

# Fusion of Neural Network positioning and Database Correlation in localizing a Mobile Terminal.

Claude Takenga, Chen Xi  
Institute of communications engineering  
University of Hannover  
Appelstrasse 9A, 30167 Hannover, Germany

Kyandoghene Kyamakya  
Transportation Informatics  
Alpen Adria University Klagenfurt  
Universitaetsstr. 65, A-9020 Klagenfurt, Austria

**Abstract** - Mobile phone location activities have been intensified recently. This work will examine the possibility of combining the Neural Network (NN) positioning and the database correlation (DC) localization system in order to realize a robust RSS based positioning system within a GSM network context. The NN is trained first offline before it use in the positioning unit. Function approximation and classification properties of the NN will be investigated and the best NN architecture will be chosen from a series of candidates. During positioning phase both positioning results from NN and DC are fused with a robust tested method in order to increase the precision of the position estimate. Results show that, the fusion process reduces the positioning error.

**Keywords**- data fusion, database correlation, localization, neural network, mobile positioning.

## I. Introduction

Most accurate positioning systems today are achieved using satellite-based positioning. However, the GPS has to be used only in case of a clear sky. This makes its use impossible or not reliable in urban areas, mountainous terrains and covered spaces. Many researchers have worked on alternative positioning systems suited for such environment. There are several types of Localization Based Systems:

- Cell ID, the precision of this method is 300m in urban areas, 2km in suburban areas and 3-4km in rural zones.
- Enhanced Cell ID With this method one can get a precision similar to Cell ID, but for rural areas, with circular sectors of 550 meters.
- Time of Arrival and Angle of arrival, which suffer in case of severe multipath.
- E-OTD: This is similar to TOA, but the position is estimated by the mobile phone, not by the

base station. The precision of this method depends on the number of available LMUs in the networks, varying from 50 to 200 m.

- Fingerprint methods have been preferred as they perform better in areas with severe multipath propagation compared to others [1-4], moreover, they do not require any additional equipment for their implementation.

This paper fuses both NN position estimates and the location estimations from the database correlation in order to get better results. The major problem for both NN and DC localization methods is the creation of the fingerprint database. The signal fingerprint can be collected either by measurements or by a computation network planning tool. Measurements collection is time consuming but produces more accurate fingerprint data. In this work we used both methods. Collected data are used for the NN training and predicted data [5] for the database correlation method. Our algorithm is tested with real collected data.

## II. Literature overview

A number of papers have paid attention on positioning of a mobile system within a GSM environment using the received signal strengths [1-4]. Some works exploit the use of data fusion techniques combining information from different sources in order to have better results [6-8]. Other works focus on the dead reckoning technique, which estimates the present position by projecting heading and speed from a known past position [9,10]. This method suffers from an accumulation of the positioning error. A constant update can be made in order to remedy to this issue.

In [11] a mobile positioning using GSM cellular phone and artificial NN has been presented. Classification and Function approximation properties of NN have been exploited, but only one method for classification was considered.

In this work, a survey of different neural training algorithms and architectures adapted to this RSS fingerprint positioning is conducted. Function approximation and classification properties of the NN are investigated in order to consider the best candidate suited for a given scenario. Two types of classifications are presented and implemented for the conducted

experiments. The benefit from a fusion process between the NN position estimate and Database correlation position estimate is exploited.

### III. NN positioning method

The NN acquires intelligence during training. This process is performed off line before using it in the positioning algorithm. During training, the NN parameters are set in order to make a good mapping between RSS and location with the least squared error. Positioning using NN can be tackled in two different ways: as a function approximation problem and as a multi-class classification problem.

#### A. Function approximation

During training, the NN parameters are set in order to approximate a function which represents the mapping between the RSS and positions (x-y) with the least squared error.

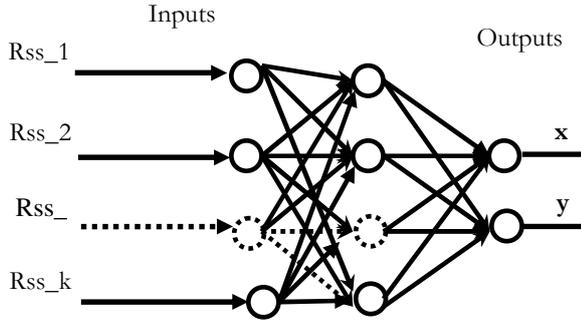


Figure.1 NN architecture for function approximation

The number of inputs corresponds to the number of Cell antennas considered. In our case, 10 inputs were used. The outputs in this function approximation case are two (i.e. x,y-coordinates of the location position) as shown on Fig.1. The number of hidden layers and their corresponding number of neurons are fixed experimentally.

