Research Article

# LOCATION ESTIMATION ON MOBILE NETWORKS

**Ahmed Hakan Kılıç[†], Ali Boyacı[††], Serhan Yarkan[‡]**

[†] İstanbul Commerce University, Engineering Faculty, Department of Computer Engineering, Istanbul, Turkey.

[††] İstanbul Commerce University, Engineering Faculty, Department of Computer Engineering, Istanbul, Turkey.

[‡] İstanbul Commerce University, Engineering Faculty, Department of Computer Engineering, Istanbul, Turkey.

**ahakan.kilic@istanbulticaret.edu.tr, aboyaci@ticaret.edu.tr, syarkan@ticaret.edu.tr**

0000-0002-5523-5363, 0000-0002-2553-1911, 0000-0001-6430-3009

**Atıf/Citation:** Kılıç, A., H., Boyacı, A., Yarkan, S., (2022). Location Estimation On Mobile Networks. Journal of Technology and Applied Sciences 4(2), 199-204

## ABSTRACT

This paper reports the location estimation on Mobile networks using Base Station (BTS) data. Processed data have been collected from the field as TA (Timing Advance), RSRP (Reference Signal Received Power), and RSRQ (Reference Signal Received Quality) measurements. We also gathered the corresponding Global Positioning System (GPS) to the measurements. Location estimation results compared to the actual location. We gathered the accurate sites of the users to increase the service quality of the BTS. This article was produced from the thesis titled “Location Estimation on Mobile Networks.”

**Keywords:** Base Station, GPS, Mobile Networks, Machine Learning, RSRP, RSRQ, Timing Advance (TA)

# MOBİL AĞLARDA LOKASYON TAHMİNİ

## ÖZET

Bu makale, Baz İstasyonu (BTS) verilerini kullanarak Mobil ağlarda konum tahminini raporlamaktadır. İşlenen veriler sahadan TA (Timing Advance), RSRP (Reference Signal Received Power) ve RSRQ (Reference Signal Received Quality) ölçümleri olarak toplanmıştır. Ölçümlere karşılık gelen Global Konumlandırma Sistemi (GPS) verileri de toplanmıştır. Lokasyon tahminleri gerçek lokasyonla (GPS) karşılaştırılmıştır. Baz istasyonunun hizmet kalitesini artırmak amaçlanmıştır. Bu makale “Mobil Ağlarda Lokasyon Tahmini” başlıklı tezden üretilmiştir.

**Anahtar Kelimeler:** Baz İstasyonu, GPS, Mobil Ağlar, Makine Öğrenmesi, RSRP, RSRQ, Timing Advance (TA)

| | | |
|---|---|---|
| **Geliş/Received** | : | **03.08.2021** |
| **Gözden Geçirme/Revised** | : | **13.08.2021** |
| **Kabul/Accepted** | : | **13.08.2021** |

## 1. INTRODUCTION

Location estimation on Mobile networks is becoming an essential topic in recent years. Whether it is an emergency, improving service quality, or providing an alternative way of location estimation. BTS data can be used to locate the users connected to it. BTS does not feature a location-providing service, but we can manipulate and fine-tune the data they provide to estimate the user's location. There are several ways to estimate the location of the user using BTS data.

Whether using Microcell Zone concept (Samarah, 2016), using TDOA and AOA measurements with Nelder-Mead algorithm (Kyunghyun Lee, 2019), applying Kalman Filtering or triangulation method which data from three BTS is used to triangulate the coordinates of the user relative to the BTS (Anisetti, Ardagna, Bellandi, Damiani, & Reale, 2011). Although the triangulation method seems the easiest way of locating user, it has some hard aches. BTS are not located in a similar geographic location, and they should be configured in the environment they placed to cover all the areas to improve the service quality. The data gathered from the rural areas may be different from a crowded city center.

In our approach, we used actual data gathered from the field. There is no need for special hardware or updates on the BTS. BTS are not dedicated to providing location service, and we just used the data already available on the BTS. Our methodology is based on the geographical position of three BTS and triangulation. In addition to triangulation, we also used the GPS coordinates corresponding to the measurement data we have. We had an opportunity to compare our result with the actual location. We located the users connected to the BTS, and by doing this we may be able to increase the service quality. Configuration of the BTS for the environment can further improve estimating the location of the user. BTS can improve the service quality by locating the user's location.

The remainder of this paper is organized as follows. Literature review, data and problem, solution and methodology and conclusion sections.

