Estimating Coverage and Capacity of High Frequency mobile networks in Ultradense Urban Areas

Gabriele Gemmi[a,*], Michele Segata[b], Leonardo Maccari[a]

[a] *Ca' Foscari University of Venice, Italy*
[b] *University of Trento, Italy*

Abstract

High frequency communications (mmWave and TeraHz) in urban areas require a higher density of base stations compared to pre-5G mobile networks, but open the way to a quantum leap in increased throughput and reduced latency. However, we currently have no indication of how much we need to densify the deployment, and on the trade-off between the density of base stations and the performance improvement. This paper studies the problem of base stations placement to guarantee coverage to vehicles and pedestrians in urban areas when using high frequency communications. Our novel methodology takes advantage of vehicular traffic simulations and precise urban maps to generate a realistic demand model for vehicles and pedestrians in urban areas. We use a bounded error heuristic to find the maximal coverage that can be achieved with a given density of base stations, primarily using line-of-sight communications. We implemented the heuristic using Cuda libraries on Nvidia GPUs and evaluated the coverage in an urban area in the city of Luxembourg, for which vehicular traffic patterns are available. We focus on coverage and capacity analysis for the mmWave frequency, but the results are easily extended to TeraHz communications.

Our results are the first to show that a reasonably low density (15 base stations per km^2) is sufficient to provide coverage for vehicles in urban environments. However, optimizing on vehicles or on pedestrians are competing objectives: the operator needs to choose which one to target based on its business model when designing the network infrastructure. Our algorithms, code and open data can be used to perform this task and reproduce our results in different settings.

Keywords: vehicular communication, 5G, mmWave, gNB placement

*This work was partially supported by NGIAtlantic.eu project within the EU Horizon 2020 programme under Grant No. 871582.

*Corresponding author

Email addresses: gabriele.gemmi@unive.it (Gabriele Gemmi), michele.segata@unitn.it (Michele Segata), leonardo.maccari@unive.it (Leonardo Maccari)

1. Introduction

In order to meet the increasing demand for mobile connectivity, the next-generation access networks (XGs) will rely on the use of very high frequencies (mmWave and TeraHerz) and on the densification of the existing access network, by increasing up to 10 times the number of deployed base stations [1]. These new communication technologies are much more susceptible to obstruction and they need Line-of-Sight (LoS) to function reliably. For these reasons, the placement strategy of the base stations is crucial, and as already shown in previous works, an optimal choice of such locations can lead to substantial savings for network operators [2].

One of the future applications enabled by XGs is the use of Cooperative Autonomous Vehicles (CAVs). To be really effective, cooperative driving will not only require vehicles to exchange basic data such as position, speed, heading, etc., but raw sensor data as well. This will permit vehicles to implement Cooperative Perception (CP) [3], i.e., to be able to construct a view of the surrounding environment that goes beyond the field of view of their sensors. Sharing raw sensor data rather than pre-processed data enables vehicles to take decisions on their own or to come up with a consensus on how to classify certain objects, which can lead to safer and more efficient driving (see the *boar and the hare* example [4]).

While this vision is technologically stimulating, we still miss a reliable estimation of the effort needed to deploy a new, denser infrastructure that will provide mostly LoS coverage to mobile terminals in urban areas. Manufacturers suggest that 5G will require more than 100 base stations - or next Generation NodeBs (gNBs) using the 5G terminology - per km^2, up from the roughly 10 per km^2 in LTE [1], and even more in 6G. Some works consider densities even higher than a hundred gNB/km^2 [5].

Without algorithms, data, and code to tackle this problem, the research community can not design protocols and applications with credible results. This paper contributes to provide this missing link. We use a data-driven approach to find an optimal placement for gNBs by taking into account a demand model built for vehicles (based on street traffic patterns) and for pedestrians (isolating sidewalks and pedestrian areas on detailed maps). Given a demand model, we devise a new heuristic that exploits it to find the optimal location of the gNBs. We take advantage of open geographical data, specifically OpenStreetMap (OSM) vectorial maps and Digital Surface Model (DSM) to evaluate different gNB placements on real-world data. While the analyses have been conducted only in the city of Luxembourg, the availability of open-data together with the source code we release will enable anyone else to reproduce the analyses in different areas. [1]

We advance the state of the art with the following findings:

- A "*reasonably low*" density (15 gNBs/km^2) can be used to provide 95%

[1] `https://github.com/UniVe-NeDS-Lab/TrueBS/tree/vehicular_mod`

coverage to vehicles, while pedestrian areas require a higher density (35 gNBs/km^2);

- Non Line-of-Sight (NLoS) links can be useful to fill the gap but the performance penalty is extremely high;
- Providing connectivity to vehicles and pedestrians are two competing objectives, as a small part of the walkable area is on the streets. If the operator needs to guarantee service continuity also to pedestrians, it needs to adopt a dedicated covering strategy.

This work extends and improves the initial results obtained in a previous conference paper [6] in which we introduced the methodology and measured the availability of LoS from gNBs to ground points. This work uses the same approach but focuses on the estimation of the link capacity. We use realistic path loss models suggested by ETSI to provide much more insightful results compared to our previous ones.

2. State of the Art

While the placement of base stations is a widely investigated matter [7], the LoS requirements introduced by the newer communication technologies have reignited the attention on the subject, with several works taking advantage of similar techniques [8, 9, 10, 11, 2]. However, to the best of our knowledge, no other study is focused on investigating different placement strategies to optimize mobile coverage for vehicles using realistic traffic data, and for pedestrians using detailed city maps. This is because vehicular networks are a very specific application for which research on field is extremely expensive and the available simulation tools fall short in the analysis of a three-dimensional scenario, that is fundamental to evaluate LoS in XGs. The most popular open source simulator used to simulate vehicular networks is Veins, powered by the Omnet++ discrete event simulator [12]. Veins uses a flat 2D map, so it does allow to perform LoS estimation but it is mostly useful for vehicle-to-vehicle communications, assuming that vehicles in communication range are at roughly the same elevation. For some other applications however, the lack of a 3D approach has been recognized as a limitation [13]. There have been some attempts to improve Veins to introduce the third dimension but this is limited to the use of loss models that are more accurate in considering the ground elevation (like the n-ray ground model) and not necessarily the shape of buildings [14]. Only very recently Veins has been extended to introduce simple prisms as three dimensional building shapes [15] with a limited computational penalty. In comparison, our approach is GPU-based and can exploit data of arbitrary accuracy. Finally, the CARLA simulator is a recent instrument that heavily relies on 3D models [16], however the goal of CARLA is not to model realistic networks of vehicles, but to improve their intelligence with the reconstruction of 3D objects. The most similar research, from Jaquet et al. [17] is focused on enhancing vehicular networks by taking advantage of unmanned aerial vehicles.