#### B. Classification

In classification case, the experimental area is divided into squared sections. During training, predicted RSS together with their corresponding section numbers are provided to the NN which makes a generalization. In positioning phase, the location estimation of the MS section using received power levels from different base stations is the classification task.

The difference here is that the number of outputs corresponds to the number of sections to which we divides the experimental area.

A set of RSS is given at the input. If these data were taken at a point belonging to sector number 2, then all the target inputs are 0 except for the second target which value is 1

as shown on Fig. 2. This series of 0100 is referred to in this work as codeword.

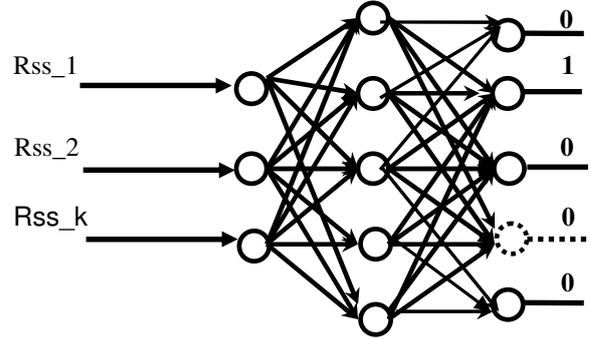


Figure.2 NN architecture for classification

During positioning phase, it appears that these values of the codeword change and are not 0 at all the positions and 1 at the rank corresponding to the location estimate section number. They take values lying in interval 0-1. This is due to the fact that, the training error is not null and the training and testing points are different.

Two different methods of deciding on the section estimate at the NN output can be applied: First, choose the rank of the maximum value in the codeword to be the section number. We will refer to this method as Absolute Maximum method. The Second method, consider all rank with values greater than a certain weighted value. These considered ranks contribute to the choice of the section estimate by finding their center. We will refer to this as Weighted Method.

At the NN output, the so called weighted quadratic

average value ( $q_{average}$ ) and so called weighted variance ( $\sigma$ ) are calculated at every tested point. All the rank of the output values laying in the interval  $[q_{average} - \sigma, q_{average} + \sigma]$  are taken into account. The center of the selected sections is then calculated and considered to be the location estimate.

$$q_{average} = \frac{\sum_{i=1}^n s_i^2}{\sum_{i=1}^n \frac{s_i}{s_{max}}} \quad (1)$$

$$\sigma = \frac{\sum_{i=1}^n \left[ (s_i - q_{average}) \cdot \frac{s_i}{s_{max}} \right]^2}{\sum_{i=1}^n \frac{s_i}{s_{max}}} \quad (2)$$

Where  $s_i$  is the  $i$ th output of the codeword.

$s_{max}$  is the max value of the output of the codeword.  
 $n$  – number of sections in which our area is divided

### C. Comparative results

#### Experiment 1

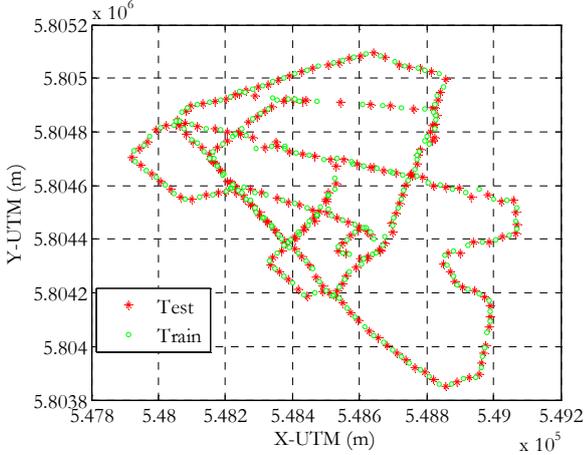


Figure 3. Sample positions used for training and test.

