Automatic NLOS Classification from Virtual 3D City Models

Ivo Müürsepp
Thomas Johann Seebeck Department of Electronics
Tallinn University of Technology
Tallinn, Estonia
ivo.muursepp@ttu.ee

Muhammad Mahtab Alam
Thomas Johann Seebeck Department of Electronics
Tallinn University of Technology
Tallinn, Estonia
muhammad.alam@taltech.ee

***Abstract* – The paper describes the use of publicly available 3D models of buildings for possible mobile positioning accuracy enhancements. The lack of the line of sight (LOS) between base stations (gNB) and user equipment (UE) is one of main reasons for poor performance of positioning algorithms. If the lack of LOS is automatically detected then given measurement can be simply discarded. Alternative option is mitigation of the lack of LOS to the final positioning estimate. Method uses virtual 3D city models in CityGML (Geography Mark-up Language) format to automatically detect if the straight line from gNB to UE intersects any surface of the surrounding buildings or not. Method was tested with the measurement results collected with the portable 5G test scanner. Classification results were verified visually.**

***Keywords* – CityGML, LOS/NLOS classification, 3D models, 5G NR, Machine Learning.**

I. INTRODUCTION

Radio navigation, in some form or another, has been around almost for a century. Implementation of global navigation satellite systems (GNSS), like Global Positioning System (GPS), has made accurate positioning widely available. Common GNSS receivers can offer positioning accuracy within few meters on almost everywhere on the planet Earth. With enhancements like real-time kinematic positioning (RTK) application, positioning accuracy can be increased up to only few centimetres.

Positioning accuracies mentioned above are not always available but only obtainable under suitable conditions. Multipath propagation and the lack of the line of sight (LOS) in dense urban areas leads to large errors in location estimate. GNSS are also not suitable for indoor positioning needs. Creation of dedicated, ground-based radio navigation system to complement GNSS where it performs poorly, could be solution. Unfortunately, such systems will not be economically feasible.

Authors' version of the paper published in the Eighth International Conference on Fog and Mobile Edge Computing (FMEC 2023). DOI: 10.1109/FMEC59375.2023.10306158 ©2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

At the same time there are terrestrial mobile communication networks available, covering large areas. Idea to use those communication networks also for the radio signal-based positioning is not a new one. As mobile networks are designed for voice- and data communication, not for navigation, then achievable positioning accuracy has been poor. So poor indeed that only practical applications so far have been emergency- and internal security related. For example, the GSM mobile communication system, utilizing the Timing Advance for trilateration, achieved average positioning accuracy of 500 m [1].

Fifth generation (5G) mobile communication system is the first one where the mobile positioning, with usable accuracy, is a built-in feature. Release 16 of 3rd Generation Partnership Project (3GPP) specifies requirements for the 5G new radio (NR) based positioning accuracy. As a starting point for commercial indoor use cases the horizontal accuracy must be better than 10 m and vertical accuracy better than 3 m. For outdoor cases, both the horizontal- and vertical accuracies, must be better than 3 m [2]. Numerous simulations by many different counterparts, summarized in 3GPP TR 38.855 [3], indicate that those demands for positioning accuracies are, at least in principle, achievable.

Influence of non-line-of-sight (NLOS) propagation on positioning accuracy in 5G networks were simulated in [4]. Results showed, unsurprisingly, that the lack of direct line-of-sight LOS between anchor nodes and user equipment (UE) can cause large differences between estimated and actual location of the device. Same holds, in general, for all radio signal-based positioning systems and methods. This in turn, has sparked interest in different NLOS detection and mitigation methods. Few possible approaches are described next.

If there are more measurements available than the minimal number required for positioning then such redundancy can be used to mitigate errors as shown in [5], [6] or [7]. As errors caused by NLOS propagation have one-sided distribution then there are many papers, similar to [8], that are accounting for this constraint. The available knowledge about the geometry of area, where the positioning takes place, can be used in map-based methods like [9].