3. Problem Formulation and Solution

We consider a 3D shape of an urban area and a set Λ of points in the ground that can be potentially covered with a LoS connection from a gNB. Each point corresponds to an (x, y, z) triplet, in which the (x, y) coordinates are quantized using one point per squared meter. Points are selected to be outside any building shape and only in public areas (streets, roundabouts, street parking, sidewalks) and not in private areas. Each point in Λ is uniquely identified by the (x, y) couple, as we set the z value to a height of 1.5 m, as such, with a little abuse of notation we may use $(x, y) \in \Lambda$ when it simplifies the description.

The problem we tackle can be summarized in three steps described in the next two sections:

1. For each point in Λ define a weight, the higher the weight the higher the probability the point will be covered. The matrix that associates (x, y) to a generic weight is called $\boldsymbol{\tau}$;
2. Identify the set $\mathcal{P}$ of points in space in which a gNB could potentially be placed. Each position $p_i \in \mathcal{P}$ is defined by an (x, y, z) triplet, we consider only points on the facades of buildings;
3. Find an algorithm that chooses the minimal number of gNBs so that the coverage is maximized according to some metric.

Step one is a novel contribution of this paper. In a previous work [2], all points in the city had the same importance, whereas here we weight locations considering traffic demand and pedestrian areas, choosing gNB locations depending on that. The solution to the second step comes from a previous publication in which we introduced the problem of coverage as a variation of the classical maximum subset coverage problem [2], while the third step modifies the solution proposed therein to take into account the weights introduced in step one. In the remainder of this section, we will formalize point 3 (the gNB placement problem) and provide an algorithmic solution for a generic weights matrix $\boldsymbol{\tau}$. Then, in Sect. 4 we detail the different strategies to obtain 2 realistic weights matrices for vehicles and for pedestrians.

For convenience, we list all the mathematical symbols in Tab. 1, some of which are also represented graphically in Fig. 1. In what follows we use the $|\cdot|$ operator on sets and matrices. On sets it counts the number of elements, while on matrices it refers to the 1-norm.

Limitations. It is important to note that all our results are to be interpreted as a base-line on which to produce fine-grained further results. In this sense, we recognize that our conclusions are not directly applicable to other cities, because at the time of writing the Luxembourg scenario is the only one for which we own all the necessary data, however the paper methodology is fully repeatable. Future works will extend our approach using data coming from other cities and more complex ways of estimating the performance of the links, such as estimating the blockage due to obstacles that are not buildings or using more elaborate ray-tracers that take into account also reflections and diffraction.

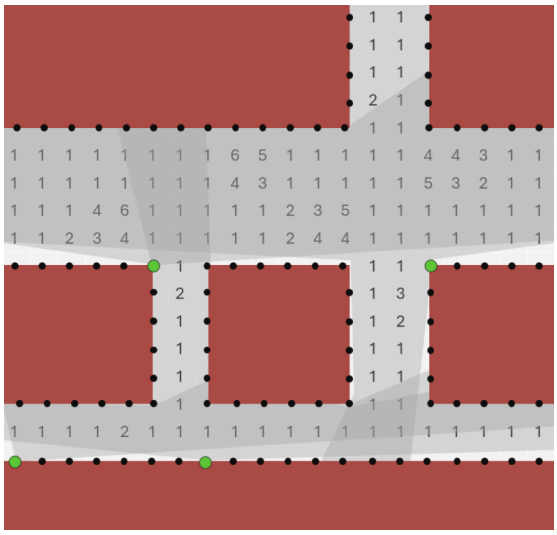


Figure 1: Graphical depiction of the weighted placement problem: The matrix of weight ($\boldsymbol{\tau}$) associated with each point is represented by the numbers in the background; the set of Buildings ($\mathcal{B}$) is represented by the rectangles in brick red; the set of candidate locations ($\mathcal{P}$) is represented by the small black points on the edge of the buildings; the locations selected by the algorithm are represented by the larger dots colored in green, their viewsheds $\boldsymbol{\sigma}^i$ are represented by the grey shadows projected on the streets.

3.1. gNB Placement

We briefly recall what was proposed in Gemmi et al. [2] that is at the base of this work. Let $\mathcal{B} = \{b_i\}$ be the set of buildings extracted from the OSM dataset. Let also $\phi(b_i)$ be a function that extracts a set of coordinates that compose the perimeter of the building b_i, with points spaced on average one meter away from each other, placed 1 m below the height of the roof. The elevation of the roof is derived using open data repositories that provide a surface model (a DSM file). We can then define the set of candidate locations $\mathcal{P}$ as:

$$\mathcal{P} = \bigcup_{b_i \in \mathcal{B}} \phi(b_i) \tag{1}$$

Once the set of candidate locations is determined, we need to evaluate the coverage from each of them. In order to do so, we take advantage of a viewshed algorithm implemented using the CUDA library on NVidia GPUs [18]. This algorithm, applied on a highly precise DSM computes the presence of LoS from the candidate location p_i to each point in Λ given a maximum distance ($d_{max} = 300\,\text{m}$) from p_i. We consider this value of distance as it has been used as an upper bound in other works [19]. Let $\boldsymbol{\sigma}^i = \Upsilon(\Lambda, p_i)$ be a 2-dimensional binary matrix that associates each point in Λ to a non-zero value if there is LoS from the point p_i to the point of coordinate $(x, y, z) \in \Lambda$. Υ corresponds to the application of the viewshed algorithm from the point p_i over the set of points Λ. We call $\boldsymbol{\sigma}^i$ the *viewshed* matrix from point p_i. If we apply Υ to all the points

