INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2026, VOL. 72, NO. 2, PP. 1-7
Manuscript received January 05, 2026; revised April 2026. doi: 10.24425/ijet.2026.157920

# Mobile cellular network-Based Positioning using Machine Learning

Reda Yagoub, Samia Bentaieb, Roumaissa Anberi, and Fouzia Bakhti

*Abstract*—**Positioning systems are essential for various applications, ranging from navigation to location-based services. In the case of mobile cellular networks. This study explored the use of mobile cellular network signals to develop a positioning system for beehives. Our research introduces a new approach for positioning using mobile cellular (LTE, UMTS, and GSM) radio signal data. We trained a machine learning model to predict geographic coordinates (latitude and longitude) based on various parameters extracted from the intercepted radio signals. These parameters include the Cell Identity, Mobile Country Code (MCC), Mobile Network Code (MNC), Location Area Code (LAC), and additional relevant identifiers.Our study offers a novel approach for precise position determination by utilizing information provided by mobile cellular signals.**

*Keywords*—**LTE; GSM; UMTS; Positioning; Machine Learning; Data preprocessing; Random Forest; Cell Id; localization**

## I. INTRODUCTION

POSITIONING systems play a crucial role in modern technology, enabling a wide range of applications, from navigation to emergency services.

Mobile cellular networks, including GSM, UMTS, and LTE technologies, serve as viable alternatives for positioning. These networks are widely available and provide extensive coverage, making them promising candidates for location estimation applications. By utilizing signal metrics from cellular networks, it is possible to determine the location of a device even in scenarios where GPS signals are weak or inaccessible.

This study aims to explore the feasibility of mobile cellular network signal metrics for developing a positioning system. We focused on a set of features derived from these signals, including the Mobile Country Code (MCC), Mobile Network Code (MNC), and Location Area Code (LAC), to train a predictive model. Specifically, we employed a Random Forest algorithm to predict geographic coordinates (latitude and longitude) based on these features.

Our research contributes to the existing body of knowledge by demonstrating the potential of using mobile cellular networks for positioning. We provide detailed insights into the model training process, evaluation metrics, and overall performance of the system. The findings of this study have significant implications for the development of reliable and efficient positioning systems that can operate in various environments. The main contributions of this study can be summarized as follows:

1. This study involved rigorous data collection and processing of LTE, UMTS, and GSM signals, then preparing the metrics for localization tasks.
2. For the prediction of geographic locations, we tested Random Forest (RF), after verifying its reliability with various machine learning algorithms for consistent performance.
3. Comprehensive error evaluation was conducted by defining and analyzing error ranges with results further illustrated through Folium-based visualizations that clearly compared ground-truth and predicted positions.

## II. STATE OF THE ART

Positioning has long relied on GNSS (e.g., GPS) for outdoor accuracy, and alternative and complementary technologies such as Wi-Fi, Bluetooth, and cellular signals have been thoroughly investigated. According to recent surveys, cellular-based positioning is appealing because of its widespread infrastructure and capacity to provide coarse-to-fine location estimates in various settings [1].

Traditional cellular methods include Cell-ID, Timing Advance (TA) / TOA, TDOA, and AOA. These geometric/time-based methods can achieve good accuracy in well-instrumented networks but typically require synchronization, dense base station deployments, or antenna arrays, constraints that limit their practicality in many real-world deployments. Cell-ID and Enhanced Cell-ID remain widely used for their low complexity; however, their metric accuracy is constrained by cell size and topology [1].

With the rapid advancement of machine learning, data-driven positioning approaches have recently gained increasing attention. Instead of relying solely on geometric models, these approaches exploit large datasets of signal measurements to learn the complex relationships between signal features and geographical locations. Fingerprinting solutions construct signatures from RSSI/RSRP/RSRQ/RS-SNR, cell identifiers, multiband measurements, and contextual features, and then apply KNN/WKNN, Random Forests, gradient boosting, or deep neural networks to perform localization. Deep learning variants increasingly transform signal fingerprints into image-like representations (e.g., multi-channel “fingerprint images”) and feed them to convolutional or hybrid architectures, reporting substantial improvements, particularly in complex outdoor or urban settings [2],[3]. The results demonstrate that machine learning-based approaches can significantly outperform classical techniques.

