

09/158296

EUROPEAN COOPERATION
IN THE FIELD OF SCIENTIFIC
AND TECHNICAL RESEARCH

COST 259 TD(98) 043

April 99

EURO -COST

Source: Spain

Telefónica

Mobile Location using coverage information: Theoretical analysis and results

J. Jiménez, E. Alonso, H. Gonzalez, M. Romero, I. Berberana

Telefónica I+D. Emilio Vargas 6, 28043 Madrid, Spain.

jjimenez @tid.es

Abstract:

This contribution addresses the location problem using coverage information in a cellular network context, particularly GSM. As it is well known, mobile location usually requires more detailed information than just coverage, due to the high variability of this kind of information. However, recent advances in propagation prediction, the availability of new tools, etc. make it coverage information a good candidate for mobile location, at least for a first approximation.

This contribution makes a preliminary analysis of the possibilities of using propagation information for location purposes, studies its requirements and limitations. In particular, several algorithms are considered and the effect of different propagation conditions is analysed.

As can be expected, location using coverage information can only be undertaken when a good knowledge of the propagation environment is available. In most cases, an "ad-hoc" measurement campaign is required. Nevertheless, when good information is available, location errors smaller than 100-150 metres are obtained for 70% of cases.

Introduction

In a previous contribution to COST [1], the use of propagation information for "hot spot" location has been reported. The basic conclusion of the report is that locating hot spots using just coverage data was indeed possible and relatively small errors are obtained.

In this contribution, a more careful analysis of the possibilities of using coverage data for location information is undertaken. The objective is to analyse what is the influence on the location error of several factors: the most important being the error on the coverage information data, but it has also been analysed the number of measurements available, the mobile speed, the correlation between measurements, different location algorithms etc.

Location algorithms

As it is well known, one of the most interesting characteristics of GSM is the fact the MS reports back to the BS (MSC) the information of the receiving level of the six BS received with greater level. This is done to facilitate the hand-over of the communication in course when the MS is leaving the serving BS coverage area. This information is not usually recorded by the network, unless specifically instructed to do so via the O&M system. Even in those cases, most O&M systems can only record the measurement reports prior to the occurrence of a call event: (Call set up, hand-over.). For that reason, MS location has to be done with a relatively small amount of measurement reports – in the region of 10-15.

Several algorithms were suggested in [1] in order to find the MS location based on that reduced amount of information, those algorithms are summarised here for reference:

In mathematical terms the objective is to find a candidate position $r(x,y)$ for the mobile station taking into account the collection of reported measurement levels v_m to a number of bases (normally around 6). The algorithm compares those reported levels to a collection of planned or previously measured values v_p for every location of the area under study, usually of some square kilometres.

The estimated location of the MS is the position r that makes minimum the distance between the measured (reported) v_m and the planned v_p .

If there is only one set of measured v_m (corresponding to one location) the usual approach is to give as candidate position with the minimum square value of the distance between the reported and planned values.

$$f(r) = \sum_{i=1}^M (v_{im} - v_{ip})^2 \quad (1)$$

(being M the number of BS with reported and planned measurements)

Instead of the minimum square value, different metrics can be used: absolute value, minimax, etc.

If the number of reported measurements of a mobile is greater than one the location finding algorithm can be refined. Now v_m is a collection of k measurements to M base stations.

Several approaches can be used.

1.- The optimum approach would be using a new expression for $f(r)$ that would optimise in the whole set of combinations for the components of r .

$$f_1(r) = \sum_{\text{all possible paths}} \sum_{i=1}^M (v_{im} - v_{ip})^2$$

This approach implies first selecting all possible paths in an area of some square kilometres. The number grows exponentially with the number of measurements. Some reduction is obtained using the fact that the MS cannot move at too high speed, but even so the calculation burden is too high.