Many authors have proposed statistics-based NLOS detection and/or mitigation methods. Distribution of measured times of Arrival (ToA) are compared against statistical models in [10]. Ultra-wideband (UWB) technology allows to measure channel impulse response (CIR) and many parameters of received signal. One or more of those parameters can then be used to distinguish between LOS and NLOS channels like it is described in [11] or [12]. Paper [13] describes the use of difference in average channel power for LOS/NLOS detection in case of 5G-based indoor positioning.

Parameters of statistical models are dependent on the exact propagation environment. If system is used in different location then detection algorithm must be retuned for it. Luckily, the machine learning (ML) methods are specifically useful in such a scenario where some complex model must be learned automatically. Because of this, the ML is often suggested as possible solution to LOS/NLOS classification problem. In paper [14] a deep feedforward neural network is used for classification based on measured CIR of UWB channel. Paper [15] addresses issue of speeding up machine learning in new areas based on previously learned relationships.

Training of any supervised ML method, like neural network for example, needs a lot of labelled data. If we want to train LOS/NLOS classifier then there is a need for a large amount of correctly labelled measurement results. For small-scale experiments such labelling can be done manually by human observer. For more practical applications this process must be automated.

In current paper we propose method for automatic collection and labelling 5G NR measurements for LOS/NLOS classification. Labelling itself can be done automatically using publicly available 3D databases of buildings in CityGML (Geography Mark-up Language) format.

The idea to use 3D computer models of buildings for radio propagation estimation and modelling is not totally new. One good example of modelling mobile communication channel in such a way is given in [16]. Google Earth application allows to display 3D visualization of building on map but models itself are not accessible for outsiders. Due this there are methods developed that allow to obtain this information from Google Map, one example being paper [17]. Top shapes of buildings can be imported from the Overpass Application Programming Interface provided by Open Street Map. This approach is used to build custom made application in [18]. It is shown that 3D models of buildings can be used to improve positioning accuracy of global navigation satellite systems [19].

In current work we are using publicly available, high-quality data instead of building models based on flat 2D images as in [18] or trying to somehow acquire restricted data like in [17].

Rest of the paper is organized as follows. Section II of the paper describes the proposed approach and experimental verification of it. Results are discussed in section III followed finally by the conclusions.

II. APPROACH

In case of 5G NR based positioning, the UE has limited number of measurements available that can be used to estimate if there is LOS link between it and selected base station. Those measurements available were synchronisation signal based reference signal received power (SS-RSRP), -signal quality (SS-RSRQ) and signal to noise ratio (SS-SINR) along with channel state information (CSI) reference signal-based quality of downlink channel. This is quite limited data compared, to what is available for UWB devices about signal and channel. In other hand, all 5G terminals are performing those measurements regularly anyway, so using them for improving positioning accuracy demands no additional effort.

Values for those measurements were gathered under both LOS and NLOS conditions. Labelling was done manually by experimenter who visually verified lack or existence of LOS to the 5G base station (gNodeB - gNB) antenna. Analysis of obtained result indicated that LOS and NLOS sets have some small overlap. Most of measurements were separable thus making classification possible. As the decision rule itself is complex and will vary along with the exact area, then the use of the machine learning is reasonable solution to classification problem.

If we want to use ML based classifier in practical solutions then LOS/NLOS labelling of training data must be done automatically. Proposed solution is explained next.

Training data itself is collected by the high-end UE equipped with the GNSS receiver. Along the values of the measured signal parameters, also the positions of the measurements are recorded. Those measurements can be carried out specially for this purpose alone or data can be collected by regular users during everyday operation.

It is clear that 5G positioning is probably not needed for the devices that has built-in GNSS module. But trained classifier will be useful for low-end IoT and reduced-capacity (RedCap) devices without satellite navigation receivers.