---

Reda Yagoub, Roumaissa Anberi and Fouzia Bakhtiare are with University of Ain Témouchent - Belhadj Bouchaib, Po Box 284, 46000, Algeria (e-mail: reda.yagoub@univ-temouchent.edu.dz, anberiroumaissa@gmail.com, bakhtifouzia37@gmail.com ). Samia Bentaieb is with University of Ain Témouchent - Belhadj Bouchaib , Po Box 284, 46000, Algeria, Laboratory Signals and Images (LSI) (e-mail: samia.bentaieb@univ-temouchent.edu.dz).

In Algeria and other developing countries, limited research has been conducted on cellular-based positioning using real-world operator data. The present study extends this line of research by collecting a large-scale dataset across the three main Algerian operators (Mobilis, Djezzy, and Ooredoo) and applying advanced machine learning methods to improve accuracy. This study highlights the potential of data-driven positioning for practical deployment in regions with heterogeneous network conditions.

## III. LOCALIZATION METHODS BASED ON MOBILE CELLULAR NETWORK

Several localization techniques based on mobile cellular network parameters have been proposed, each with its own advantages and disadvantages.

Time of Arrival (TOA): This technique determines the position by measuring the travel time of signals between the base station and mobile phone. While it typically achieves 50-200 meter accuracy, it requires precise time synchronization between transmitter and receiver, which poses implementation challenges [4].

Time Difference of Arrival (TDOA): This method measures the difference in arrival times of signals at several base stations. Each TDOA measurement defines a hyperbola representing the potential positions, and the predicted location is indicated by the intersection of these hyperbolas. Even if it has the ability to achieve great accuracy, a dense network of synchronized base stations is required. Usually offers precision within a range of 100-200 meters [5], [6].

Angle of Arrival (AOA): The AOA creates a line of position by determining the direction in which the signal arrives at the base station. The position of the mobile device can be triangulated by combining the AOA data from various base stations. To precisely measure the direction of transmission, this approach requires base stations to be equipped with antenna arrays. This method achieves accuracy levels between 100 and 300 m [6], [7].

Cell Identity (Cell ID): The simplest method, Cell ID positioning, locates mobile devices by identifying their connected cell towers [8]. Although this method is straightforward and does not require additional hardware, its accuracy is limited by the size of the cell, which can be large in rural areas. E. Trevisani, A. and Vitaletti [9], presents an experimental study of the Cell-ID positioning technique in the United States and Italy. The results obtained show that the average distance between the GPS location and the estimated Cell ID location was approximately 500 m in Italy and approximately 800 m in the United States.

Enhanced Cell ID (E-CID): This technique improves the basic Cell ID positioning by incorporating additional metrics, such as the Timing Advance (TA). E-CID can significantly enhance accuracy, particularly in urban environments with dense cell tower deployments. Research by Borenovic et al. [10], demonstrated that the accuracy of E-CID improves with the number of base stations providing TA values, achieving errors as low as 105 meters with 8-9 base stations.

Table I shows a comparison of positioning methods based on mobile cellular network signals in terms of accuracy.

Table I
Comparison of Various Positioning Techniques based on mobile cellular network

| Method | Accuracy | Strengths | Limitations |
|---|---|---|---|
| Time of Arrival (TOA) | 50-200 meters | High accuracy | Requires precise time synchronization |
| Time Difference of Arrival (TDOA) | 100-200 meters | Accurate with dense network | Needs synchronized base stations |
| Angle of Arrival (AOA) | 100-300 meters | Effective in urban areas | Requires antenna arrays at base stations |
| Cell Identity (Cell ID) | 500-2000 meters | Simple and straightforward | Limited accuracy, dependent on cell size |
| Enhanced Cell ID (E-CID) | As low as 105 meters | Improved accuracy with multiple base stations | Depends on availability of multiple TA values |

## IV. METHODOLOGY

### A. *Description*

The data used in this study were collected from the mobile cellular networks of the main Algerian mobile network operators, Mobilis, Djezzy, and Ooredoo. These three mobile operators account for the vast majority of the mobile telecommunications market in Algeria and provide coverage in both urban and rural areas. Data collection was achieved using a smartphone-based application designed specifically for this study so that real user-level measurements could be captured in an unobtrusive and realistic manner. By using a smartphone application to collect performance data from real users, the measurements provide a closer representation of the true Quality of Experience (QoE) than the performance statistics reported by the operators.