2.- A second possibility would be to use a derivative of a Viterbi-like algorithm. First a candidate starting point r is determined using (1). The following elements $r_j, j=2...k$, can be determined taking into account they cannot be separated more than the maximum mobile velocity (vel) multiplied by the measurement interval. This implies that (1) will be evaluated only for those values v_p inside a circle of radius $vel \cdot \Delta t$.

Since those measurements with low attenuation tend to give better estimates, the best approach is to start with r , closer to one BS.

The procedure could be generalised. Instead of just using the r_i as starting point, it could be possible to define a number l of starting points. Then there will be l different trajectories. Finally, the trajectory with smallest global error could be chosen as the selected solution.

The algorithm can be rather time consuming for large l . One possible way of reducing computing time is only to pursue the analysis of those trajectories with small accumulated error in a manner similar to the Viterbi's Algorithm.

3.- A completely different approach tries to eliminate the Rayleigh or Shadowing component in the reported measurement. That can be done performing a smoothing of the received data v_m . For example, let us consider the reported levels to station j

$$v_{1j}, v_{2j}, v_{3j}, \dots, v_{kj}$$

Those values usually observe a relatively large dispersion. Those values could be corrected via a "moving average" of size m ($m < k$) procedure

$$v_{mj} = \frac{\sum_{i=m-m+1}^{m+m} v_{ij}}{2 \cdot m + 1}$$

4.- A collection of r can be determined using (1) independently for every v_m . Then, the estimated values (in the x-y plane) can be smoothed following a procedure suggested in [2]

The basic idea is to find a least squares approximation $r_e(t)$ to the collection of estimated locations $r(t_j)$: assuming they correspond to a collection of samples taken at a constant velocity vel .

If v is assumed constant, it can be estimated using the least squares criteria as

$$vel = \frac{\sum_{j=1}^k (t_j - \bar{t}) [r(t_j) - \bar{r}]}{\sum_{j=1}^k (t_j - \bar{t})^2}$$

Being

$$\bar{t} = (1/k) \sum_{k=1}^k t_j \text{ and } \bar{r} = (1/k) \sum_{k=1}^k r(t_j)$$

Then, the estimated positions are

$$r_e(t) = r_e(0) + v t$$

Note the smoothing procedure used in 4 can be also be applied to 1-3.

Simulation results

In the following, the results of using algorithms 2-4 are compared. A simulated area of 4 x 6 km with seven trisect BS is assumed. The coverage is simulated using Hata model (no shadowing). Measured values are evaluated from the actual values adding a lognormal random variable. BS are indicated in the map by a 'b' symbol. Those BS are trisectorial (120 degrees)

In the first set of simulations, the probability location function is obtained.

To obtain it, a MS at a given location and with a given speed is simulated. The MS, located originally at $x = 30,20$ and moving along the vector (1,1) sends 10 sets of 5 measurement reports, each set is sent every 28 meters. Then, the MS is located, using one of the algorithms described in the previous section and the estimated location saved. The process is then repeated 300 times and a collection of estimated locations obtained. Then we represent the locus of those locations, normally the estimated locations are close to the "true" location, but the "spread" gives an indication of the form and value of the expected error