Table 1: List of symbols

Symbol	Meaning
Λ	Set of ground points potentially to be covered
$\boldsymbol{\tau}$	Matrix of weights associated to points in Λ
$\mathcal{P}$	Set of all candidate locations for gNBs
$\mathcal{B} = \{b_i\}$	Set of buildings extracted from the OSM dataset
$\phi(b_i)$	Set of coordinates that compose the perimeter of the building b_i
$\boldsymbol{\sigma}^i = \Upsilon(\Lambda, p_i)$	Υ is a function returning a binary matrix $\boldsymbol{\sigma}^i$ indicating if there is LoS between a point p_i and a set of points Λ
Ω	Set of all viewsheds corresponding to all possible gNB positions
$\Omega^\star$	Quasi-optimal set of viewshed computed by the optimization algorithm
$\boldsymbol{\pi}, \boldsymbol{\gamma}$	Weight matrices for the pedestrian and vehicular strategies, respectively
$\Lambda_{\boldsymbol{\pi}}, \Lambda_{\boldsymbol{\gamma}}$	Set of ground points to be covered by the pedestrian and vehicular strategies, respectively. Named as walkable and drivable areas in the text
λ	Desired gNB density, in gNBs per squared kilometer
S	Size of the area considered in the analysis [km^2]
$k = \lambda S$	Maximum number of gNBs to deploy
$\lvert\Gamma\rvert$, Γ being a set	Set size operator, counts the number of elements in the set
$\lvert\boldsymbol{\tau}\rvert$, $\boldsymbol{\tau} \in \mathcal{R}^{m,n}$	1-norm operator over a matrix (sum of absolute values), i.e., $\sum_{i=1}^{m}\sum_{j=1}^{n}\lvert\boldsymbol{\tau}_{i,j}\rvert$

in $\mathcal{P}$ we obtain a collection of matrices that represent all the possible viewsheds from all the potential positions of gNBs, as in Eq. (2):

$$\Omega = \bigcup_{p_i \in \mathcal{P}} \Upsilon(\Lambda, p_i) \tag{2}$$

We then obtain a collection of matrices $\Omega = \{\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \ldots, \boldsymbol{\sigma}^m\}$ in which $\boldsymbol{\sigma}^i_{x,y} = 1$ means that a terminal in position (x, y) has LoS with a gNB placed in the point p_i. Remember that we extend the classic definition of LoS with the additional constraint of being within a maximum distance d_{max} and that to each point of coordinate (x, y) we attribute an elevation z given by the DSM file, plus 1.5m.

3.2. Quasi-optimal gNB placement

Given a parameter k indicating the maximum number of gNBs an operator is willing to deploy, we want to find a subset $\Omega^\star \subseteq \Omega$ whose size is lower than k

that maximizes the coverage $\left|\bigvee_{\boldsymbol{\sigma}^i \in \Omega^\star} \boldsymbol{\sigma}^i\right|$, where $\bigvee$ is the OR operator between binary matrices. In order to take into account the traffic patterns, consider a generic non-negative weight matrix $\boldsymbol{\tau}$ with the same shape of $\boldsymbol{\sigma}^i$. We can formulate the maximization objective as

$$\max_{\Omega^\star} \left| \boldsymbol{\tau} \odot \left(\bigvee_{\boldsymbol{\sigma}^i \in \Omega^\star} \boldsymbol{\sigma}^i \right) \right| \quad \text{with } |\Omega^\star| \leq k, \tag{3}$$

where $\odot$ is the Hadamard product (the element-by-element multiplication between two matrices). This will lead to a choice of the optimal k viewsheds in $\Omega^\star$ to cover the roads with the highest traffic. The problem is a so-called weighted maximum coverage problem.

Note that if we call $\mathbb{1}$ the matrix made of all one elements, and we set $\boldsymbol{\tau} = \mathbb{1}$ then the problem converges to the classical unweighted maximum coverage problem, in which we try to cover the largest portion of the points in Λ treating all of them equally.

3.3. Heuristic solution

Since the above-described coverage problem is NP-Hard in our past work we relied on a polynomial greedy heuristic with bounded error to efficiently find a quasi-optimal solution [2].

Here we modify the greedy heuristic as described in detail in Algorithm 1 to take into account the weight matrix. The heuristic proceeds as follows: we start by defining a coverage matrix $\mathbf{C}$ of the same size of $\boldsymbol{\tau}$, initialized with zeros (Line 2). Each iteration of the loop in Line 4 will choose the position of one gNB. For each candidate location p_i and the corresponding viewshed $\boldsymbol{\sigma}^i$ we derive the so-far uncovered elements $\bar{\boldsymbol{C}}$ as the negation of the coverage matrix (Line 7). The idea is, at each step, to progressively add the gNB that provides the largest additional coverage with respect to the already covered area. We define $\mathbf{C}^\star$ that represents the so-far uncovered elements that would be covered by the candidate location, with their weight given by $\boldsymbol{\tau}$ (Line 8). Note that *bool*() is a function that makes an integer matrix a boolean one, $\neg$ is the boolean NOT operand. We then provide a score for p_i as the norm-1 of the coverage matrix (in Line 9).

Then, the element with the maximal ranking that is not already in $\Omega^\star$ ($\boldsymbol{\sigma}^j \notin \Omega^\star$) is chosen and the corresponding values of the viewshed matrix are added to $\mathbf{C}$ (Line 13). Note that this makes $\mathbf{C}$ a non-boolean matrix. Finally, the viewshed with the maximal ranking $\boldsymbol{\sigma}^\star$ is added to the set of optimal viewsheds (Line 14). The loop is repeated till the number of desired locations is reached. The operation at line Line 8 has complexity $|\Lambda|$, and is repeated at most $k \times |\mathcal{P}|$ times, so the overall complexity is $O(k|\mathcal{P}||\Lambda|)$.

This algorithm is referred to as $\Gamma(\Omega, k, \boldsymbol{\tau})$ and in the next section we use it with two weights matrices $\boldsymbol{\pi}$ or $\boldsymbol{\gamma}$, producing two sets of quasi-optimal viewsheds optimized for pedestrians or vehicles respectively.