To better understand the distribution of the measured performance data, Fig.1 shows the distribution of the records collected according to the three main mobile network operators. As can be seen, Mobilis accounted for almost 45% of the records, followed closely by Djezzy with just under 40% and then Ooredoo with just over 15%. This distribution of records reasonably reflects the availability of users and network coverage at the time of data collection. While the distribution of achievable records is informative, it is important to consider that the significant imbalance between operators may impact the relative representativeness of the measurements and will thus be accounted for in the analysis that follows to mitigate any potential bias.

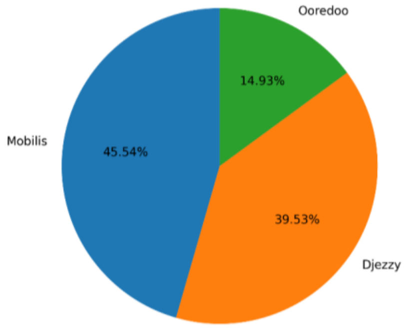

Fig. 1. Distribution of Dataset Records by Mobile Network Operator

### *B. Data Collection*

The measurements were carried out over several weeks in various geographic locations, such as urban areas, highways, and rural areas, at different times of the day to capture the spatiotemporal variability of the radio environment. Each record contains a geographic position (latitude and longitude) and several radio signal characteristics. The dataset consists of more than 19,000 records each contains a range of radio signal characteristics as well as contextual information that are summarized in Table II.

Table II
Summary of Dataset Features and Measurement Units

| Acronym | Description | Unit |
|---|---|---|
| MCC | Mobile Country Code | - |
| MNC | Mobile Network Code | - |
| LAC | Location Area Code | - |
| Cell ID | Cell Identifier | - |
| RNC ID | Radio Network Controller Identifier | - |
| PSC | Primary Scrambling Code | - |
| ASU | Arbitrary Strength Unit | - |
| RSS | Received Signal Strength | dBm |
| Accuracy | GPS Accuracy | meter |
| Speed | Speed | km/h |
| Altitude | Altitude | meter |
| Network Type | Network Type (UMTS, LTE, 4G) | - |
| RSRP | Reference Signal Received Power | dBm |
| RSRQ | Reference Signal Received Quality | dB |
| RSSI | Received Signal Strength Indicator | dBm |
| RSSNR | Reference Signal Signal-to-Noise Ratio | dB |
| Ec/No | Energy per Chip to Noise Density Ratio (for UMTS) | dB |

Fig.2 shows some GPS points of the collected data in university of Ain temouchent.

Fig. 2. Some GPS points of collected data.

### *C. Pre-processing*

For data preprocessing used in this study, the feature set was divided into two groups: numerical and categorical features. The numerical features used are: MCC, MNC, LAC, Cell_ID, RNC_ID, PSC, ASU, dBm, Accuracy, Speed, Altitude, RSRP, RSRQ, RSSI, RSSNR, and Ec/No.

The numerical variables were standardized using the StandardScaler method to formulate the data such that they had a mean of zero and unit variance, according to Equation 1. This means that, regardless of their original scale, all numerical attributes are treated equally during the learning process.

$$x' = \frac{x - \mu}{\sigma} \tag{1}$$

where $x$ is the original feature value, $\mu$ is the mean of the feature, $\sigma$ is the standard deviation of the feature, and $x'$ is the standardized feature value.

The categorical feature Net_Type is assigned to a specific technology (e.g., LTE, UMTS) and is transformed into binary indicator variables using the OneHotEncoder, which is designed to create training data that remains robust even with unseen categories/values during the test set. The preprocessing pipeline with the full steps of normalizing all numerical and categorical attributes was finalized prior to model training.

### *D. Machine Learning-based Prediction Method*

Machine learning (ML) refers to computational algorithms that learn feature patterns or relationships from data to make predictions or decisions without explicit programming. These models are expected to generalize from the training data to work successfully on unseen data. Hence, they allow a wireless device to intelligently sense its environment and predict how the network will behave in different situations, with the overall goal of using intelligent prediction to maintain the quality of service in mobile communication systems. Predictive capabilities enable devices to detect variations in logical wireless channels, networks, and users if they have observed a few observations per channel, per traffic pattern, and across network composition, content requests, or user context, thereby allowing the device to proactively manage the service in wireless environments to ensure a quality user experience.

