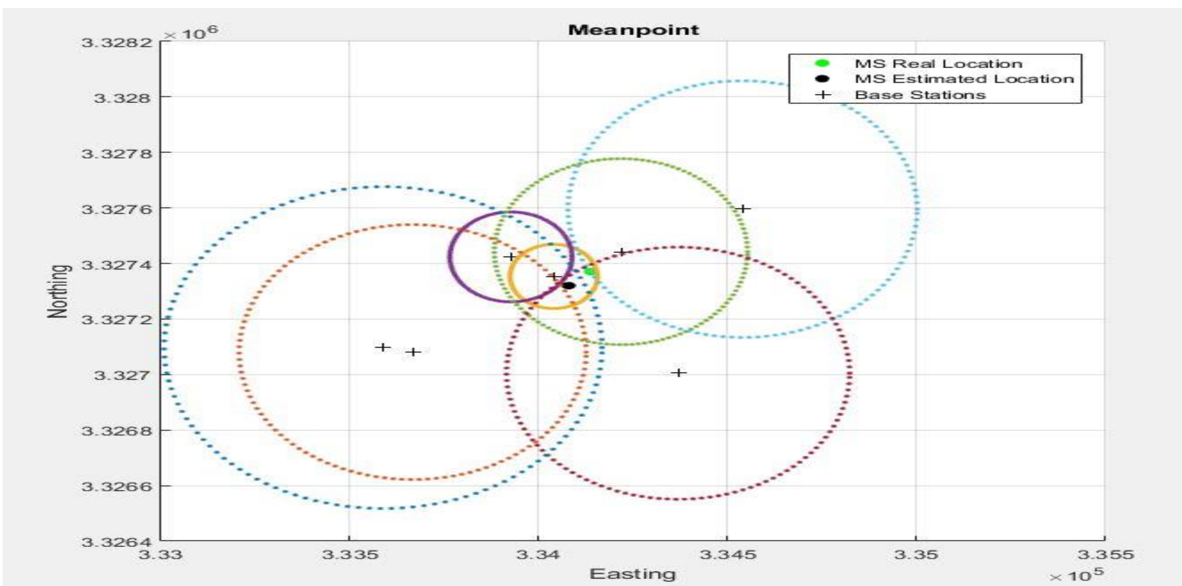


Figure 8 Mean point is 30.0658, 31.2787 error 77m

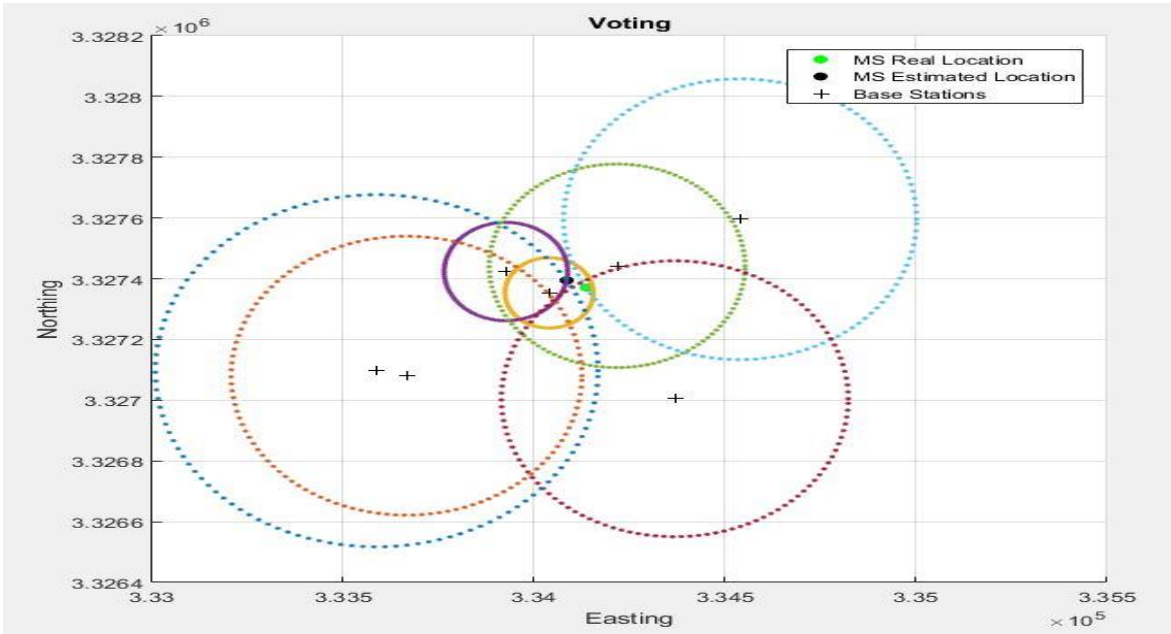


Figure 9 Voting is 30.0664, 31.2787 error 57m

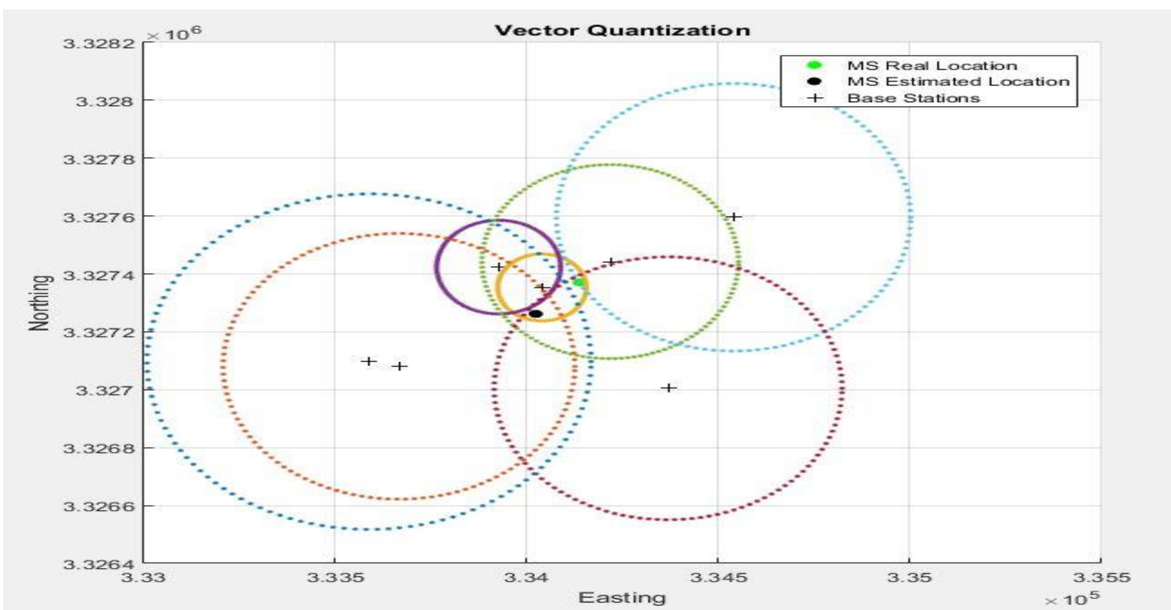


Figure 10, Vector Quantization is 30.0652, 31.2781 error 157m

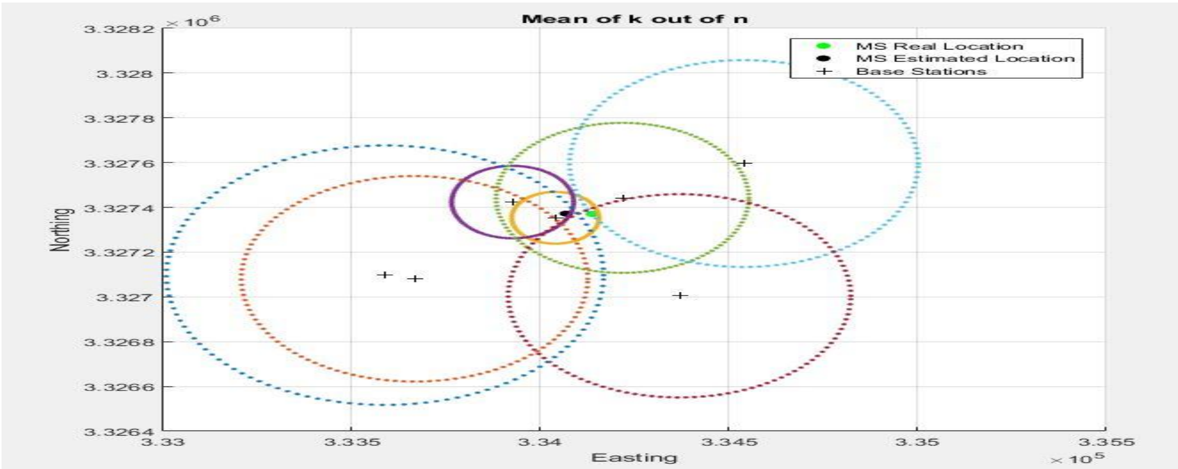


Figure 11, Mean of k out of n is 30.0662, 31.2786 error 68m.

Table (7) Summery of the second point.

| Method | Error in meters |
|---|---|
| Least Squares | 111 |
| Midpoint | 69 |
| Mean | 123 |
| Voting | 62 |
| Vector Quantization | 50 |
| Mean of k out of n | 86 |

To present the effect of fast fading on distance estimation, 10 measurement reports are taken in each point. Due to the fast fading or multipath fading phenomenon, the received signal strength from the same base station at different times is not the same. This leaded to different distances