4.1 Chapter Overview

The accuracy of the first step via the utilization of an advanced CellID MS RSS estimation model. This model is based on instantaneous field measurements, and takes into account the complicated propagation mechanisms that limit the CDR base positioning accuracy in the offline CDRs analysis method. That there is close interconnection between the two stages, the initial BTS CDR Location estimate is necessary in the second step as an initialization parameter, in order to facilitate the accurate estimation of the cellular device in GSM/CDMA network increment data mining.

4.2 Network online CDR analysis

The first of step of GSM/CDMA network data mining accuracy in offline network but some time need to online network data analysis, now second step Cell ID Increment data to improve the accuracy of network navigation, this step demonstrate the incremental data mining, the association rules of data mining on transactional database is usually an offline process since it is costly to find the association rules in large databases [41], incremental updating techniques should be developed for maintenance of the discovered association rules to avoid redoing mining on the whole updated database [147][92]. A database may allow frequent or periodic updates and such updates may not only invalidate existing association rules but also activate new rules. Thus it is nontrivial to maintain such discovered rules in large or small databases [122].

Incremental CDR Data mining is the process of extracting knowledge structures from continuous, rapid data records. A data stream is an ordered sequence of instances which have limited computing and storage capabilities[289]. In many application of data increment mining, it can be read only once or small number of times (J Joseph 2011) [133]. Figure 4.1 show the association rule of datamining process.
In this research, network CDR incremental update with received signal strength. The propose algorithm for discovering the complete set of GSM/CDMA network frequent patterns in CDR (Particular time series databases), researcher discover the frequent patterns over the entire time series in contrast to applying a sliding window over a portion of the time series. The proposed approach has the ability to discover frequent patterns that contain gaps between patterns' items (CDR data field) with a maximum user-defined gap size. With the arrival of each new data item, the algorithm updates the existing mining results incrementally. Define a set of states for the patterns in the database depending on whether they are frequent or non-frequent. Based on the transitions among these states, the algorithm takes certain actions to update the existing frequent patterns.

![Diagram of association rule mining process]

Figure-4.1 Process of association rule for mining

Offline CDR analysis has been provide a suspect network navigation on previous time or past time movement of suspect or behavior, geo position of the communication, Online CellID CDR data is the three directions (any one) Alfa, Beta Gamma (α, β, γ), data changes from the
real time and network latitude and longitude change in real time, in other word to say that network CDR update in real time. CellID (alpha,beta,gama) it is provide the direction of suspect geographic movement. Frequency information provides the approximate distance of BTS to suspects. Increment CDR Data is to receive signal strength (Generally called Received Power) area of the cell ID (BTS), IMEI with IMSI number real time geographic position generate in CDRs (RSS) data. In a particular time lat& long geo position indicate the group communication increase rapidly update this case group increment clustering help to behavior of this group for example syndicate of cricket match batting , during the match particular lat & log position communication has increase that condition incremental algorithms sound support to indentify this position of group and incremental identify number CDR , self increment CDR indicate the relevant BTS suspect direction ,suspect distance in BTS ,suspect behavior in enter BTS communication ,more BTS communication behavior. The online increments data provide by company, if any user knows self incremental CDR then longing account online prepaid or postpaid, it is seem to real time CDR data, clustered information of the user, But investigation agencies know the suspects that condition special like connected investigation agency central server, and they know the increment user CDR data and investigation agency location estimate by frequency calculate methods.

In a received communication signal at the mobilephone device is normally define in dBm. Figure 4.2 show receives signal strength. Clearly two different service provider RSS is -75 dBm asu 19, -69 dBm asu 22, because cell phone receive the signal two different place cell ID. Localization using the received signal strength (RSS) of GSMA/CDMA (IEEE 802.11 standard) local area networks with a priori information of the coordinates of the routers access points is addressed.