To achieve the localization of mobile cellular networks, the study presented in this paper employs a set of radio and network features that together indicate system identifiers, signal quality, and contextual mobility. The system-level identifiers include Mobile Country Code (MCC), Mobile Network Code (MNC), Location Area Code (LAC), Cell Identifier (Cell ID), Radio Network Controller (RNC), Primary Scrambling Code (PSC). Signal measurements were represented by Arbitrary Strength Unit (ASU), signal strength units in dBm, Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), Received Signal Strength Indicator (RSSI), Reference Signal Signal-to-Noise Ratio (RSSNR), and Ec/No (for UMTS). The contextual features include GPS accuracy, user speed, altitude, and network type (2G, UMTS, 4G, etc.) to capture the device mobility patterns and environmental context. These features may or may not be used to obtain localization; however, some features are designed for comparisons in the mobility space.

To utilize the features in device localization, a set of four ML models were employed: the K-Nearest Neighbors (KNN) model, Support Vector Regressor (SVR) with Radial Basis Function (RBF) kernel, Random Forest (RF), and XGBoost. Collectively, the models learn complex spatial relationships with the RF model and how these relationships occur over time within the radio environment. Each model is expected to improve the location accuracy in mobile cellular networks.

KNN is a similarity-based machine learning algorithm. Unlike parametric models that learn fixed parameters, KNN stores the training dataset and predicts based on the nearby data points. When classifying a new instance, it identifies the K closest neighbors and assigns the majority class among these neighbors as the predicted label. The choice of K is crucial to the model performance. The distance metric used in KNN determines the nearest neighbor. While the Euclidean distance is common, other measures, such as the Manhattan distance or cosine similarity, can be used based on the data. KNN's simplicity of KNN makes it suitable; however, it can be computationally expensive and sensitive to irrelevant features, requiring careful feature selection [11], In this study, the model was applied with to simultaneously predict both latitude and longitude.

SVR is the regression version of the Support Vector Machine (SVM) algorithm. Its key idea is to find a function that deviates from the actual targets by at most a small margin while also being as flat as possible [12]. The RBF kernel (radial basis function) is commonly used because it allows SVR to model nonlinear relationships by mapping data into a higher-dimensional feature space

Random forest, introduced by Breiman, is an ensemble learning method that extends the concept of decision trees by constructing multiple trees and aggregating their predictions [13]. This approach involves creating independent decision trees, each trained on a randomly selected subset of the training data. The random selection process, known as bootstrap aggregating or bagging, helps reduce overfitting and increase the model's generalization capability.

In addition to using random subsets of data, the random forest incorporates feature randomness. At each node split in a tree, only a random subset of features is considered for the split decision-making. This feature randomness further increases the diversity among the trees, making the ensemble more robust to noise and outliers. The final prediction of a random forest is typically obtained by averaging the predictions of all individual trees for regression tasks or by majority voting for classification. This aggregation of diverse trees leads to improved predictive accuracy and stability compared to single decision trees and other simpler ensemble methods.

In this study, the RF model was applied using decision trees. To address the prediction of multiple continuous outputs, namely, latitude and longitude, the algorithm was adapted using a multi-output regression framework.

XGBoost is an ensemble learning algorithm that leverages decision trees and the boosting technique, where errors from earlier models are addressed by those that follow the model. The final model was constructed as a weighted combination of weak learners, refined through gradient boosting to minimize a specified loss function [14]. One of the key strengths of XGBoost lies in its ability to handle various types of data and learning tasks, including regression. The algorithm employs regularization techniques to prevent overfitting and improve generalization performance. Additionally, XGBoost incorporates built-in cross-validation and early stopping mechanisms, allowing for efficient model tuning and selection. Its scalability and performance optimization make it particularly well-suited for handling large-scale datasets and complex prediction tasks.

In this study, the XGBoost model was implemented with 100 boosting rounds, and a multi-output regression framework was adopted to simultaneously predict the latitude and longitude.

## V. RESULTS AND DISCUSSION

In order to evaluate the accuracy of predictive models, it is essential to use error metrics that have both mathematical and geographical meaningful. In this study, we focused on two complementary measures: the Mean Absolute Error (MAE) and geodesic error (expressed in meters). Additionally, we computed the standard deviation (STD) to assess the variability of the errors.