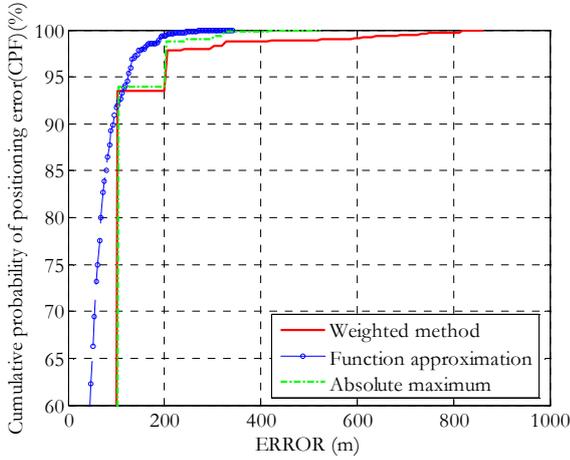


Figure 4. Cumulative distribution functions for the three methods

Fig.3 shows the sample positions in our experimental area which were used for the NN training and which were used in positioning phase. The training and testing points are 10m distant one from another. The three methods perform good as we can see in fig.4. Absolute maximum appears to perform better than the weighted, because of a good representation of the training data in the entire concerned area. As a result, in the codeword at the NN output, the rank of the maximum value is more likely to be the true position.

We can see also that for both types of classification methods, we can get true estimation of the section with 93% of probability. The size of the sectors is considered 100mx100m which makes a maximum positioning error of 100m.

#### Experiment 2

In Fig.5, we have another case. Here, the training is made with a small quantity of data and which are not uniformly located on the experimental area. This is a case whereby, data at the entire area can not be available for training or

in order to remedy to time consuming process during data collection, only a small sample is used for this purpose.

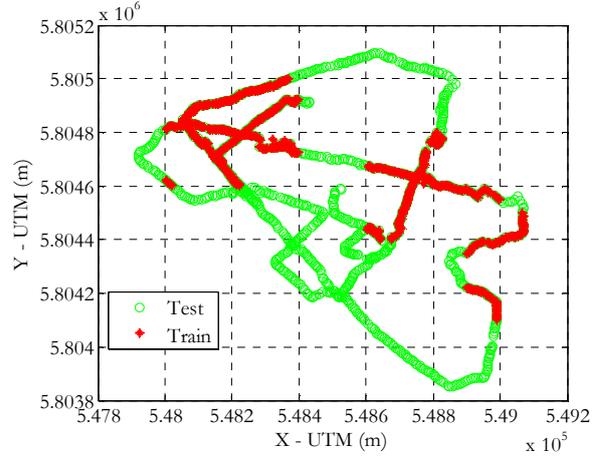


Figure 5. Sample positions used for training and test.

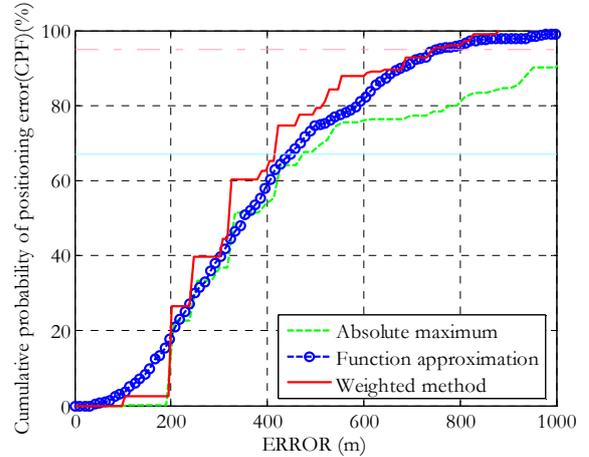


Figure 6. Cumulative distribution functions for the three methods

In Fig.6, it turns out that the Weighted Method could be the best solution. This can be explained by the fact that, the error is more likely to be bigger and the max value at the NN output codeword does not indicate the true section estimate. A contribution of values bigger than a weighted limit could be a better indication for the location estimate.

### IV. Database correlation method

The key Idea of the DC is to store any location-dependent signal information that can be measured by a MS and corresponding positions in a database for the whole interested area that is used by a location server [3]. In this work we use predicted RSS as the signal fingerprint. When the MS need to be located, the necessary measurements are performed and transmitted to the location server. The location server then calculates the MS location by comparing the transmitted fingerprint and the fingerprint of the database. In DC method, searching time should be reduced by limiting the searching process around the area where the mobile unit is located. The

Timing Advance and Cell Identification information help to limit the searching area Fig.7. Other robust pattern matching methods for database correlation such as, those using Bayer filter have been investigated in [4] and could be used for better results. In this work, the metric used to decide on the matching location is the mean quadratic error. The coordinates associated with the fingerprint that provides the smallest mean quadratic error is returned as the estimate of the position.