Algorithm 1 Greedy algorithm for the weighted maximum coverage problem.

Input: Ω (Set of viewsheds), k (number of gNBs), $\boldsymbol{\tau}$(weighted traffic matrix)
Output: $\Omega^\star$ (Set of the viewsheds from optimal locations)

procedure $\Gamma(\Omega, k, \boldsymbol{\tau})$
  $\mathbf{C} = \mathbf{0}$
  $\Omega^\star = \{\}$
  for $i \leftarrow 1$ to k **do**
    $h^\star = -\infty$
    for $\boldsymbol{\sigma}^j \in \Omega$ **do**
      $\bar{\boldsymbol{C}} = \neg bool(\mathbf{C})$
      $\mathbf{C}^\star = \bar{\boldsymbol{C}} \odot \boldsymbol{\sigma}^j \odot \boldsymbol{\tau}$
      $h_j = |\mathbf{C}^\star|$
      if $h_j > h^\star$ and $\boldsymbol{\sigma}^j \notin \Omega^\star$ **then**
        $\boldsymbol{\sigma}^{i^\star} = \boldsymbol{\sigma}^j$
        $h^\star = h_j$
    $\mathbf{C} = \mathbf{C} + \boldsymbol{\sigma}^{i^\star}$
    $\Omega^\star = \Omega^\star \cup \{\boldsymbol{\sigma}^{i^\star}\}$
  return $\Omega^\star$

4. A Demand Model for Vehicles and Pedestrians

Obtaining realistic traffic data, for both pedestrians and vehicles, is always a challenging task, as data collected by cities is rarely released to the public. One possibility, which is the one we consider in this work, is to generate traffic data using microscopic traffic simulators and realistic scenarios. We use the urban traffic simulator SUMO [20] to generate realistic mobility traces of the city of Luxembourg. In particular, we make use of the Luxembourg SUMO Traffic (LuST) scenario [21], a publicly available scenario generated from traffic data provided by the Luxembourg government which includes both public and private transportation over a period of 24 h. Unfortunately, since the model does not include pedestrian mobility, we had to resort to a different approach to model it. In the next sections, we detail both models.

The reason for choosing Luxembourg lies in the fact that, in addition to the demand model, our solution requires buildings height data, which is available for this city. We obtain such information from 3D lidar data available from the Luxemburg open-data platform [2], enabling us to associate a precise height to every building extracted from the OSM map. We considered including other cities in our analysis, but finding both the demand model and the lidar data is unfortunately very unlikely. As an example, we find lidar data for the city of Turin, but the SUMO scenario [22] produces traffic deadlocks with consequent

[2]The lidar data is available at: `https://data.public.lu/fr/datasets/lidar-2019-modele-numerique-de-terrain-mnt/`

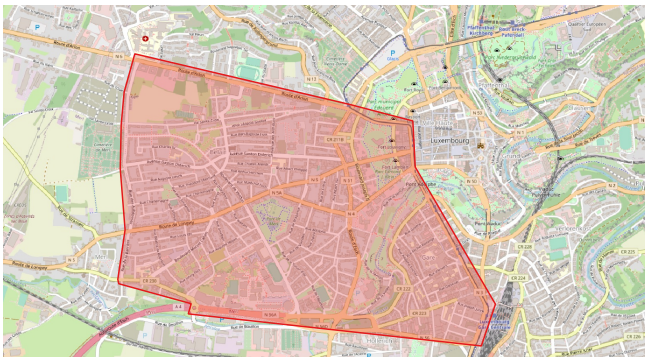


Figure 2: Area of the city of Luxembourg over which traces are collected.

"teleportations"[3], which would make the traces unrealistic. On the contrary, there exists a very realistic SUMO scenario of the Principality of Monaco [23], but no 3D lidar data.

4.1. Vehicular Model

To obtain traffic traces, we run the scenario over the full 24 h for a total of 286 215 vehicles moving on the streets. The simulation sampling time is set to 1 s and, at the same frequency, we log the positions of the vehicles in the area of the city shown in Fig. 2, corresponding to an area S of roughly 4 km^2. We collect traces using GPS (latitude/longitude) coordinates and then convert them to a `.gpx` file for later processing. We map each point in a trace to a cell in the discretized space given by Λ. We obtain a matrix $\boldsymbol{\gamma}$ with the same shape as $\boldsymbol{\tau}$, where $\boldsymbol{\gamma}_{x,y} = n$ means that n vehicles have passed in cell (x, y) during the whole simulation[4].

Fig. 3 shows the empirical pdf of the values of the cells with non zero value, binned with bins of size 0.125 passages/minute, rescaled to the number of passages per minute for readability. It can be seen that the majority of the cells have less than one passage per minute, with the 95th percentile roughly at 0.55. The distribution is pretty skewed, with about 5 orders of magnitude between the largest and the lowest frequency.

We call $\Lambda_{\boldsymbol{\gamma}}$ the set of points $(x, y, z) \in \Lambda$ in which $\boldsymbol{\gamma}_{x,y} \neq 0$, and we refer to it as the *drivable* area.

[3]Teleportation is used in SUMO to resolve deadlocks or collisions. If a vehicle is involved in a collision or it is stopped for a time longer than a threshold, SUMO moves it to the next edge in its path.

[4]In our previous work [6] due to memory limitations of our GPU we had to re-scale the frequency values in $\boldsymbol{\tau}$ to be within the allowed range of one byte: $[0, 255]$. In this work we resort to a GPU with a larger memory, enabling us to work with 16 bit integers and get rid of such re-scaling.

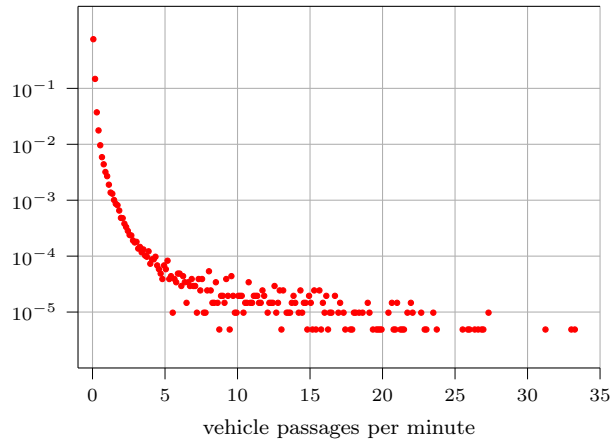


Figure 3: Empirical p.d.f. of the vehicles passages per minute per cell.