Given $n$ samples, with true latitude and longitude $(lat_i, lon_i)$ and predicted coordinates $(\widehat{lat}_i, \widehat{lon}_i)$, the MAE is computed as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left(\left|lat_i - \widehat{lat}_i\right| + \left|lon_i - \widehat{lon}_i\right|\right) \tag{2}$$

Since degrees are not always meaningful in terms of real-world distances, we also calculate the geodesic error. For each pair of true and predicted coordinates $(lat_i, lon_i)$ and $(\widehat{lat}_i, \widehat{lon}_i)$, the error is defined as:

$$d_i = \text{geodesic}((lat_i, lon_i), (\widehat{lat}_i, \widehat{lon}_i)) \tag{3}$$

where $d_i$ is expressed in meters.

In addition to the mean error, the standard deviation (STD) of the errors provides information on the variability of the predictions. For geodesic errors $d_i$, the STD is calculated as follows:

$$STD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2} \tag{4}$$

where $\bar{d}$ is the mean geodesic error. A lower STD indicates that the errors are more consistent across all predictions, whereas a higher STD suggests larger variability.

To ensure robust and reliable results, a five-fold cross-validation approach was employed. This technique involves partitioning the dataset into five subsets, using four for training and one for validation in each iteration. By rotating through all possible combinations, this method mitigates the risk of overfitting and provides a more comprehensive assessment of the model performance across different subsets of the data. The reported errors represent the average values obtained from these multiple iterations with their corresponding STD.

The results presented in Table III indicate that Random Forest and K-Nearest Neighbors (KNN) achieved the best performance, with mean geodesic errors of 0.28 meters and 0.22 meters, respectively. While the MAE values in degrees are extremely small (on the order of $10^{-6}$), the geodesic error provides a more intuitive interpretation in terms of real-world distance. KNN achieved the lowest mean error, although with a higher variability (standard deviation of 5.9 meters) compared to the RF (3.8 meters). In contrast, SVR performed poorly, yielding an average geodesic error of approximately 133 meters, while XGBoost showed moderate results with a mean error of 27.3 meters.

TABLE III
PERFORMANCE COMPARISON OF ML MODELS IN TERMS OF MAE AND GEOLOCATION ERROR

| Model | MAE mean (°) | MAE std (°) | Geo error mean (meters) | Geo error std (meters) |
|---|---|---|---|---|
| RandomForest | 0.000002 | 6.48e-07 | 0.277 | 3.834 |
| KNN | 0.000001 | 7.21e-07 | 0.219 | 5.922 |
| SVM | 0.000849 | 4.11e-06 | 132.794 | 61.033 |
| XGBoost | 0.000176 | 3.44e-06 | 27.257 | 21.069 |

Fig.3 shows the relative importance of the features in the RF predictive model. The results indicate that altitude and speed are the most significant features for the predictive model, with the highest importance scores for the variables. This indicates the strong effects of mobility dynamics and elevation differences on the predictive model. Of additional importance is the fact that the accuracy of the GPS measurements seemes to be critical, indicating the relevance of spatial aspects in wireless localization applications. The other features (RNC, RSSNR, and Cell ID) provided good contributions, yet the signal variables (RSRP, RSRQ, PSC, and RSSI) had relatively low importance scores.

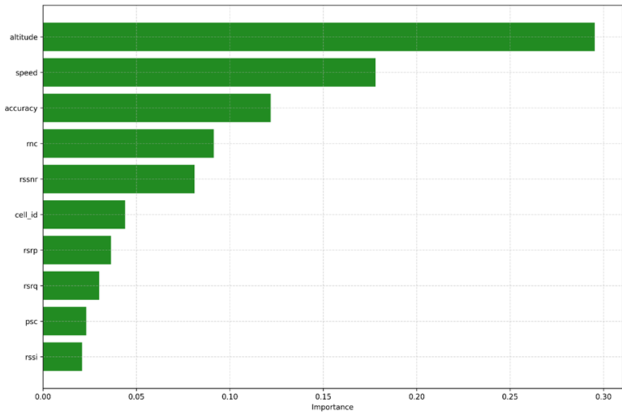

Fig. 3. Top 10 features ranked by importance in the Random Forest model.

The memory footprints for all machine learning models considered in this study are shown in Fig.4. The models explained a fair amount of variability, with SVM using only 0.005 MB and Random Forest using 24 MB or more memory per sample. XGBoost and KNN fell between SVM and Random Forest with memory footprints of 0.26MB and 6.36MB, respectively.