## V. Fusion

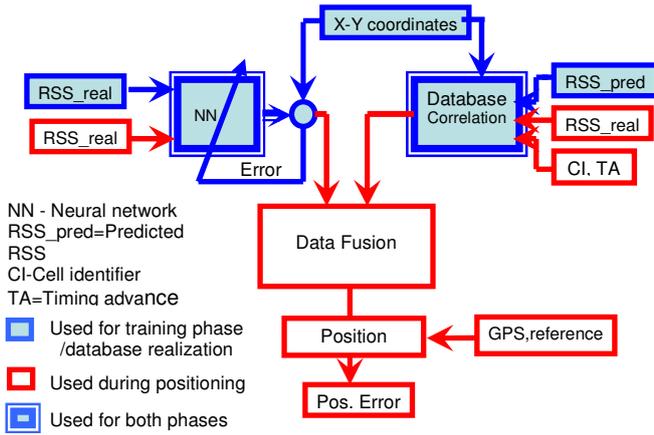


Figure 7. Block diagram of the fusion process

The training of the NN is made off line and takes time for a good generalization. After training, a positioning test can be performed in order to decide whether it satisfies our requirements or the training set should be reselect, NN architecture should be changed for a new training to be performed. As result, we have at the NN output the position estimate  $A$  with a known variance  $\sigma_1^2$  got from the performance tests. From the probability theory, we know that, in case of a Gaussian density function, 68% of the probability is contained within the band  $\sigma$  unit to each side of the mean. The same way, DC method provides us with an output  $B$  having a certain variance  $\sigma_2^2$  which is also known after some tests. Having both positions and their corresponding variances, a fusion unit processes these information to produce a new location estimate  $C$  with a variance  $\sigma_3^2$ .

The new position estimate  $C$  and variance are found through the formulas bellow:

$$C = \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\right).A + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right).B \quad (3)$$

$$\frac{1}{\sigma_3^2} = \left(\frac{1}{\sigma_1^2}\right) + \left(\frac{1}{\sigma_2^2}\right) \quad (4)$$

We should note that  $\sigma_3$  is less than either  $\sigma_1$  and  $\sigma_2$ , which is to say that the error in the position estimate has been decreased by combining the two pieces of information [5-7].

## VI. Experimental settings and fusion results

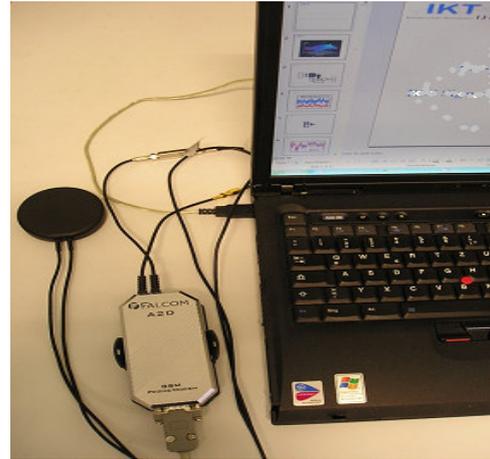


Figure 8. GSM-GPS antenna and Modem used

The experimental environment is an urban area with some high buildings. Experiments were performed in an area of size, 3km by 3km in which 10 GSM cell antennas were considered. The collected and predicted RSS used are from E-plus mobile network, Germany. A deterministic method was applied to predict the rss of our concerned area [5]. The pixel resolution of these provided predicted RSS is 5m.

In positioning phase, real collected RSS with a GSM-GPS antenna and modem were used in the algorithm, Fig.8.

Fig.9. shows us the fusion impact on the positioning accuracy. For this part of the experiment, we have chosen one of the worse cases. We trained the NN with data available from small parts of the whole area. This scenario is similarly to fig.5, but taking the training data as testing and vice versa.