The goal of this research is to navigate a user’s present receive strength track base region position. Clearly manages with signal strength fluctuation over time and allows control over system accuracy by means of specification of confidence intervals. Position based on received signal strength patterns is difficult. Scattering and reflection of electromagnetic waves can strongly distort the signals shielding severely attenuates the signal strength RSS is the knowledge about the track show in the figure 4.3 there is a strong correlation between the distance of a base station and its signal strength, which can be exploited by a suitable positioning system
Figure 4.2 - Measuring cellular device receive signal strength cell phone Micromax –A50, and HTC.
But some environment factor or variable factor such as walls, buildings, or mountains, cell signal fluctuate other variable (Day Hour/ Peak Load If the suspect’s movement or hide an area, near a cell tower, but still have weak cell signal strength or receive signal strength affected row material, building materials for example metal, concrete, and low-e glass reduce or block cell signal from entering into a building) effect the suspect position. In this case position is difficult and never finds the suspect position, now the CDRRSS effective positing methods are more effective and result is positive direction. Our research approach is absolutely based on signal patterns received from base stations.

4.3 Arbitrary Strength Unit (asu part)

Figure 4.2 shows Arbitrary Strength Unit (ASU) is an integer value relative to the received signal strength determined by the cellular device. It is possible to calculate the real signal strength measured in dBm (and thereby power in Watts) by a formula. There are different
formulas use to determine 2G and 3G networks, ASU is equal to the RSSI (received signal strength indicator).

\[
\text{dBm} = 2 \times \text{ASU} - 113
\]

Table 4.1 Network ASU RSS

<table>
<thead>
<tr>
<th>Sr.no.</th>
<th>Network Name</th>
<th>Measurement</th>
<th>Usable ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>GSM</td>
<td>-113 to -51 dBm</td>
<td>GSM -113dBm (0ASU) to -51dBm(31ASU)</td>
</tr>
<tr>
<td>2.</td>
<td>CDMA</td>
<td>-100 to -75 dBm</td>
<td>CDMA -100dBm(1ASU) to -75dBm(16ASU)</td>
</tr>
</tbody>
</table>

4.4 Online CDR with RSS

All cell phone devices working a mobile network code (MNC) is used in combination with a mobile country code (MCC) (also known as a "MCC / MNC ") to uniquely identify a mobile phone operator/carrier using the GSM/CDMA mobile networks [114] [appendices iv]. The MS (cell phone devices) is in idle mode the MS always knows the Cell-ID, the network knows only the LAI. The network becomes aware of the Cell-ID only when the MS (Mobile Station) switch into dedicated mode, namely when the channel is used to actually establish a call.

GSM/CDMA network cell identification (Cell-ID) navigation analysis, where the cell, which is cell phone is registered. This information is available in the network and at the cell phone. Cell-ID is then converted to a geographic position using knowledge of the operator's network, such as coverage database at the serving mobile location centre (SMLC).

\[
\text{CDR}_{\text{SMLC}}
\]  
(Where CDR SMLC serving mobile location center)

4.4.1 Cell size power density
Cell tower antennas transmit in the frequency range of 869 - 894 MHz (CDMA), 935 - 960 MHz (GSM900) and 1810 – 1880 MHz (GSM1800). Also, 3G has been deployed in a few cities, in which base station antenna transmits in the frequency range of 2110 – 2170 MHz. Cell phone operators divide a region in large number of cells, and each cell is divided into number of sectors. The base stations are normally configured to transmit different signals into each of these sectors. In general, there may be three sectors with equal angular coverage of 120 degrees in the horizontal direction as this is a convenient way to divide a hexagonal cell. Every cell phone tower transmit the frequency signals in power (watt), Power density of cell phone antenna

The accuracy level thus generally depends on the cell size, but may also depend on other factors such as the wireless network type (GSM or CDMA), cell type (Omni-directional or sectored), cell kind (macro cells large-scale cells - or pico cells - small-scale cells), and so on.

\[
\text{CS}_{\text{MS}} \tag{4.2}
\]

(Where Cell Size Mobile Station (Cellphone) )