Estonia is a small and tech savvy Northern-European country. At the homepage of Estonian Land Board [20] there is open data about 3D building models for all the country. Models are available for download in three different formats. In current work we used CityGML format in which the data about the buildings is stored in both human- and machine-readable XML format. Data in CityGML format can have five different levels of detail (LOD). Lowest of them, LOD0 shows only projections of buildings foundations projected on map. LOD1 approximates buildings with rectangular prisms while LOD2 can show the outlines of buildings with great details and it is currently the highest level available for use. Higher LOD values will increase additional detail, as windows, door, internal walls and even furniture. As they are not available for current research then they are not discussed in greater detail here. Similar models are also available for all of the United States (in LOD1) and many other parts of the world.

CityGML is simply XML format used to store geological information about the buildings and other man-made structures. Files itself can contain a lot of additional information like name of the administrative area, street name and address of each building, type or use of the structure and so on. For our purpose's most of this metadata is not necessary.

Single CityGML file contains usually all buildings in specified geographical location like city, town or region. Every single building in the file is stored as a separate GML multi surface object. This object itself contains number of non-self-intersecting polygons as surface members. Those polygons are representing walls, floors, roof and other possible surfaces of the specific building. Exact number of the polygons depends on the shape of the building and used LOD.

Each polygon is stored as a linear ring type object. A linear ring is a simple geometric shape bounded by at least four points connected by a series of segments. Two-dimensional example of such a shape is shown in Figure 1. The coordinates of the ring's first and last points must be the same in order to close the surface. Again, the exact number of points depends on the shape of the actual building and on the LOD represented. All points have three coordinates: latitude, longitude and height. Locations of the base station antennas, P_{gNB} are known for the mobile network operator. While the position P_{UE} of the UE, used for training data collection, can be acquired by the built in GNSS receiver.

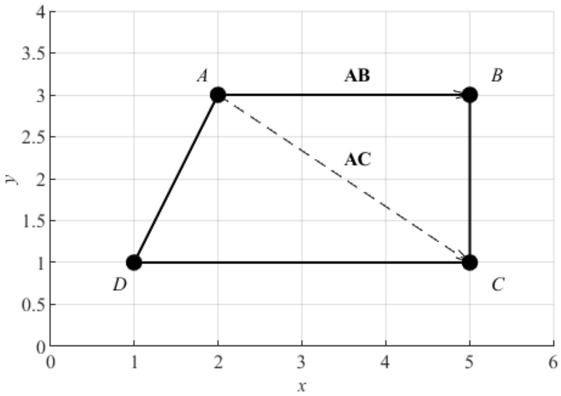

Figure 1. Linear ring type object

Algorithm that is used to determine if there is LOS between UE and any given gNB will be described next. Let the vector **a** be pointed from the position of UE to the position of gNB

$$\mathbf{a} = P_{gNB} - P_{UE}.$$

If the dot product of this vector with the normal vector **n** of polygonal surface is zero, then the line from UE to gNB is parallel to given surface and thus that surface cannot block our view. In order to find the normal vector, we need coordinates of three points *A*, *B* and *C* on a given polygon (see Figure 1). Any three points will do as long as they are not positioned along the same line. Result is vector product of vectors **AB** and **AC** normalized by the product of their corresponding lengths

$$\mathbf{n} = \frac{\mathbf{AB} \times \mathbf{AC}}{|\mathbf{AB}||\mathbf{AC}|}.$$

At first it might seem that reasonable choice of points is just first three elements of linear ring. But as the surfaces are generated automatically, based on Lidar (Light Detection and Ranging) measurements, then even rectangular polygon might have much more than minimal necessary five points. Many redundant points will indeed lie on the same line and thus cannot be used. So, the additional step of checking validity of selected points must be added to the algorithm.