estimated for the same point of measurement from the same base station. The overall mean distance and standard deviation per cell per point are in table 8, 9 for the first two points.

Table (8): Measurements reports for first point.

| Cell (Cell ID, LAC) | Mean Distance (m) | Standard Deviation (m) |
|---|---|---|
| 16646, 1293 | 366.8 | 171.6 |
| 16645, 1293 | 364.5 | 169.8 |
| 21156, 1293 | 351.1 | 164.1 |
| 19893, 1293 | 363.1 | 186.8 |
| 651, 1293 | 376.3 | 186.4 |
| 16647, 1293 | 372.1 | 191.6 |
| 653, 1293 | 353.7 | 171.7 |

Table (9): Measurements reports for second point.

| Cell (Cell ID, LAC) | Mean Distance (m) | Standard Deviation (m) |
|---|---|---|
| 16646, 1293 | 345.2 | 148.4 |
| 21156, 1293 | 343.6 | 142.9 |
| 16645, 1293 | 345.6 | 132.9 |

| | | |
|---|---|---|
| 651, 1293 | 329.5 | 133 |
| 21155, 1293 | 341.7 | 140.9 |

| | | |
|---|---|---|
| 655, 1293 | 339.17 | 149 |
| 653, 1293 | 359.1 | 152.9 |

Overall the 10 points of testing of the first experiment. The best, worst, and mean accuracy for each method in meters are in table 10.

Table (10) overall ten points testing

| Method | Best | Worst | Mean |
|---|---|---|---|
| Least Squares | 49.1 | 272.4 | 135.9 |
| Midpoint | 43 | 129.1 | 81.8 |
| Mean | 21.9 | 158.9 | 107.1 |
| Voting | 42.6 | 158.2 | 97.5 |
| VQ | 50.8 | 289 | 127.6 |
| Mean of k out of n | 51 | 138.2 | 88.18 |

***Experiment 2:***

In the second experiment a moving car was used in short trip in the Fifth Settlement, Cairo, Egypt. Through the trip power reports collected every 10 seconds and associated with GPS position. Overall, the number of power reports was 373. The set of best (over all methods), where error < 50m, made 81 points with percent 21.7%. Figures 13, 14 depict one of these cases using voting and k out of n methods.