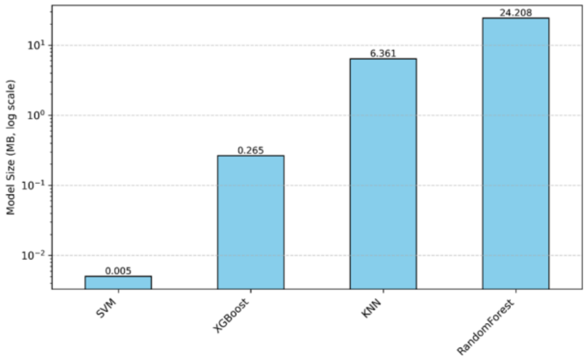

Fig. 4. Comparison of model sizes for the different models.

The reliability of the proposed approach was evaluated beyond the initial dataset using a test set of 572 samples gathered two weeks after the original data collection and the obtained RF model. This temporal separation allowed for an assessment of the model's performance on more recent, potentially evolved data, providing insights into its robustness and generalizability over time.

The distance between the real and predicted positions is presented in Table IV. A total of 68.5% of the samples fall within a very short range of less than 5 metres, which is indicative of high positioning accuracy in most instances. Very few records were located in the 5–30 meters range (0.9%–2.3%), while 6.0% were in the 40–100 meters range. There was a marked gradation from 100 to 200 meters (13.8%) and 5.3% to records greater than 200 meters. The distribution shows that although most measurements provide high accuracy, a notable number of outliers are found at larger distances. Outliers at these larger distances may invigorate the localization potential.

TABLE IV
PREDICTED POSITIONS BY DISTANCE RANGE

| Distance Range (meters) | Number of Records | Percentage of Total |
|---|---|---|
| Less than 5 meters | 391 | 68.5% |
| Between 5 and 10 meters | 5 | 0.9% |
| Between 10 and 20 meters | 10 | 1.8% |
| Between 20 and 30 meters | 13 | 2.3% |
| Between 30- 40 meters | 9 | 1.6 |
| Between 40 and 100 meters | 34 | 6.0% |
| Between 100 and 200 meters | 79 | 13.8% |
| More than 200 meters | 30 | 5.3% |

Compared to traditional methods, such as Cell ID positioning or Enhanced Cell ID (E-CID), this approach offers improved accuracy. The use of machine learning enables the integration of diverse signal metrics, thereby enhancing precision even in challenging environments.

We examined the mean inference time per sample to assess the viability of the proposed model for deployment in real-time applications. Our results indicate that, on average, the prediction time for one sample reached 6.083 ms. This result indicates that the proposed model can process multiple samples per second with minimal delay. Therefore, the model can be deployed within real-time wireless communication systems with rapid, continuous prediction to support adaptive decision-making and sustained quality of service.

Fig.5 offers a comprehensive visual representation of the Random Forest (RF) localization performance through four distinct cases mapped using Folium. The color-coding scheme employed for each sample enhanced the clarity and distinguishability of the data points. In each pair:

- The actual position is denoted by a star (★).
- The predicted position is indicated by a plus sign (+).
- Each sample was assigned a unique color, facilitating the easy matching of real and predicted points within the same pair.

Subfigures (a)–(d) showcase a spectrum of localization outcomes, ranging from highly accurate to significantly erroneous predictions. Subfigures (a) and (b) exemplify cases of excellent and good localization, where the predicted positions closely align with the actual locations, resulting in errors of less than 5 meters and between 5 and 10 meters, respectively. Subfigure (c) presents a scenario of moderate accuracy, with the predicted location deviating from the actual one by 40-100 meters. In contrast, subfigure (d) illustrates a case of poor localization, where the predicted position significantly diverges from the actual location, resulting in a substantial error of more than 200 meters.

Fig. 5. Some examples of ground-truth locations and RF predicted positions. (a) excellent positioning (b) good positioning (c) moderate positioning (d) poor positioning

## VI. CONCLUSION

This study demonstrated the feasibility of using LTE, UMTS, and GSM signal metrics for precise positioning by employing a Random Forest model to predict geographic coordinates. The results showed satisfactory accuracy, highlighting the potential benefits of using mobile cellular networks as a supplementary or alternative approach to estimate localization rather than GPS. This is especially relevant in cases where the GPS may be weaker, blocked, or unavailable. In addition to emphasizing the benefits of RF-based positioning for mobile localization, this study considers the need to manage large localization errors.