4.2. Pedestrian Model

Due to the lack of a pedestrian mobility model in the LuST scenario, we had to manually define certain public areas where the pedestrians would transit. We decided to focus on sidewalks and public open-air areas such as parks, gardens, squares, etc.

To model the sidewalks we started from the lines characterizing the OSM roads, which were used to center two symmetrical 2-meter wide sidewalks with a distance from the center of the road depending on the type of road. To model the public open-air areas, on the other hand, we relied on the public land use database from OSM, where we selected the public areas with one of the following types: `cemetery`, `forest`, `grass`, `heath`, `meadow`, `orchard`, `park`, `recreation_ground`. Again, we mapped those points to a cell in the discretized space given by Λ and obtained a matrix $\boldsymbol{\pi}$ with the same shape as $\boldsymbol{\tau}$, where $\boldsymbol{\pi}_{x,y} = 1$ when the element corresponds to one of those areas and 0 otherwise. Fig. 4 shows both sidewalks and public areas in green, the roads dedicated to vehicles in yellow, and buildings in gray. We call $\Lambda_{\boldsymbol{\pi}}$ the set of points (x, y, z) in which $\boldsymbol{\pi}_{x,y} \neq 0$, and we refer to it as the *walkable* area.

Note that in a realistic deployment our algorithm allows to be manually tailored by an operator, according to local needs and strategies. This could be achieved by manually modifying the demand model: the operator could increase (or decrease) the weights to improve the coverage of a certain area that is more (ol less) important, based on its business or operating needs. Once the weights are manually modified, the optimization algorithm does not need modifications.

5. Experimental Setup and Metrics

We consider two different settings, one in which we optimize the coverage for the vehicular traffic ($\boldsymbol{\tau} = \boldsymbol{\gamma}$), and another one in which we optimize for the

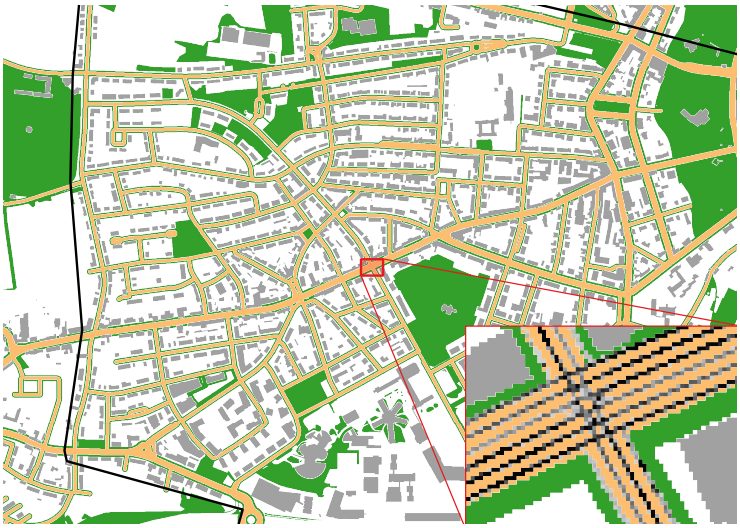


Figure 4: Detailed view of a portion of the area considered in the analysis. Yellow highlights roads, green indicates public areas for pedestrians (sidewalks and parks), grey indicates buildings, and white indicates private areas. The zoomed portion shows the vehicular traces with a grayscale representing the number of passages per point (darker means higher).

Parameter	Value
Area size	$3.98\,\mathrm{km}^2$
Carrier frequency	28 GHz
Bandwidth (B)	400 MHz
Thermal Noise (T=300 K)	−87.8 dBm
Noise Figure (N_t)	5 dB
3GPP Channel Model	ETSI TR 38.901 Urban Micro
Reception gain	3 dBi
MIMO layers (μ)	2
Transmission power	30 dBm
Transmission gain	10 dBi
Maximum distance for LoS links	300 m

Table 2: Simulation Parameters

areas where pedestrians might be located ($\boldsymbol{\tau} = \boldsymbol{\pi}$). We apply Algorithm 1 to compute the optimal locations for the gNBs, increasing their number k. We consider a density λ of gNBs per squared km going from 5 to 45 at steps of 5, and we set $k = \lambda S$. Remember that S indicates the size of the area we consider.

We obtain two solutions for the coverage:

$$\Omega^{\star}_{\lambda,\boldsymbol{\gamma}} = \Gamma(\Omega, \lambda S, \boldsymbol{\gamma}) \tag{4}$$

$$\Omega^{\star}_{\lambda,\boldsymbol{\pi}} = \Gamma(\Omega, \lambda S, \boldsymbol{\pi}) \tag{5}$$

that we aggregate with the OR operator to obtain a full coverage matrix:

$$\Phi_{\lambda,\boldsymbol{\gamma}} = \bigvee_{\boldsymbol{\sigma}^j \in \Omega^{\star}_{\lambda,\boldsymbol{\gamma}}} \boldsymbol{\sigma}^j; \quad \Phi_{\lambda,\boldsymbol{\pi}} = \bigvee_{\boldsymbol{\sigma}^j \in \Omega^{\star}_{\lambda,\boldsymbol{\pi}}} \boldsymbol{\sigma}^j \tag{6}$$

In brief, the elements in Φ indicate whether a point (x, y) is covered by at least one gNB.