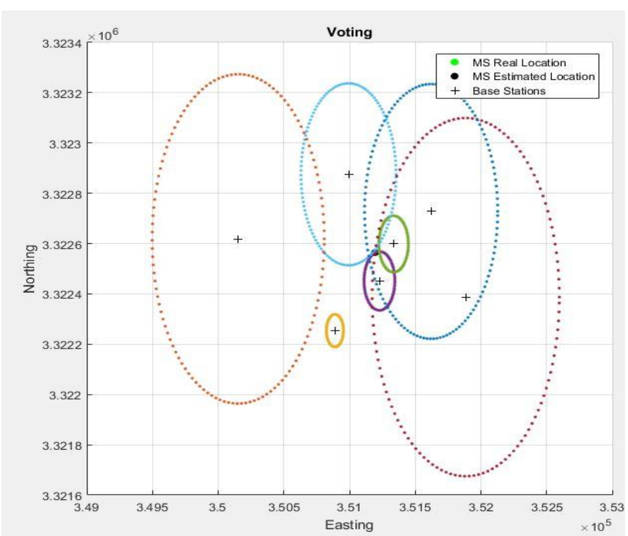


Figure 13: case 30.0251, 31.4568 voting error 3 m

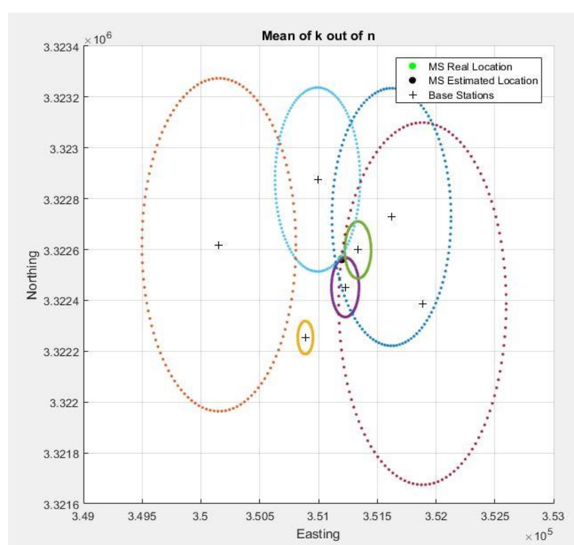

Figure 14: case 30.0251, 31.4568 k out of n
error 7 m

***Experiment 3:***

Here, like the second experiment, a moving car was used in short trip in El-Manial, Cairo, Egypt. Through the trip power reports are collected every 10 seconds and associated with GPS position. Overall the number of power reports was 300. Table (11) contains the statistics for different algorithms.

Table (11) experiment 3 statistics.

| Error/Method | Least Squares | Midpoint | Mean | Voting | Vector Quantization | Mean of k out of n | Best of all Methods |
|---|---|---|---|---|---|---|---|
| Below 50m | 7% | 14% | 10% | 13% | 12% | 10% | 34% |
| 50-150m | 29% | 50% | 36% | 45% | 38% | 50% | 45% |
| 150-250m | 15% | 21% | 22% | 23% | 24% | 25% | 14% |
| 250-350m | 14% | 12% | 19% | 11% | 15% | 11% | 7% |
| 350-500m | 15% | 3% | 9% | 7% | 7% | 4% | 0% |
| Above 500m | 20% | 0% | 4% | 1% | 4% | 0% | 0% |

***Experiment 4:***

In the Fourth experiment, combined with previous collected data, trips are made in different areas with different that made overall 2000 points. Table (12) summarizes the experiment.