If the dot product $\mathbf{a}\cdot\mathbf{n}$ is not zero, then line from UE to gNB must intersect the plane, on what the polygon under interest lies. The coordinates of this intersection point P_i must be calculated next. At first, we need to find value of the parameter

$$t = -\frac{|\mathbf{G}|}{|\mathbf{H}|},$$

where the matrixes **G** and **H** are formed as

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ A_x & B_x & C_x & P_{UEx} \\ A_y & B_y & C_y & P_{UEy} \\ A_z & B_z & C_z & P_{UEz} \end{bmatrix}$$

and

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ A_x & B_x & C_x & a_x \\ A_y & B_y & C_y & a_y \\ A_z & B_z & C_z & a_z \end{bmatrix}.$$

Coordinates of the intersection point can now be found simply as

$$P_i = P_{UE} + t\mathbf{a}.$$

For next we must just check if the intersection point P_i lies within given polygon or not. This is typical point in polygon (PIP) problem that can be solved in many ways. For example, by ray casting or winding number algorithms. Solving PIP problem by either method requires both polygon and intersection point being located on 2D plane instead of 3D space. This requirement can be fulfilled by projecting both polygon and intersection point into a 2D plane. To minimize possible effects of the limited precision it is recommended to project onto plane that results in largest area of the projected polygon. If the largest component on normal vector **n** is *x*-component then projection should be made onto plane $x = 0$, if the n_y is the largest then plane should be $y = 0$ and so on.

If the intersection point is within polygon then the line, defined by the positions of UE and gNB, must pass through this surface. This in order means that line of sight is blocked by the specific surface unless bot UE and gNB are on the same side of it. Is this the case or not is thus the last condition that must be checked.

For every new location of the UE we assume initially that there is LOS between it and gNB under interest. Then steps described above are repeated for every surface of every building. If intersection with any surface is detected then we already know that there is no LOS and further checking is not necessary anymore. If all surfaces are handled and no intersection was detected then this means that there must be direct visibility between UE and gNB.

There is actually no need to check against every surface in the city. As the coordinates of UE are available then only buildings at its close vicinity should be included. As the coordinates of each building are also included in metadata then such preselection is easy task to perform.

In order to test the described algorithm a test-drive was performed in and around the campus of Tallinn University of Technology (TUT). Local mobile network provider Telia Estonia has 5G base station with three radio units (RU) at the top of the tallest campus building. Rohde & Schwarz TSME6 drive and walk test scanner was used to collect all necessary measurements.

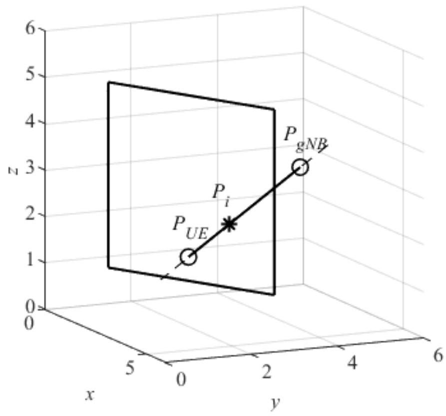

Figure 2. Geometry of LOS determination

Both the trajectory of the test drive and the known locations of the RUs are shown in Figure 3.

Measurement at each point of the trajectory contained GPS coordinates of measurement location, time of the measurement and the values of 5G signal parameters. Coordinates and the height of the receiver was used for current work, 5G signal parameters were used for classification algorithm. Later is out of the scope of the current paper.

Results of the classification can be verified visually on the site, during the measurements. If the measurement testbed is on the ground then visual LOS from its location to the gNB antenna can be confirmed by the experimenter.

III. RESULTS

Proposed method was tested with data collected during measurements with the portable 5G test scanner. Classification results were verified visually to be correct indeed. This automatically labelled data was later used to train different ML algorithms for automatic LOS/NLOS classification in case when low-end 5G device has no built-in GNSS device. Those results were quite promising but are not in the scope of the current paper.

Figure 3 gives an example of used method. Three red circles are the locations of the RUs at the rooftop of the seven-story building. Red dots are the locations of the measurements that were made without the LOS between setup and any of the antennas. Blue dots are the locations of those measurements where there was direct LOS from antenna to testbed. Automatic classification results are in good correspondence with the visually acquired ones.

IV. CONCLUSIONS

It is shown that we can indeed automatically label the 5G signal measurements either to be made under the LOS or NLOS conditions. For the labelling itself a high-end UE with built in GNSS receiver is required. Gathered and labelled data in other hand, can later be used to increase the mobile positioning accuracy of the low-end devices like IoT sensors or other similar.