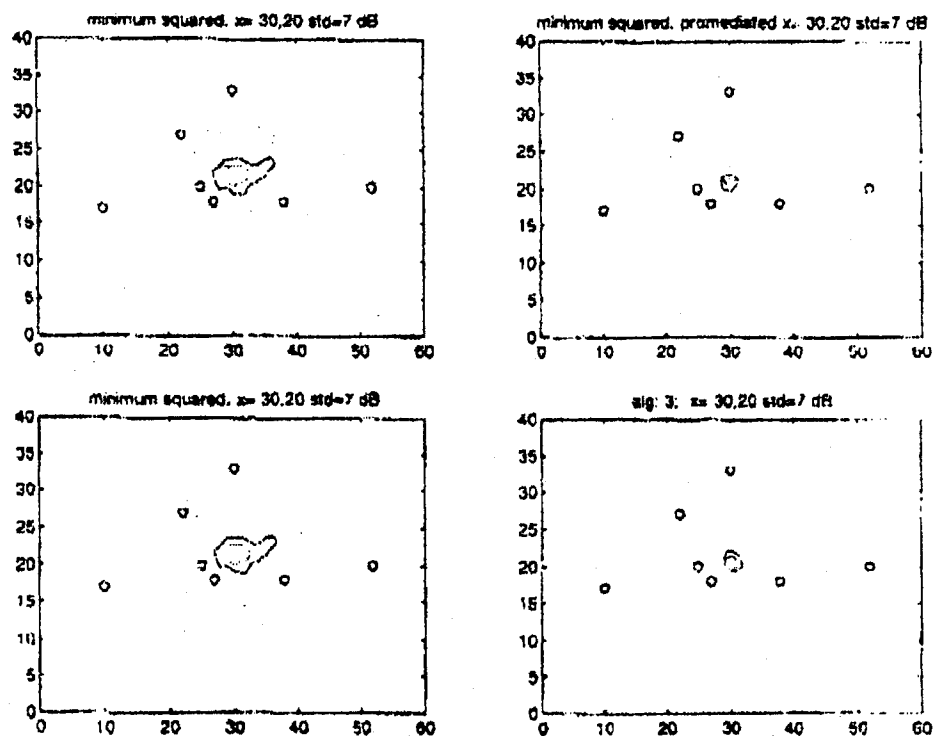


Figure 1. Probability location function. Comparison between different algorithms.

- a) Direct algorithm
- b) Direct algorithm with averaging (assumed at speed zero)
- c) Algorithm number 2
- d) Algorithm number 3

The results, shown in figure 1 (a-d) show the probability location function. The first area (light grey) is the 99% location area, the second corresponds to a 66% and the third to 33%.

It can be seen that different algorithms tend to behave rather similarly, however, algorithm 2 gives the best results, but possibly not enough for the added complexity. Promediate is always a good solution, reducing significantly the indetermination area, which can be as small as 50 m for 66% of cases. If the number of measurements is small, the mobile cannot be too away from the original location and it is not necessary to estimate speed.

However, these good results are not obtained when a MS far from a BS is located. The following set of pictures show the results for a location $x=45,30$

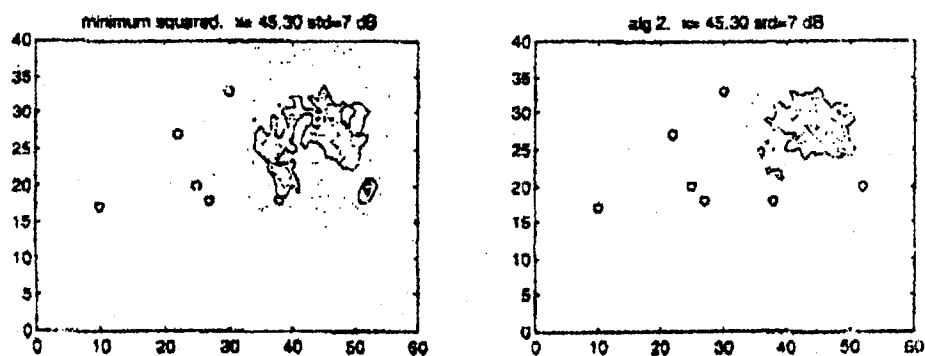


Figure 2. Probability location function. Situation for far locations

- a) Direct algorithm
- b) Algorithm number 2

In this set of results becomes apparent how those locations are not correctly reported. The variability is very large and this explains the erratic behaviour of the location function that seems scattered around.

Next, the mean value of the error is investigated over the whole area of interest. The following set of figures (3 onwards) is obtained repeating the previous process at every location. The mean error value (and its dispersion) is obtained and represented in a 3-D graph.