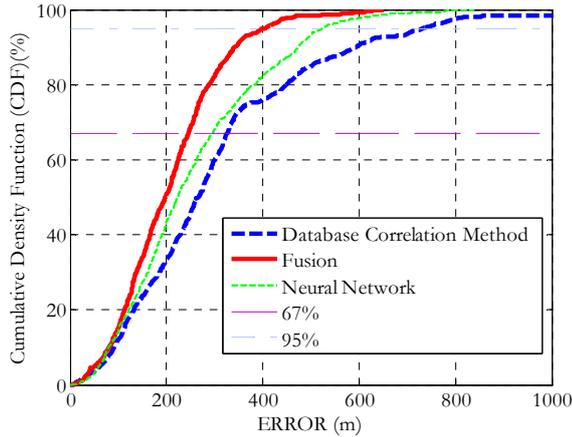


Figure 9. CDF for NN,DC and Fusion NN-DC

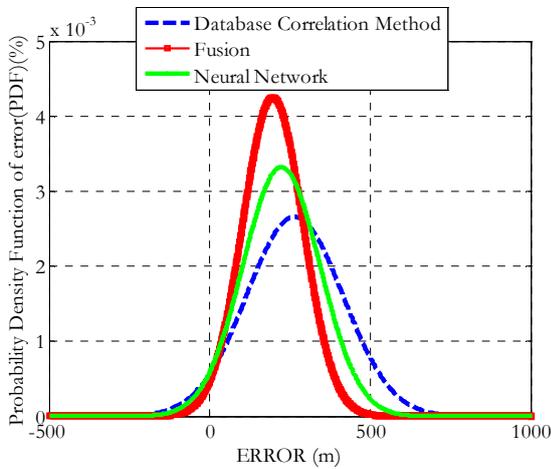


Figure 10. PDF for NN,DC and Fusion of both methods

We have at the NN output the position estimate with a variance  $\sigma_1^2 = (120.2\text{m})^2$  got from the performance tests. The same way, DC method provides us with an output having a variance  $\sigma_2^2 = (150.2\text{m})^2$ . The new position estimate got from the fusion result as shown in (3) has a variance  $\sigma_3^2 = (93.8\text{m})^2$  which is less than both  $\sigma_1$  and  $\sigma_2$  as shown in Fig.10. This leads to a conclusion that, the error in the position estimate has been decreased by combining the two pieces of information [5-7].

## VII. Conclusion

The accuracy of different methods to yield the position estimate of a MS in a GSM network has been presented. An investigation of different NN positioning method was undergone for the given problem. Results in the second experiment show that a classification NN positioning method can give better results compared to function

approximation in case where training data are available for only some sectors of the whole experimental area.

A conventional database correlation positioning method was presented. The performance of this DC method could be further increased by involving more robust pattern recognition algorithms such as Bayer Filter.

A fusion theory was investigated in order to combine positioning results from NN and DC sources for the accuracy improvement. Results illustrated on Fig.9,10 show that, positioning error can be reduced by fusing positioning information coming from the two units. The new positioning estimated error is less than the error at both NN and DC output, even if the DC error is quite big. With these results, we can conclude that, even poor quality data provide some information and should thus increase the precision of the final output after fusion.

In order to reduce further the positioning error, both NN positioning unit and DC should be further improved. The training data can be reselected in order to cover the entire experimental area. For the DC, a use of more robust pattern matching algorithms could improve its accuracy. Future works are focused on the post processing of these DC results using an extra NN in order to further reduce the DC positioning error.

## References

- [1] C. Takenga, K. Kyamakya " Location Fingerprinting in GSM network and Impact of Data Pre-processing," 'in press' WMC'06, Munich, 2006.
- [2] R. Yamamoto, H. Matsutani, H. Matsuki, T. Oono, and H. Ohtsuka, "Position Location Technologies using signal strengths in cellular system," presented at VTC-Spring, 2001, Rhodes Island..
- [3] H.Laitinen, J.Lahteenmaki,T.Nordstrom, "Database Correlation Method for GSM location", presented at IEEE VTC'01
- [4] M.Khalaf-Allah, K. Kyamakya "Database Correlation using Bayes Filter for Mobile Terminal Localization in GSM Suburban Environments", Proceedings of the 63<sup>rd</sup> IEEE Vehicular Technology Conference (IEEE VTC 2006-Spring), May 7-10, 2006, Melbourne, Australia.
- [5] T. Kurner and A. Meier, "Prediction of outdoor and outdoor-to-indoor coverage in urban areas at 1.8 GHz," IEEE on selected areas in communications, vol. 20, 2002.
- [6] P.S.Maybeck, "Stochastic models, estimation, and control" Academic Press Inc, 1999
- [7] V. Belur. Dasarathy, "Multisensor, multisource information fusion: architectures, algorithms and applications", Bellingham, Wash. : SPIE, 2005
- [8] D.L.Hall, S.A.H.McMullen "Mathematical Techniques in Multisensor Data Fusion", Artech House, Boston, 2004
- [9] Cliff Randell Chris Djiallis Henk Muller, "Personal Position Measurement Using Dead Reckoning".
- [10] Gerald P. Roston and Eric P. Krotkov, " Dead Reckoning Navigation for Walking Robots". Raleigh, NC July 7- 10,1992
- [11] Z. Salcic, E.Chan, "Mobile Station Positioning using GSM cellular phone and artificial Neural Networks", Wireless Personal Communications, vol..14, p.235-254,2000