Table (12) experiment 4 statistics.

| Error/Method | Least Squares | Midpoint | Mean | Voting | Vector Quantization | Mean of k out of n | Best of all Methods |
|---|---|---|---|---|---|---|---|
| Below 50m | 4% | 10% | 8% | 8% | 8% | 10% | 28% |
| 50-150m | 17% | 43% | 36% | 34% | 31% | 39% | 44% |
| 150-250m | 15% | 23% | 24% | 25% | 23% | 25% | 17% |
| 250-350m | 14% | 12% | 15% | 13% | 15% | 13% | 7% |
| 350-500m | 15% | 7% | 11% | 12% | 11% | 8% | 3% |
| Above 500m | 35% | 5% | 6% | 8% | 12% | 5% | 1% |

Overall, the mid-point of intersection points method came in the first place as it is more robust against distance errors, cell’s locations errors, and errors caused by fading. The mean of closest n (mean of k out of n) came in the second place. Mean, voting, and vector quantization came after, respectively. The least squares method comes in the last place as it showed the worst accuracy results compared to the other four methods as it is very sensitive to error in distance calculations as well as error in cells locations. In best cases, where distances are well estimated and cell locations errors was minimum, mean of k out of n and voting algorithms shows best accuracies.

***Experiment 5:***

In this experiment we study the enhancement in position as the time goes, assuming fixed position or slow-moving targets. The experiment was taken in metropolitan area, school campus, for several points. The data collected every 10 seconds for a period of 200 seconds. The output of an average performance point is in Table (13).  From the table we can easily notice that the mean of k out n and voting are stable in measurements compared to others.  Table (14) contains methods under study error statistics. The last column is the error in average location. The improvements in the errors come from the fact that the estimated location is not biased that is of type closed to white noise. There is an additional note on the latest table for the standard deviation in VQ method that comes from the fact that we use codebook size of two. Therefore, it outperforms others when none of cells suffers large biased error and vice versa.

Table (13) Point: 30.063562, 31.279018

| Time | Least Squares | Mid | Mean | Voting | Vector Quantization | Mean of k out of n |
|---|---|---|---|---|---|---|
| 10 s | 135 | 79 | 109 | 102 | 162 | 79 |
| 20 s | 84 | 83 | 53 | 6 | 144 | 54 |
| 30 s | 101 | 104 | 131 | 90 | 221 | 37 |
| 40 s | 102 | 57 | 95 | 81 | 34 | 64 |
| 50 s | 132 | 104 | 107 | 65 | 15 | 100 |
| 60 s | 115 | 103 | 118 | 61 | 63 | 75 |
| 70 s | 100 | 80 | 112 | 40 | 113 | 74 |
| 80 s | 111 | 75 | 113 | 61 | 97 | 63 |
| 90 s | 123 | 44 | 102 | 17 | 32 | 65 |
| 100 s | 241 | 58 | 69 | 40 | 32 | 69 |
| 110 s | 105 | 39 | 106 | 82 | 130 | 71 |
| 120 s | 186 | 75 | 88 | 30 | 199 | 50 |
| 130 s | 250 | 102 | 83 | 72 | 14 | 86 |
| 140 s | 120 | 56 | 112 | 30 | 70 | 68 |
| 150 s | 111 | 36 | 67 | 21 | 57 | 55 |
| 160 s | 107 | 44 | 70 | 13 | 34 | 51 |
| 170 s | 134 | 2 | 60 | 40 | 42 | 27 |
| 180 s | 118 | 61 | 90 | 18 | 8 | 76 |
| 190 s | 135 | 59 | 73 | 23 | 56 | 53 |
| 200 s | 131 | 46 | 59 | 26 | 49 | 50 |

Table (14) Overall methods statistics

| Method | Mean of the error | STD | Error in the Mean |
|---|---|---|---|
| Least Squares | 132 | 28 | 85 |
| Midpoint | 66 | 30 | 60.8 |
| Mean | 90 | 22 | 70 |
| Voting | 45 | 20 | 18 |
| VQ | 77.8 | 60 | 70.6 |
| Mean of k out of n | 49 | 16.2 | 16.8 |