5.1. Coverage Metrics

We use four metrics to compare the results in terms of coverage (again, $\odot$ is the Hadamard product and $|\cdot|$ the norm-1). The first two refer to the drivable area and are:

$$Dcov_{\mathrm{ve}}(\lambda) = \frac{|\boldsymbol{\gamma} \odot \Phi_{\lambda,\boldsymbol{\gamma}}|}{|\boldsymbol{\gamma}|}; \quad Dcov_{\mathrm{pe}}(\lambda) = \frac{|\boldsymbol{\gamma} \odot \Phi_{\lambda,\boldsymbol{\pi}}|}{|\boldsymbol{\gamma}|} \tag{7}$$

The metrics in Eq. (7) tell how good the coverage of drivable areas is when we optimize for vehicles ($Dcov_{\mathrm{ve}}$) or when we optimize for pedestrians ($Dcov_{\mathrm{pe}}$). $Dcov_{\mathrm{pe}}$, in practical terms, tells us what happens if we try to optimize the coverage for pedestrians but we measure the results only on the points where the vehicles pass (with their multiplicity). Of course, we expect $Dcov_{\mathrm{ve}}(\lambda)$ to be larger than $Dcov_{\mathrm{pe}}(\lambda)$, yet we are interested in the difference.

We use two more metrics to evaluate the complementary set-up:

$$Wcov_{\mathrm{pe}}(\lambda) = \frac{|\boldsymbol{\pi} \odot \Phi_{\lambda,\boldsymbol{\pi}}|}{|\boldsymbol{\pi}|}; \quad Wcov_{\mathrm{ve}}(\lambda) = \frac{|\boldsymbol{\pi} \odot \Phi_{\lambda,\boldsymbol{\gamma}}|}{|\boldsymbol{\pi}|} \tag{8}$$

Both metrics express how good the coverage of walkable areas is, in the first case when we optimize for pedestrians ($Wcov_{\mathrm{pe}}$) while, in the second, when we optimize for vehicles ($Wcov_{\mathrm{ve}}$).

5.2. Channel Capacity Model

Besides the evaluation on pure coverage, we also evaluate the quality of the coverage in terms of capacity, using the Shannon channel capacity formula. To obtain the received signal strength we consider a transmission power of 30 dBm at a frequency of 28 GHz. We also assume the gNB to have an isotropic antenna with transmission gain equal to 10dB. In practice, this would be made possible by having between one and three sectorial antennas covering the unobstructed area around the gNB . To compute the pathloss between a gNB located at p_i and a point (x, y, z) in the city, we employ the Urban Micro model defined in ETSI TR 38.901 [24]. The model provides the path loss for both LoS and NLoS conditions, and we can easily distinguish between the two cases thanks to the

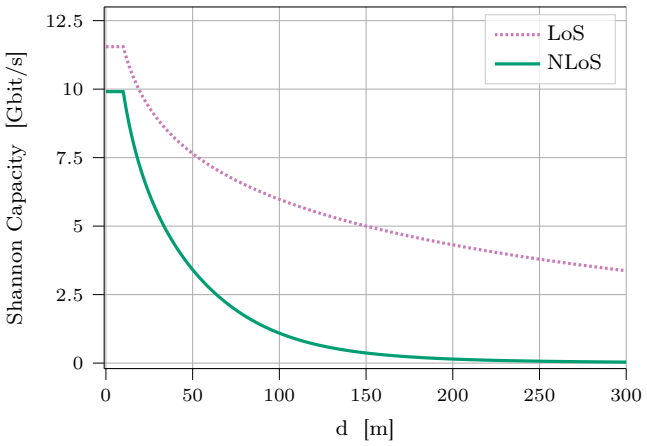


Figure 5: Shannon channel capacity as a function of the distance between two points d and the LoS/NLoS conditions.

pre-computed viewsheds. In addition, we consider a gNB to be in LoS with a point only if the distance between them is smaller than 300 m. We set the noise floor for a bandwidth $B = 400\,\text{MHz}$ to be $N_t = -87.8\,\text{dBm}$ and we compute the Signal to Noise Ratio (SNR) and the corresponding Shannon channel capacity $\boldsymbol{C}_{\text{SH}}(d)$ for $\mu = 2$ MIMO layers:

$$\boldsymbol{C}_{\text{SH}}(d) = \mu B \log_2 \left(1 + \frac{S(d)}{N_t}\right) \tag{9}$$

In Eq. (9), $S(d)$ refers to the received signal power at distance d, as per the LoS/NLoSpath loss model. Tab. 2 details all the parameters.

Fig. 5 shows a graph of the capacity as a function of the distance and LoS/NLoS. We report this known curve for three reasons. The first is that it quantifies the difference between the maximum bit-rate in the two cases. The second is that it highlights that a terminal continuously switching between LoS and NLoS would experience a constant change in network performance that would make it very hard to support any application, so a partial LoS coverage may be worse than a fully NLoS coverage. The third is that if the NLoS curve is shifted right, it can intersect the LoS one. This translates into the fact that in certain points it may be convenient to choose a gNB that is in NLoS but is physically closer to the closest one in LoS. This is more likely if the NLoS link is very short, and impossible if it is longer than roughly 50 m. It is interesting to evaluate if and how often this event happens.

5.3. Capacity Metrics

Consider a point $(x, y, z) \in \Lambda_{\boldsymbol{\gamma}}$ in the drivable area, the set $\Omega^{\star}_{\lambda,\gamma}$ of the quasi-optimal viewsheds at density λ optimized for the vehicles, and the corresponding positions of gNBs. We call $LC_{x,y,\gamma}(\lambda)$ the highest capacity we can achieve from

any gNB to (x, y, z) using only LoS links[5]. If $LC_{x,y,\gamma}(\lambda) = 0$ then there is no LoS link between (x, y, z) to any gNB. We call $NC_{x,y,\gamma}(\lambda)$ the highest capacity when using only NLoS links, which is always larger than zero (there are always NLoS links to any (x, y, z)). The same metrics are also defined for walkable areas using the $\boldsymbol{\pi}$ pedix. We collect all the capacities in four sets, divided by LoS/NLoS:

$$Dcap^L_\lambda = \{LC_{x,y,\boldsymbol{\gamma}}(\lambda) \; \forall \; (x, y, z) \in \Lambda_{\boldsymbol{\gamma}} \mid LC_{x,y,\boldsymbol{\gamma}}(\lambda) > NC_{x,y,\boldsymbol{\gamma}}(\lambda)\} \tag{10}$$