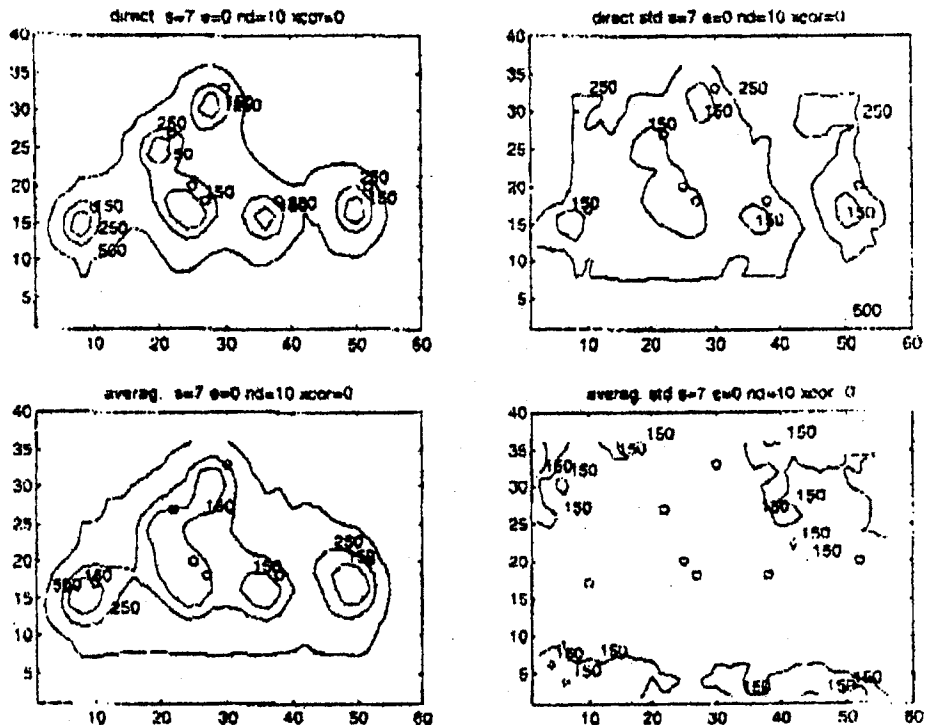


Figure 3. Location squared error. Mean and dispersion

- a) Direct algorithm
- b) Dispersion for a)
- c) Direct algorithm averaged
- d) Dispersion for c)

First, the error at every position is represented in fig 3-a with standard conditions (lognormal std. =7 dB, no extra errors, $e=0$, 10 reports assumed no correlated). It can be seen that only areas close to a BS are acceptable. In this set of results, the averaging effect appears clear. Fig. 3 a shows the average error (with algorithm 1). It can be seen that most areas have an average error smaller than 500 m. However, the dispersion (shown in 3-b), is also around 250, leading to a very bad accuracy. When averaging is used (3 c) the mean error improves but, most significant, there is a large decrease in dispersion (less than 150 m). Location errors in the central area can be in the region of 250m for that lognormal standard deviation of 7 dB.

Next, an study of different parameters is undertaken and briefly reported in Figure 4.