## V.2 Comparative Study:

There are many researches, recently, in GSM positioning as it exists within the intersection of interests of Mobile service providers, law enforcements, rescue, position aware mobile application and others. The researcher's methods could be coarsely divided into two main categories power map based [26] [29] and to-cell distance based [29] [26]. Power map-based methods there are reports of 5m errors, indoor navigation as well as floor and within floor localization. Our study to such methods that is true if and only if: a) perfect setup measurements and dense map b) the same device used in measurements d) Dense measurements distribution to the target area e) non metropolitan areas. The accuracy comes basically from the fact that the problem became re-measurement in the same place or an interpolation from close neighboring points. That is the problem, in this case, is not propagation problem with all former mentioned complexities. In our study, the power map interpolation doesn't come to such accuracy or even close within the metropolitan's areas. The to-cell distance-based methods which are subject of comparison with ours reports are far different from the first group. In [27] five methods were understudy, from Figure 7 of that study, the cumulative distribution function is 0.2 at error 150m case of rural areas. In [28] the error 200m represents less than 0.2 cumulative probabilities. Tables 11 through 14 performance measurements competes with all in class proposed methods.

## VI. Conclusion

The proposed positioning methodologies contain two major steps: estimation of the cells transmission parameters needed and using this data together with online cell database and network transmitted RSS to estimate target location. In the first step, GSM modem equipped with GPS receiver is used to collect data about the working area cells. Then, the collected data with cells location database is used to estimate cell transmission power based on the inverse of the propagation models and least square method. The propagation models used in the study are Okumura Hata with the extension of it COST 231 and COST 231 and Walfisch Ikegami. In the second step, the RSS reports received during the regular communication process or enforced communication processed in two stages to get the most probable target location. The first stage is applied to the RSS burst to estimate the close to real values since they vary duties numerous factors. In this stage, the most effective two methods found to be voting and k out m methods. In the second stage, distances to cells are estimated based on the propagation models that lead to many candidates of target location. Six methods are used to infer the most likely target position. A simple tool was engineered for experimentation. The experimentations with tool were done in campus, around the campus and within different suburbs of Cairo. The results of study showed that mean of k out of n has better performance as contains two qualities, outlier removal and averaging that has bias cancelation effect. The Mid and Least Square are considered the worst as they have not any of the former mentioned qualities. The rest of the methods constitute an intermediate case as they include one of the qualities. The realistic testing of the proposed methods, in general, comparable to or outperform the same class to our knowledge.

# References


[1]- Nishith D. Tripathi, Jeffrey H. Reed, “Cellular Communications, A Comprehensive and Practical Guide”, John Wiley & Sons, 2015.

[2]- Balaram Singh, Soumya Pallai, Susil Kumar Rath, “A Survey of Cellular Positioning Techniques in GSM Networks”, Joint Venture College Computer Education, B.B.Mahavidyalaya, Utkal University, India, 2012.

[3]- Gunnar Heine, “GSM Networks, Protocols, Terminology, and Implementation”, Artech House, 1999.

[4]- ETSI GSM 05.08 (1996). Digital cellular telecommunications system (Phase 2+); “Radio subsystem link control”. Version 5.1.0 July 1996.

[5]- ETSI GSM 05.10 (1996). Digital cellular telecommunications system (Phase 2+); “Radio subsystem synchronization”. Version 5.0.0 Release May 1996.

[6]- http://cellidfinder.com/

[7]- Theodore S. Rappaport, “Wireless Communications Principles and Practice”. 2nd ed., Pearson Education Inc 2002.

[8]- Govind Sati, Sonika Singh, “A review on outdoor propagation models in radio communication”. International Journal of Computer Engineering & Science, Vol 4, Issue 2, pp. 64-68, March 2014.