Cell size Power density \( P_d \) at a distance \( R \) is given by

\[
P_d = \left( P_t \times G_t / 3.1428 \times R^2 \right) \text{ watt/m}^2 \tag{4.3}
\]

where, \( P_t \) = Transmitter power in Watts

\( G_t \) = Gain of transmitting antenna

\( R \) = Distance from the antenna in meters

For example is \( P_t = 20 \),

\( G_t = 17 \) and dB50
Now cell size is power density show in the table 4.2

Table 4.2 Power density at various distances from the transmitting tower

<table>
<thead>
<tr>
<th>Sr.No</th>
<th>Distance R (m)</th>
<th>Power density $P_d$ in W/m²</th>
<th>Power density $P_d$ in $\mu$W/m²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>79.6</td>
<td>79,600,000</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>8.84</td>
<td>8,840,000</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>3.18</td>
<td>3,180,000</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>0.796</td>
<td>796,000</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0.0318</td>
<td>31,800</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>0.008</td>
<td>7,960</td>
</tr>
<tr>
<td>7</td>
<td>500</td>
<td>0.00318</td>
<td>318</td>
</tr>
</tbody>
</table>

The power density values given in Table 4.2 are for a single carrier and a single operator. To measure the power at a distance $R$, an antenna is used to receive the power and a power meter is used to measure received power. Power Received $P_r$ by an antenna at a distance $R$ is given by:

$$P_r = P_t \times G_t \times G_r \times \text{Area (Alfa, beta, gamma)}$$

$$P_r = P_t \times G_t \times G_r \times \left(\frac{\lambda}{4 \times 3.142 \times R}\right)^2$$

Where

$\lambda$, Electromagnetic waves travel at the speed of light (299,792,458 meters per second) and their frequency and wavelength can be determined by the formulas:

$$F = \frac{C}{\lambda}$$

‘$F$’ is denoted frequency in cycles per second,
‘C’ is denoted light velocity 299,792,458 mtrs/seconds

\(\lambda\)  is denoted represents wavelengths in meters

Received power is directly proportional to the transmitted power, gain of transmitting and receiving antennas, and square of wavelength of the signal and it is inversely proportional to square of distance. Table 4.3 show the power

Table 4.3 show the power in dBm

<table>
<thead>
<tr>
<th>Power dBm</th>
<th>Power (dBW)</th>
<th>Power (watt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-100 dBm</td>
<td>-130 dBW</td>
<td>0.1 pW</td>
</tr>
<tr>
<td>-90 dBm</td>
<td>-120 dBW</td>
<td>1 pW</td>
</tr>
<tr>
<td>-80 dBm</td>
<td>-110 dBW</td>
<td>10 pW</td>
</tr>
<tr>
<td>-70 dBm</td>
<td>-100 dBW</td>
<td>100 pW</td>
</tr>
<tr>
<td>-60 dBm</td>
<td>-90 dBW</td>
<td>1 nW</td>
</tr>
<tr>
<td>-50 dBm</td>
<td>-80 dBW</td>
<td>10 nW</td>
</tr>
<tr>
<td>-40 dBm</td>
<td>-70 dBW</td>
<td>100 nW</td>
</tr>
<tr>
<td>-30 dBm</td>
<td>-60 dBW</td>
<td>1 W</td>
</tr>
<tr>
<td>-20 dBm</td>
<td>-50 dBW</td>
<td>10 W</td>
</tr>
<tr>
<td>-10 dBm</td>
<td>-40 dBW</td>
<td>100 W</td>
</tr>
<tr>
<td>-1 dBm</td>
<td>-31 dBW</td>
<td>794 W</td>
</tr>
<tr>
<td>0 dBm</td>
<td>-30 dBW</td>
<td>1.000 mW</td>
</tr>
<tr>
<td>1 dBm</td>
<td>-29 dBW</td>
<td>1.259 mW</td>
</tr>
<tr>
<td>10 dBm</td>
<td>-20 dBW</td>
<td>10 mW</td>
</tr>
<tr>
<td>20 dBm</td>
<td>-10 dBW</td>
<td>100 mW</td>
</tr>
<tr>
<td>30 dBm</td>
<td>0 dBW</td>
<td>1 W</td>
</tr>
<tr>
<td>40 dBm</td>
<td>10 dBW</td>
<td>10 W</td>
</tr>
<tr>
<td>50 dBm</td>
<td>20 dBW</td>
<td>100 W</td>
</tr>
<tr>
<td>60 dBm</td>
<td>30 dBW</td>
<td>1 kW</td>
</tr>
<tr>
<td>70 dBm</td>
<td>40 dBW</td>
<td>10 kW</td>
</tr>
<tr>
<td>80 dBm</td>
<td>50 dBW</td>
<td>100 kW</td>
</tr>
<tr>
<td>90 dBm</td>
<td>60 dBW</td>
<td>1 MW</td>
</tr>
<tr>
<td>100 dBm</td>
<td>70 dBW</td>
<td>10 MW</td>
</tr>
</tbody>
</table>