$$Dcap^N_\lambda = \{NC_{x,y,\boldsymbol{\gamma}}(\lambda) \; \forall \; (x, y, z) \in \Lambda_{\boldsymbol{\gamma}} \mid NC_{x,y,\boldsymbol{\gamma}}(\lambda) > LC_{x,y,\boldsymbol{\gamma}}(\lambda)\} \tag{11}$$

$$Wcap^L_\lambda = \{LC_{x,y,\boldsymbol{\pi}}(\lambda) \; \forall \; (x, y, z) \in \Lambda_{\boldsymbol{\pi}} \mid LC_{x,y,\boldsymbol{\pi}}(\lambda) > NC_{x,y,\boldsymbol{\pi}}(\lambda)\} \tag{12}$$

$$Wcap^N_\lambda = \{NC_{x,y,\boldsymbol{\pi}}(\lambda) \; \forall \; (x, y, z) \in \Lambda_{\boldsymbol{\pi}} \mid NC_{x,y,\boldsymbol{\pi}}(\lambda) > LC_{x,y,\boldsymbol{\pi}}(\lambda)\} \tag{13}$$

There are two important things to note in these definitions. The first is that we estimate the capacity on the drivable (walkable) area only when we optimize on vehicles (pedestrians), so we do not consider the cross-metrics like in Sect. 5.1 and capacity metrics don't have the *ve*/*pe* subscript like coverage metrics. The second is that we consider NLoS links only when a LoS link is not available or it offers a lower capacity than a NLoS link. This last observation leads to the definition of a further metric: the fraction of points for which there exist some LoS links, but a NLoS link provides a higher capacity.

$$W^{diff}(\lambda) = \frac{|\{(x, y, z) \mid NC_{x,y,\boldsymbol{\gamma}}(\lambda) > LC_{x,y,\boldsymbol{\gamma}}(\lambda) > 0\}|}{|\Lambda_{\boldsymbol{\gamma}}|} \tag{14}$$

$$D^{diff}(\lambda) = \frac{|\{(x, y, z) \mid NC_{x,y,\boldsymbol{\pi}}(\lambda) > LC_{x,y,\boldsymbol{\pi}}(\lambda) > 0\}|}{|\Lambda_{\boldsymbol{\pi}}|} \tag{15}$$

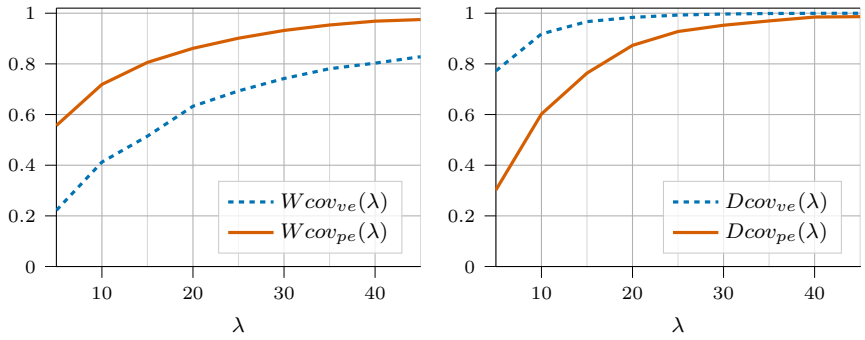
We analyze the sets of capacity metrics in the following section in terms of averages, cumulative distribution functions, and coefficient of variation to provide a link quality estimation of the coverage strategies.

6. Results

6.1. Analysis of the coverage

We start the analysis by looking at coverage metrics as observed by different areas in the city (walkable and drivable areas) for the two optimization strategies. Fig. 6a shows the coverage as perceived on walkable areas when optimizing for pedestrians ($Wcov_{pe}$) and for vehicles ($Wcov_{ve}$) as a function of the density of gNBs. Conversely, Fig. 6b shows the coverage as perceived on drivable areas when optimizing for pedestrians ($Dcov_{pe}$) and for vehicles ($Dcov_{ve}$).

[5]As previously said, we don't use z as an index in the subscript since (x, y) uniquely addresses one point.



a) Walkable areas b) Drivable areas

Figure 6: Coverage (fraction of points reached by at least one LoS link) for different optimization strategies. On the left, we measure the coverage for walkable areas ($Wcov_{pe}(\lambda)$ and $Wcov_{ve}(\lambda)$) while on the right we measure the coverage for drivable areas ($Dcov_{ve}(\lambda)$ and $Dcov_{pe}(\lambda)$).

Fig. 6b shows two very relevant conclusions. The first is that $Dcov_{ve}$ reaches 90 % coverage with $\lambda = 10$, 95 % coverage with $\lambda = 15$ and 99.9 % coverage with $\lambda = 25$ while $Dcov_{pe}$ needs 150 % and 100 % more gNB to cover 90 % and 95 % of the vehicles, respectively, and can not reach 99.9 % even with $\lambda = 45$. Considering that vehicles' coverage for autonomous driving requires high reliability, we see that there is a relevant difference when we specifically optimize for vehicles, rather than for pedestrians. The second conclusion is more generic: so far we did not have any concrete indication of how much we need to increase the density of gNBs to achieve vehicles coverage, and this result tells us that in urban areas, a reasonably low density can still be sufficient for a reliable service.

Fig. 6a instead tells a different message. There is a remarkable difference in the coverage of walkable areas when optimizing for vehicles or not. In particular, $Wcov_{ve}$ (the vehicles' optimization) allows us to cover only slightly more than 80 % of the ground. This result is worse than in our previous work [6] where we show that we could not cover more than 95 % of the ground because here we included larger public areas such as parks and squares in the analysis. On the other side, while optimizing for pedestrians and measuring on walkable areas, $Wcov_{pe}$ reaches 90 % with $\lambda = 25$ and 95 % with $\lambda = 35$.

The takeaway for the operator that needs to start deploying gNBs for LoS communications is that the goals of covering vehicles or pedestrians are competing ones. Optimizing for vehicles would reduce significantly the required density of gNBs but would not allow to reliably cover pedestrian areas.