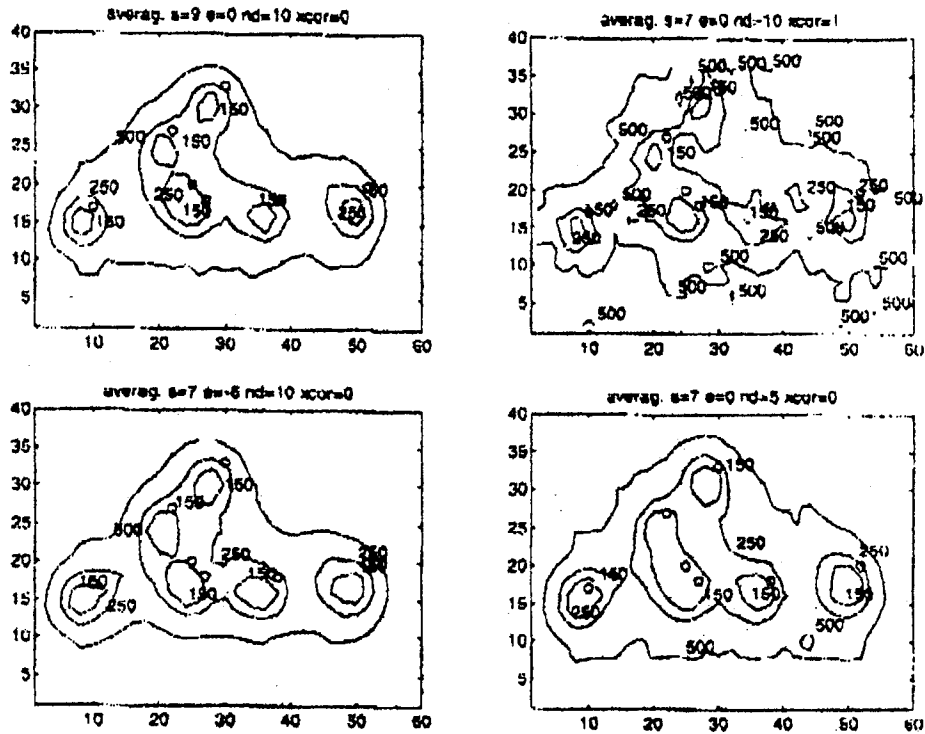


Figure 4. Dependence on different parameters

- e) Increase the lognormal to 9 dB
- f) Assume correlated samples
- g) Decrease number of processed samples
- h) Add 6 dB extra error. Fixed.

This new set of figures show the effect of increasing the lognormal deviation value from 7 to 9 dB. It should be compared with 3-c) which are the standard conditions.

Figure 4-a shows the results of increasing the standard deviation of the lognormal variable to 9 dB. The effect is not too large, but a clear shrinking of the 250 m. error zone is apparent. The area is divided into three smaller subareas and the 150 m error zone is very much reduced. Other experiments, not reported here, show a steady decrease at larger standard deviation values.

Figure 4-b corresponds to the case of assuming correlated samples. The effect is important in this case, even greater for the variance. This implies the need for *independent* reports. If the MS to be located is not moving, the location algorithm will not be effective since Rayleigh fading will mask the position.

In figure 4-c a constant error of 6 dB is added to the coverage map. This is done to analyse the possibilities of indoor location. The effect is very significant for the 150 m error area, but not so for 250 m. This is hopeful, since it implies this technique is not completely useless inside buildings, provided the error requirements are not stringent. Finally, in figure 4-d, the effect of reducing the number of samples per location to 5 is analysed. The effects are small in the mean error. However, the spread of the measures, not represented here to avoid a too large document, shows a clear increase. Of course, this is connected to the correlation study: if the number of samples is large but correlated, the location would be poor. If the samples are uncorrelated, even 5 samples can be enough for rough location purposes.

Implementation

The previous analysis have been used to develop a number of analysis tools aimed to obtaining a better understanding of the mobility behaviour of mobile users and to aid to the planning process.

Three sets of tools have been developed:

- 1) Tools for investigating the presence of hot spots in a given area.
- 2) Hot spot location.
- 3) Traffic loading analysis tools for the study of microcells introduction.

In the following, a brief overview of the use and performances of the tools is reported.

Identification of hot-spots

As indicated in the article, hot-spot identification requires a reasonably good coverage information. Errors cannot be larger than 6-7 dB. Those values are not usually achieved in an urban area unless a measurement campaign for adjusting propagation prediction values is performed. That measurement campaign is normally a costly process and it would be desirable to have an indication *whether in a given cell there is an abnormal traffic accumulation*.