Number of the buildings, even in relatively small area, can be quite high in urban environment. Also, the number of surfaces of every building can also be large, especially when building has complex size and represented level of detail is large. All this can lead to significant processing time for classification method. Fortunately, this classification must be carried out only once and it must not be in real time. Using 3D models with lower LOD level can speed up the process. The pre-filtering of buildings to be used can be also be improved in the future work.

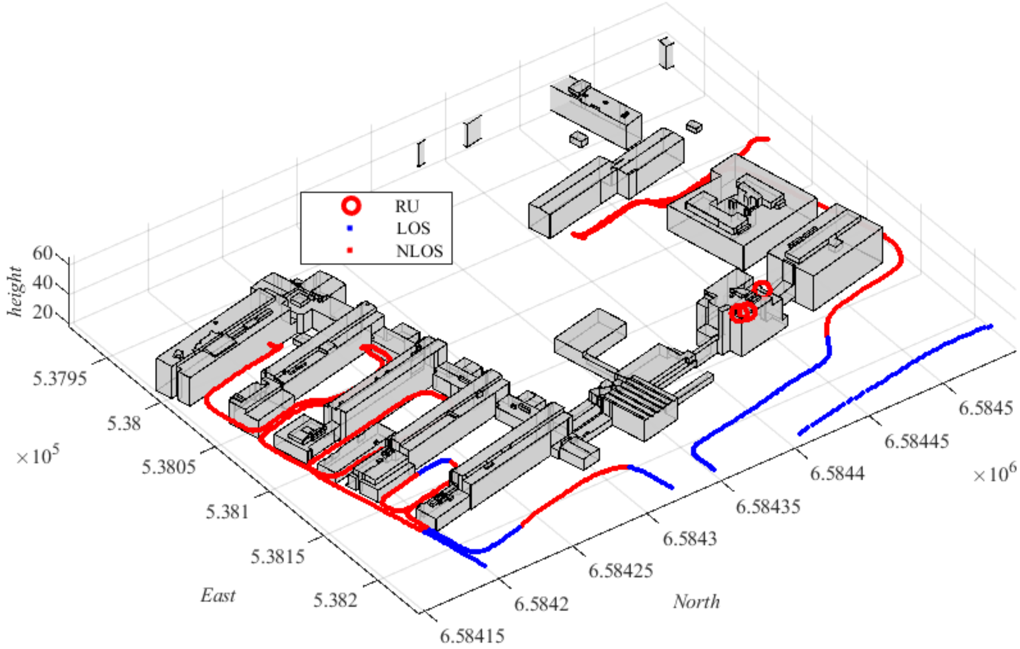

Figure 3. Example of the automatic LOS/NLOS classification

Current implementation considers only buildings but the environment can also contain a significant amount of vegetation. There are also publicly available data about the later. How to implement the Lidar based height and density measurements of vegetation is another interesting topic for the future improvements.

Available 3D models of buildings have many other possible applications in the field of mobile communications and radio signal-based positioning. For example, the models can be used with deterministic propagation models, such like ray-tracing, to calculate signal strengths and mobile coverage. Those calculations can be used for fingerprint-based mobile positioning, instead or alongside with the real measured signal values.

3D models with higher LOD can include also the openings in the buildings (windows and doors), internal structure, furniture, used building materials etc. Thus, in the future the accuracy of predictions based on those models are expected to increase.

V. ACKNOWLEDGEMENT

This work was supported in part from the European Union's Horizon 2020 research and innovation program, within 5G ROUTES project, under grant agreement No 951867, and part by the Tallinn University of Technology Development Program 2016-2022, code 2014-2020.4.01.16-0032.

REFERENCES

[1] M. Silventoinen, T. Rantalainen, „Mobile station emergency locating in GSM,“ *IEEE International Conference on Personal Wireless Communication*, New Delhi, 1996.