## 2. LITERATURE REVIEW

The document in (Samarah, 2016) presented a location estimation of a user in a Microcell Zone Concept (MZC) Mobile Station by retrieving TA from the BTS. In MZC, two or more zone sites are connected to the same BTS. This gives an advantage of the user being served with the strongest signal within the zone. When the user moves from one zone to a different zone Mobile Station Controller (MSC) changes the channel to the zone. The user will remain in the same frequency, so there will be no need for a handoff procedure. The study conducted using MATLAB simulation. But in the real environment, results may be adversely affected due to geographic location or BTS requiring more sophisticated changes to change the channel from one zone.

(Kyunghyun Lee, 2019) also performed simulations for using TDOA and AOA measurements using Nelder-Mead (NM algorithm). They proposed a method that is a combination of TDOA and AOA to improve the accuracy of location estimation. They applied NM algorithm to reduce the environmental error to enhance the accuracy of location estimation. They confirmed effectiveness with simulation results, but in real environments such as crowded locations, the results may be affected.

(Anisetti, Ardagna, Bellandi, Damiani, & Reale, 2011) proposed identify possible locations from signals using the Database Correlation Method (DCM). They then described a technique to deal with signal fluctuations to select paths with greater accuracy. Kalman filtering is also applied to gather information about roads and build a path to improve mobility approximation. They also used actual data collected from the field.

The paper in (Hernández, Arteaga, Pérez, Orozco, & Villalba, 2019) presented a system to locate user's location in an emergency. They used an android application to extract data. Their method uses known positions of BTS. But the application developed may not be compatible with all Android devices.

In our approach, we used actual data gathered from the field. We also collected the Global Positioning System (GPS) coordinates corresponding to these measurements. Then we compared the algorithm results with the corresponding GPS coordinates to the measurements.

## 3. DATA AND PROBLEM

Data we gathered from BTS are TA (Timing Advance), RSRP (Reference Signal Received Power), and RSRQ (Reference Signal Received Quality). We also collected the GPS coordinates corresponding to these measurements. By just using these data, the average distance to the BTS cannot be found. These measurements are correlated with the distance; we did fine-tune and used these measurements in the distance calculation.

TA (Timing Advance) is the time that it takes to a signal to reach from the user's device to the BTS.

BTS calculates the delay of the data, and if it sees a delay, it increases the TA value by 1. The maximum TA value is 63. This TA value can tell the BTS how far the user is from it. Each level of TA is 550 meters.

RSRP (Reference Signal Received Power) is the measurement of the power of the signal. RSRP value is measured on dbm type.

RSRQ (Reference Signal Received Quality) is signals quality, which tells us how signal compared to noise. We may have enough power (RSRP), but the quality will be lower if we also have noise in the system. RSRQ value is measured in dB. (ETSI TS 136 214 V15.5.0 (2020-01), 2020)

We also applied the formula $\left(\frac{x}{16}\times 78\right)+18$ for calculating the TA. We calculated the First TA and Last TA by using $\left(\frac{x}{16}\times 78\right)+18$ formula, x being First TA or Last TA. Taking advantage of these, we trained several machine learning algorithms.

When the calculation is made according to the TA formula, the user is within the radius. We tried to find the location by reducing the radius of this circle. The tolerance can be seen in the yellow-colored area on the fig. 1. The yellow-colored area is the margin; we want it to be more precise.

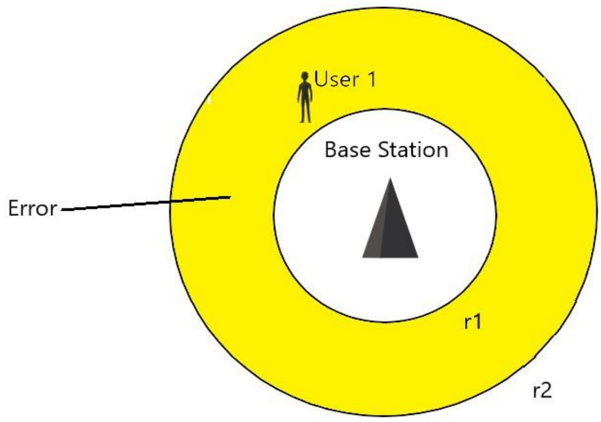


**Figure 1**. Error and Margin Presentation

R1 is the radius of the BTS, and the distance between R1 and R2 is the error margin. Users can be anywhere in the yellow-colored area. It comes at regular intervals; we can find a sharper location using these intervals to make it more precise than a circle.

User connection transfer from one BTS to the next BTS is called handover. By doing triangulation using BTS data, we can estimate the location of the user.

The problem here is every BTS has different physical conditions, different geographical structures, hills, pits, weather conditions, demographic density, and roads affect the signal. One universal model will not be sufficient to solve this problem because the situation in each BTS is different. In our approach, we used different models for each BTS.