That is the objective of the first set of tools. They are used prior to a detailed measurement campaign and the objective is just to find out the likely presence of a "hot spot" or an abnormal traffic accumulation. No coverage information is used and, therefore, the indication is fairly straightforward.

Figure 5 is an example of the type of results obtained in two typical situations, which could be considered representative. In both figures, the (x,y) axis represent the received power for the two most frequently reported stations. Z axis is the number of reports.

Figure 5-a shows an scenario where there are no clear hot-spot, whereas 5-b seems to deserve a more careful analysis.

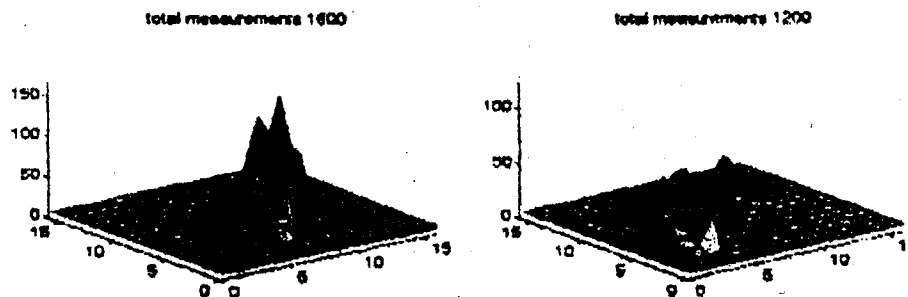


Figure 5. Analysis performed to determine the presence of a hot-spot.

5-a) cell with a clear hot spot

5-b) cell with evenly distributed traffic

Hot-spot location

(Once the above process is performed and a candidate for a hot-spot is found, a more detailed coverage map is required. In the present case, a measurement campaign has been performed. The results of the measurements are extrapolated to cover a larger area. Then, the algorithms presented in the first part of the contribution are applied.

The results, confirmed approximately by field measurements are shown in fig. 6. First an algorithm check was performed. For that purpose a measurement campaign including location positioning using GPS and in-site information was performed. Errors (difference in meters between the actual position and the reported position) are shown in fig 6-a. As can be expected, locations near to a BS have less error.

Finally, fig 6-b shows the application of the location algorithm to the real traffic in the cell. As predicted in the pre-test, the traffic is concentrated in a small number of locations. Those locations corresponds to an insurance company office where a lot of traffic can be expected.

Those results will be used by the operator to determine the possible location of a BS in the area.



Figure 6. Use of the hot spot location tool

5-a) Likely distribution of errors

5-b) Geographical traffic distribution

Conclusions

The results presented here and other experiments, it can be concluded that location based on coverage maps is feasible in certain areas in the neighbourhood of BS. Location errors are in the region of 200 m in good areas for reasonable values of coverage map errors, which could be large for some services. However, since those are mean values, they tend to compensate when a large number of mobiles are to be located – as is the case of hot spot location.

Different algorithms tend to perform similarly. Averaging is a very good measure to counteract the observed dispersion. In this contribution very simple filtering procedures have been presented. As a corollary, location based on coverage maps can only be performed with some success if more than 7-10 (uncorrelated) reports are available.

The effect of increasing imprecision on the coverage map, and the difficult location of indoor users is less important than initially feared. If location stations are properly located and errors are not too large, the results may be acceptable.

Those analysis have been used to produce a number of analysis tools. The first step is to determine whether a hot spot really exists. The second, would be to locate its position. That step requires a more specific coverage information, which normally necessitates an specific measurement campaign.

References

- [1] J. Jiménez, H. González, E. Alonso: *"Mobile Location for Hot Spot Identification"* Cost 759 TD (98) 12. Bern. Jan 98
- [2] M. Hellebrandt, R. Mathar, M. Scheibenbogen. *"Estimating Position and velocity of Mobiles in a Cellular Radio Network"*. IEEE Trans. Vehicular Tech. Feb. 1997