[2] 3GPP, „TS 22.261, Service requirements for the 5G system; Stage 1, Release 18, V18.0.0,“ September 2020.

[3] 3GPP, „TR 38.855, Study on NR positioning support, Release 16, V16.0.0," March 2019.

[4] I. Müürsepp, M. Kulmar, O. Elghary, M. A. Alam, T. Chen, S. Horsmanheimo and J. Scholliers, „Performance Evaluation of 5G-NR Positioning Accuracy Using Time Difference of Arrival Method," *IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, Athens, 2021.

[5] P.-C. Chen, „A non-line-of-sight error mitigation algorithm in location estimation," *WCNC*, New Orleans, 1999.

[6] L. Cong and W. Zhuang, „Non-line-of-sight error mitigation in TDOA mobile location," *GLOBECOM'01*, San Antonio, 2001.

[7] A. Xhafa, J. A. del Peral-Sosado, J. A. Lopez-Salcedo and G. Seco-Granados, „Evaluation of 5G Positioning Performance Based on UTDoA, AoA and Base-Station Selective Exclusion," *Sensors 22*, pp. 1-18, 2021.

[8] S. Venkatraman, J. Caffery, Jr and H.-R. You, „A Novel ToA Location Algorithm Using LoS Range Estimation for NLoS Environments," *IEEE Transactions on Vehicular Technology*, kd. 53, nr 5, pp. 1515-1524, 2004.

[9] V. Djaja-Josko and M. Kolakowski, „A new map based method for NLOS mitigation in the UWB indoor localization system," *TELFOR*, Belgrade, 2017.

[10] S. O. Al-Jazzar, J. Caffery Jr and H.-R. You, „A scattering model based approach to NLOS mitigation in TOA location systems," *VTC*, Birmingham, 2002.

[11] M. Kolakowski and J. Modelski, „First path component power based NLOS mitigation in UWB positioning system," *TELFOR*, Belgrade, 2017.

[12] M. A. Landolsi, A. F. Almutairi and K. M. A., „LOS/NLOS channel identification for improved localization in wireless ultra-wideband networks," *Telecommunication Systems*, kd. 72, pp. 441-456, 2019.

[13] S. Sosnin, A. Lomayev and A. Khoryev, „NLOS Links Detection Algorithm for Improved 5G NR Indoor Positioning," *BlackSeaCom*, Bucharest, 2021.

[14] K. K. Cwalina, P. Rajchowski, O. Blaszkiewicz, A. Olejniczak and J. Sadowski, „Deep Learning-Based LOS and NLOS Identification in Wireless Body Area Networks," *Sensors 19*, pp. 1-11, 2019.

[15] J. Park, S. Nam, H. Choi, Y. Ko and Y.-B. Ko, „Improving Deep Learning-Based UWB LOS/NLOS Identification with Transfer Learning: An Empirical Approach," *Electronics*, kd. 9, pp. 1-13, 2020.

[16] Y. Core and Y. Lostanlen, „Three-Dimensional Urban EM Wave Propagation Model for Radio Network Planning and Optimization Over Large Areas," *IEEE Transactions on Vehicular Technology*, kd. 58, nr 7, pp. 3112 - 3123, 2009.

[17] Z. Yun, S. Y. Lim and M. F. Iskander, „Radio propagation modeling in complex environments for wireless communications," *COMCAS*, Tel Aviv, 2009.

[18] A. Bounceur et.al., „CupCarbon: A new platform for the design, simulation and 2D/3D visualization of radio propagation and interferences in IoT networks," *CCNC*, Las Vegas, 2018.

[19] G. Lv, H. Zhao and J. Hu, „Real Time Correction of Multipath Error in Satellite Positioning using FPGA-Accelerated Ray Tracing," *APS/URSI*, Singapore, 2021.

[20] Estonian Land Board, „Building 3D model data," 27 October 2022. Available: https://geoportaal.maaamet.ee/eng/Download-3D-data-p837.html.

[21] 3GPP, „TR 38.857, Study on NR Positioning Enhancements, Release 17, V17.0.0," March 2021.