## 4. SOLUTION AND METHODOLOGY

We used different models for each BTS. While creating these models, we used eight machine algorithms. We used more than one machine learning algorithm because every BTS has different conditions, as mentioned Data and Problem section. These Machine algorithms are Gaussian Processor, KNeighbor Regressor, Linear Regression, Automatic Relevance Determination Regression (ARD), Least Angle Regression (Lars), Least Absolute Shrinkage and Selection Operator (Lasso), Ridge and Bayesian Ridge Regression.

Machine learning is a computer science where it uses data and various algorithms to learn and improve its accuracy. Gaussian Processor is used for classification predictive modeling (Ruan, Milstein, Blackwell, & Miller, 2017) , KNeighbor Regressor is a method which approximates the relationship between independent variables and the continuous outcome by averaging the observations in the same neighborhood (Hirose, Soejima, & Hirose, 2021), Linear Regression establishes a relationship between the dependent variable and one or more independent variables using the best fit straight line (Hirose, Soejima, & Hirose, 2021), ARD is very similar to Bayesian Ridge Regression, but it effectively prunes away redundant features (T. Van Gestel, 2001), Lars is used with high dimensional data where it finds the correlated property to the original value (Efron, et al., 2004), Lasso regression plays an important role not only in reducing overlearning, but also in feature selection (Li & Li, 2010), Ridge Regression is obtained by adding a regularization term to our cost function in linear regression. With this addition, the learning algorithm both learns the data and tries to keep the model weights as small as possible (Li, et al., 2020) and Bayesian Ridge Regression estimates a probabilistic model of the regression problem (Pereira, Abreu, & Rodrigues, 2020)

Data harvested from the downtown part of a big city in Turkey. It is a crowded city environment. Before training these machine learning algorithms, we performed data cleaning by removing 31 BTS out of 397. The removed BTS data contained some outlier values and was affecting the result of the machine learning algorithms. The machine learning algorithms are run for all 366 distinct BTS.

For regression metric we used Mean Square Logarithmic Error (MSLE). MSLE can be explained as calculation of ratio between original and predicted values. It takes log of the original and predicted values. It is a variation of the mean squared error. In MSLE relative difference between original and predicted is important. Using MSLE with regression helps avoiding large errors to be punished compared to small errors.

The formula of MSLE can be seen on fig. 2. The Gaussian and Bayes algorithm performed poorly compared to others. The other algorithms had similar performance.

$$MSLE = \frac{1}{n}\sum_{i=1}^{n}(\log(y_i + 1) - \log(\bar{y} + 1))^2$$

**Figure 2.** Mean Squared Logarithmic Error Formula

We gathered the Mean Square Logarithmic Error (MSLE) values of all the algorithms we mentioned before and used their logarithm values in fig. 3. Fitting error can be seen on the figure.