Incremental CDR problem is when CDR parameter is available and CDR all parameter update and CDR size is small but BTS cellID (alpha, Beta, Gama) change in period or change in small period latitude or longitude change than increment CDR parameter analysis and navigation
is difficult and receive power density is effected or weather effect RSS or some other variable effect

Let consider the diagram 4.4, $\text{CDR}_{\text{SMLC}}$ be a transactional database and itemset is \{Call party, Call receive party, Time Duration, IMEI, IMSI, CallType \}, Time is denoted by $T$ , Environment effect variable $E_{\text{RSS}}$

Given Parameter are

BTS=lat and long

BTS= Max CellID $\beta$

Use Education no. (4.2)

Figure 4.4 Increment $\text{CDR}_{\text{SMLC}}$
Cell size Power density $P_d$ at a distance R

Max $P_d$

When MS received RSS with ERSS

$RSS = \text{Max } P_d - \text{Max } P_r \times E_{RSS}$

Where $E_{RSS}$ calculate ASU port Experiment base

Increment algorithms apply to clustering CDR data, and navigate to GSM/CDMA network

Algorithms increment CDR 9

Step-1 Parameter (Pi)

Step-2 if parameter (call party)

{ Maximum Cell ID Maximum Call Party Maximum Received Party Minimum Call Party Minimum Cell ID Avg=CellID+ Call

$RSS = \text{Max } P_d - \text{Max } P_r \times E_{RSS}$

GIS MAP

}
if parameter (call Received party)
{
  Maximum Cell ID
  Maximum Received Party
  Minimum Received Party
  Minimum Cell ID
  Avg=CellID+ Received Party
  \[ \text{RSS} = \text{Max } P_d - \text{Max } P_r \times E_{\text{RSS}} \]
  GIS MAP
  Step-3
  IF Cell-ID (Time period)
  RSS Clustering
  GIS MAP
  Next or Stop

4.5 Latitude and Longitude GIS MAP calculation

Accuracy of increment algorithm to calculate Location on google map or bing map is required to BTS and MS latitude and longitude base distance calculation, Suppose TSP provide particular IMSI no CDRs and BTS latitude and longitude point now CDRs clustering analysis or other type of the investigation use latitude and longitude calculation formula and if available real time available RSS then produce real time GIS map. In use receive signal strength calculation this a variety of calculations for latitude/longitude points, with the formula and code fragments
for implementing them. All these algorithms are for calculations on the basis of a spherical earth (ignoring ellipsoidal effects) which is accurate enough for most purposes. The RSS depend on cell shape or cell environment dependent variable, Receive signals strength. This is based on the accuracy of the navigate method that gives different levels of accuracy and hence aims at different navigation sectors. This will include BTS and CDR based data, to give accuracies which relate to the size of the cells. A sectored cell uses a number of antennas in the cell, and these provide sector information. The larger the number of sectors, the more accurate the location of the device will be. When a cell phone device is registered to an omnidirectional cell, the location of the handset becomes exact location of the mast, that is, centre of the cell and the coverage area (cell size) of that mast becomes the accuracy level.