[9]- European Cooperative in the Field of Science and Technical Research EURO-COST 231, “Urban transmission loss models for mobile radio in the 900 and 1800 MHz bands,” Revision 2, The Hague, Sept. 1991.

[10]- M. Hata, “Empirical formula for propagation loss in land mobile radio services,” IEEE Trans. Vehicle Technology., Vol VT-29, No. 3, pp. 317–325, Aug. 1980.

[11]- J. Walfisch and H.L. Bertoni, “A Theoretical model of UHF propagation in urban environments,” IEEE Trans. Antennas Propagation, vol.36, 1988, pp.1788-1796

[12]- F. Ikegami, T. Takeuchi, and S. Yoshida, “Theoretical prediction of mean field strength for Urban Mobile Radio”, IEEE Trans. Antennas Propagation, Vol.39, No.3, 1991.

[13]- COST Final Report, http://www.lx.it.pt/cost231/

[14]- ETSI GSM 03.02 (1996). Digital cellular telecommunications system (Phase 2+);“Network architecture”. Version 5.1.0 Release May 1996.

[15]- ETSI GSM 05.02 (1996). Digital cellular telecommunications system (Phase 2+); “Multiplexing and multiple access on the radio path”. Version 5.0.0 Release May 1996.

[16]- Andreas F. Molisch, “Wireless Communications”, 2nd edition, 2011, John Wiley & Sons.

[17]- João Figueiras, Simone Frattas, “Mobile Positioning and Tracking from Conventional to Cooperative Techniques.”, A John Wiley & Sons Ltd, 2010.

[18]- Mathias Pelka, “Position Calculation with Least Squares based on Distance Measurements”, Lübeck University of Applied Sciences, Technical Report, 2015.

[19]- Eneh Joy Nnenna, and Orah Harris Onyekachi, “Mobile Positioning Techniques in GSM Cellular Networks: A Comparative Analysis ”, IJCTEE, Vol 2, Issue 6, Dec 2012.

[20]- Peter Brida, Peter Cepel, Jan Duha, “Geometric Algorithm for Received Signal Strength Base Mobile Positioning”, RADIOENGINEERING, Vol.14, NO.2, June 2005.

[21]- Syed A.Ahson and Mohammad Ilyas, “Location Based Services Handbook Applications, Technologies, and Security”, CRC Press, 2011.

[22]- John S. Seybold, “Introduction to RF Propagation”, John Wiley & Sons., 2005.

[23]- Yoseph Linde, Andres Buzo, Ropert M. Gray, “An Algorithm for Vector Quantizer Design”, IEEE Transactions on Communications, Vol. COM28, No.1, Jan 1980.

[24]- Abdelatief H. Abouali, “Object-based VQ for image compression”, Ain Shams Engineering Journal Vol.6 Issue1, March 2015.

[25]- Willam A. Porter, Abdelatief H. Abouali, “Vector Quantization for multiple classes”, Information Sciences Vol 105 Issue 1-4, March 1998.

[26]- Mohamed Ibrahim, Moustafa Youssef , “CellSense: An accurate Energy-Efficient GSM Positioning System, IEEE Transactions on Vehicular Technology, Volume: 61, Issue: 1, Jan. 2012.

[27]- Mohamed H. Abdel Meniem*, Ahmed M. Hamad**, Eman Shaaban ,”Fast and Accurate Practical Positioning Method using Enhanced-Lateration Technique and Adaptive Propagation Model in GSM Mode, IJCSI International Journal of Computer Science Issues”, Vol. 9, Issue 2, No 1, March 2012.

[28]- Huiyu Liu, Yunzhou Zhang, and others,”Mobile Localization Based on Received Signal Strength and Pearson's Correlation Coefficient”, International Journal of Distributed Sensor Networks ,Volume: 11 issue: 8, August 19, 2015.

[29]- Rafał Górak, Marcin Luckner, Michał Okulewicz and others,” Indoor Localization Based on GSM Signals: Multi-story Building Study”, Mobile Information Systems, Volume 2016