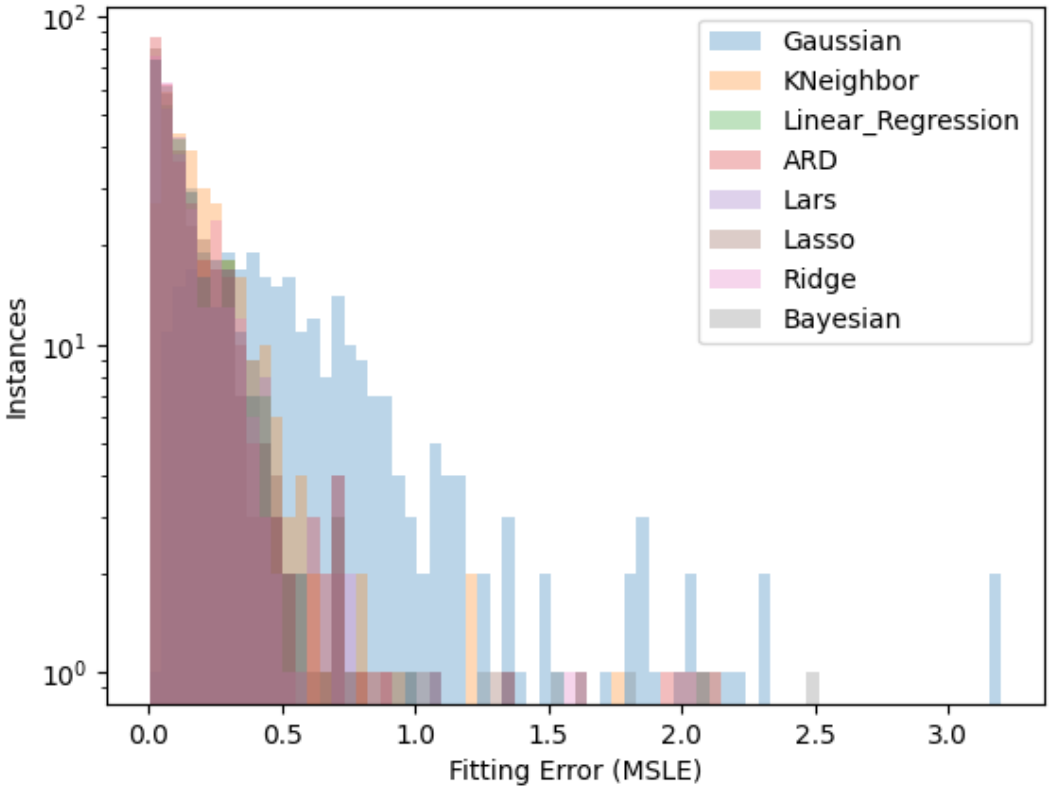


**Figure 3**. Mean Square Logarithmic Error Histogram

## 5. CONCLUSION

In this study, we gathered data from the field as TA, RSRP, and RSRQ with corresponding GPS coordinates. We then used several machine learning algorithms. The result of these machine learning algorithms compared to the GPS coordinates to measure our accuracy. The mean squared log error table can be seen in Table 1. As we mentioned earlier, for better accuracy, the geographic location of the BTS should be considered. To address that, we used different machine learning algorithms for each BTS. Most of the machine learning algorithms performed similarly on the calculation of the distance. We may include the BTS environment data to improve our results and make more precise results in our feature work.

**Table 1.** Mean Squared Logarithmic Error Results

| Machine Learning Algorithm | Mean Squared Log Error |
|---|---|
| Gaussian | 0.59 |
| KNeighbors | 0.22 |
| Linear Regression | 0.17 |
| Ard Regression | 0.17 |
| Lars | 0.17 |
| Lasso | 0.16 |
| Ridge | 0.16 |
| Bayes | 0.19 |

**REFERENCES**

Anisetti, M., Ardagna, C. A., Bellandi, V., Damiani, E., & Reale, S. (2011). Map-Based Location and Tracking in Multipath Outdoor Mobile Networks. IEEE Transactions on Wireless Communications.

ETSI TS 136 214 V15.5.0 (2020-01). (2020). Retrieved from www.etsi.org: https://www.etsi.org/deliver/etsi_ts/136200_136299/136214/15.05.00_60/ts_136214v150500p.pdf

Hernández, L. A., Arteaga, S. P., Pérez, G. S., Orozco, A. L., & Villalba, L. J. (2019). Outdoor Location of Mobile Devices Using Trilateration Algorithms for Emergency Services. IEEE.

Kyunghyun Lee, H. L. (2019). Optimised solution for hybrid TDOA/AOAbased geolocation using Nelder-Mead simplex method. IET Radar, Sonar & Navigation.

Samarah, K. G. (2016). Mobile Positioning Technique Based on Timing Advance. International Journal on Communications Antenna and Propagation (I.Re.C.A.P.), Vol. 6, N. 4.

Ruan, W., Milstein, A. B., Blackwell, W., & Miller, E. L. (2017). Multiple output Gaussian process regression algorithm for multi-frequency scattered data interpolation. *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE.

Hirose, H., Soejima, Y., & Hirose, K. (2021). NNRMLR: A Combined Method of Nearest Neighbor Regression and Multiple Linear Regression. 2012 IIAI International Conference on Advanced Applied Informatics. IEEE.

Li, C., & Li, W. (2010). Partial least squares method based on least absolute shrinkage and selection operator. 2010 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE). IEEE.

Li, D., Ge, Q., Zhang, P., Xing, Y., Yang, Z., & Nai, W. (2020). Ridge Regression with High Order Truncated Gradient Descent Method. 2020 12th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC). IEEE.

Pereira, R. C., Abreu, P. H., & Rodrigues, P. P. (2020). VAE-BRIDGE: Variational Autoencoder Filter for Bayesian Ridge Imputation of Missing Data. 2020 International Joint Conference on Neural Networks (IJCNN). IEEE.

T. Van Gestel, J. A. (2001). Automatic relevance determination for least squares support vector machine regression. International Joint Conference on Neural Networks. Proceedings. IEEE

**ACKNOWLEDGEMENT and DECLARATIONS**

This publication was produced from the Master thesis of Ahmed Hakan KILIÇ in the Computer Engineering Program of Istanbul Commerce University, Engineering Faculty