Several research directions can be explored in future studies. First, hybrid approaches utilizing Random Forest in combination with other machine learning models or deep learning architectures could help significantly reduce localization errors and improve generalizability. Second, localization can be improved by adding contextual attributes such as Wi-Fi, Bluetooth, or inertial sensor data, particularly in complex environments. Third, the proposed method may be further enhanced with adaptive and online learning, allowing the system to continuously refine its predictions as network conditions change. Finally, large-scale real-world testing that incorporates multiple geographic or environmental conditions is required to examine the capabilities of the method to appropriately train and indeed be an effective method.

## REFERENCES

[1] F. Mogyorósi, P. Revisnyei, A. Pašić, Z. Papp, I. Törös, P. Varga, and A. Pašić, "Positioning in 5G and 6G networks—A survey," Sensors, vol. 22, no. 13, pp. 4757, 2022. https://doi.org/10.3390/s22134757

[2] D. Li, Y. Lei, and H. Zhang, "A novel outdoor positioning technique using LTE network fingerprints," Sensors, vol. 20, no. 6, pp. 1691, 2020. https://doi.org/10.3390/s22134757

[3] Z. Hu, X. Chen, Z. Zhou, and S. Mumtaz, "Localization with cellular signal RSRP fingerprint of multiband and multicell," IEEE Journal on Selected Areas in Communications, vol. 42, no. 9, pp. 2380–2394, 2024. . https://doi.org/10.1109/JSAC.2024.3414000

[4] Y. Xie, Y. Wang, B. Wu, and X. You, "Mobile localisation with TOA, AOA and Doppler estimation in NLOS environments," Transactions on Emerging Telecommunications Technologies, vol. 23, no. 6, pp. 499–507, 2012. https://doi.org/10.1002/ett.2505

[5] F. Gustafsson and F. Gunnarsson, "Possibilities and fundamental limitations of positioning using wireless communication networks measurements," IEEE Signal Processing Magazine, vol. 22, no. 4, pp. 41–53, 2005. . https://doi.org/10.1109/MSP.2005.1458284

[6] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: A look at positioning aspects for future sensor networks," IEEE Signal Processing Magazine, vol. 22, no. 4, pp. 70–84, 2005.
https://doi.org/10.1109/MSP.2005.1458289

[7] M. I. Simić and P. V. Pejović, "A comparison of three methods to determine mobile station location in cellular communication systems," European Transactions on Telecommunications, vol. 20, no. 8, pp. 711–721, 2009.

[8] R. Yagoub, M. Benaissa, B. Benadda, and M. Merzoug, "Low cost RTL-SDR based positioning solution based on GSM cell ID," in Proc. 6th International Conference on Image and Signal Processing and their Applications (ISPA), pp. 1–6, 2019.
https://doi.org/10.1109/ISPA48434.2019.8966924

[9] E. Trevisani and A. Vitaletti, "Cell-ID location technique, limits and benefits: an experimental study," in Proc. 6th IEEE Workshop on Mobile Computing Systems and Applications, pp. 51–60, 2004. https://doi.org/10.1109/MCSA.2004.9

[10] M. Borenovic, M. Simic, A. Neskovic, and M. Petrovic, "Enhanced cell-ID + TA GSM positioning technique," in Proc. EUROCON 2005 – The International Conference on "Computer as a Tool", pp. 1176–1179, 2005. https://doi.org/10.1109/EURCON.2005.1630164

[11] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A brief review of nearest neighbor algorithm for learning and classification," in Proc. 2019 International Conference on Intelligent Computing and Control Systems (ICCS), pp. 1255–1260, 2019.
https://doi.org/10.1109/ICCS45141.2019.9065747

[12] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," Advances in Neural Information Processing Systems, vol. 9, 1996.

[13] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001. https://doi.org/10.1023/A:1010933404324

[14] T. Chen, T. He, M. Benesty, V. Khotilovich, Y. Tang, H. Cho, K. Chen, R. Mitchell, I. Cano, and T. Zhou, "XGBoost: extreme gradient boosting," R Package Version 0.4-2, vol. 1, no. 4, pp. 1–4, 2015. .