6.2. Analysis of the capacity

While coverage gives us a qualitative measure of the impact of different optimization strategies, it does not provide us with quantitative insights. In this section, we analyze the capacity resulting from different optimization strategies,

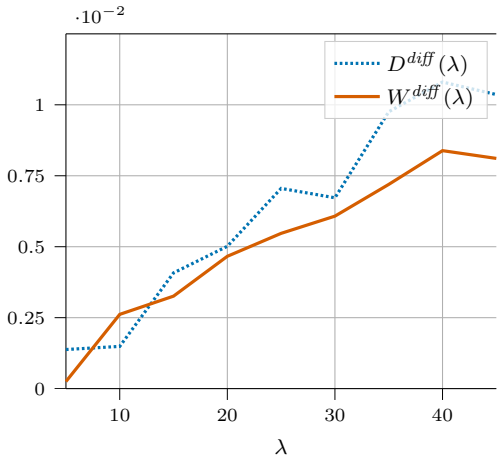


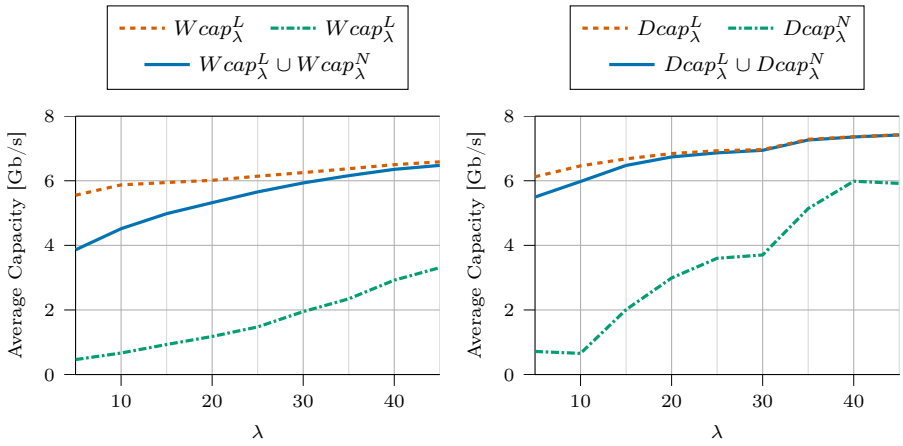
Figure 7: Fraction of points for which a NLoS gNB has been chosen even though a LoS gNB was available, for drivable areas ($D^{diff}(\lambda)$) and walkable areas ($W^{diff}(\lambda)$).

but before that, we introduce Fig. 7, which shows the fraction of NLoS links chosen even if a LoS link was available ($W^{diff}(\lambda)$ and $D^{diff}(\lambda)$). Regardless of the density, less than 1 % of the links will obtain better connectivity from a NLoS link than from a LoS link. This means that in realistic coverage conditions, the best link is almost always the LoS one and that NLoS links are of prevalent importance only when the LoS coverage is low, that is, in the leftmost part of the curves in Fig. 7. The fact that the curves in Fig. 7 grow with λ is due to the average reduction of the distance to any gNB with the growth of λ. As we already noticed this makes it more likely that the two curves in Fig. 5 intersect, as the average link length stays in the lower range.

6.2.1. Estimating the Average Capacity

For each density value λ, we compute the average capacity for LoS and NLoS links, as well as the overall average. Fig. 8a shows the average capacity measured over the walkable areas when optimizing the locations of gNBs for pedestrians. The absolute values are to be considered as an upper bound (as in conditions of interference, the capacity can be far from the Shannon limit) however the comparison is interesting, with a minimum average of 5.5 Gbit/s achieved at the lowest gNB density for LoS links, versus less than 500 Mbit/s for NLoS links. Especially at low gNB densities, this has a large impact on the overall (LoS and NLoS) average capacity: As the fraction of LoS links for $\lambda = 5$ is 55 % (Fig. 6a), the overall average is as low as 4 Gbit/s.

Fig. 8b instead shows the average capacity measured over drivable areas only when optimizing for vehicles. Again, there is a striking difference between LoS and NLoS links, however, since with the same density the drivable areas reach a better coverage, the overall average is closer to the average of LoS links. This also applies to the right end of the curve, in which a higher density reduces the



a) Walkable areas

b) Drivable areas

Figure 8: Average Capacity for different areas and LoS conditions. On the left, we measure the average capacity for walkable areas, divided LoS links ($Wcap^L_\lambda$), NLoS links ($Wcap^N_\lambda$) , and both of them ($Wcap^L_\lambda \cup Wcap^N_\lambda$)). On the right, we measure the same metrics for drivable areas.

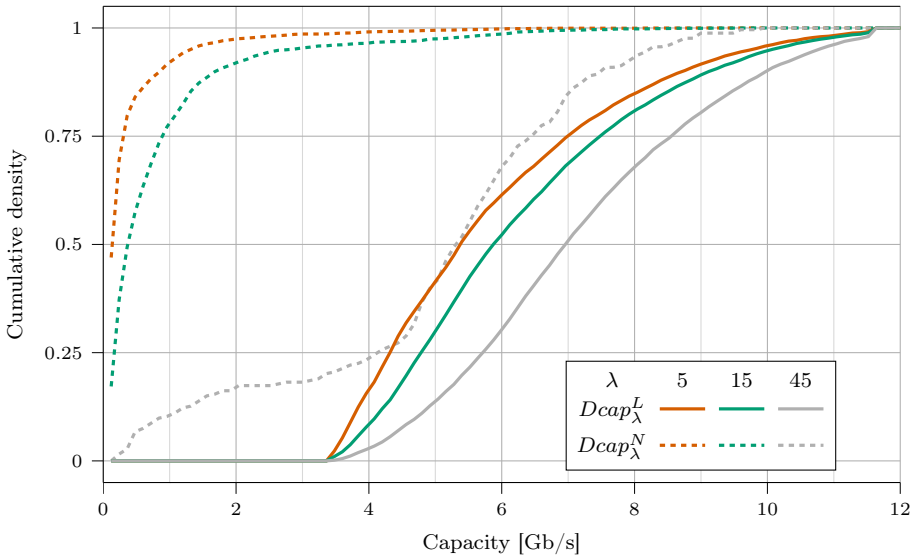
average length of all links (including NLoS) and thus the green curve gets closer to the orange one.

6.2.2. Capacity Distribution

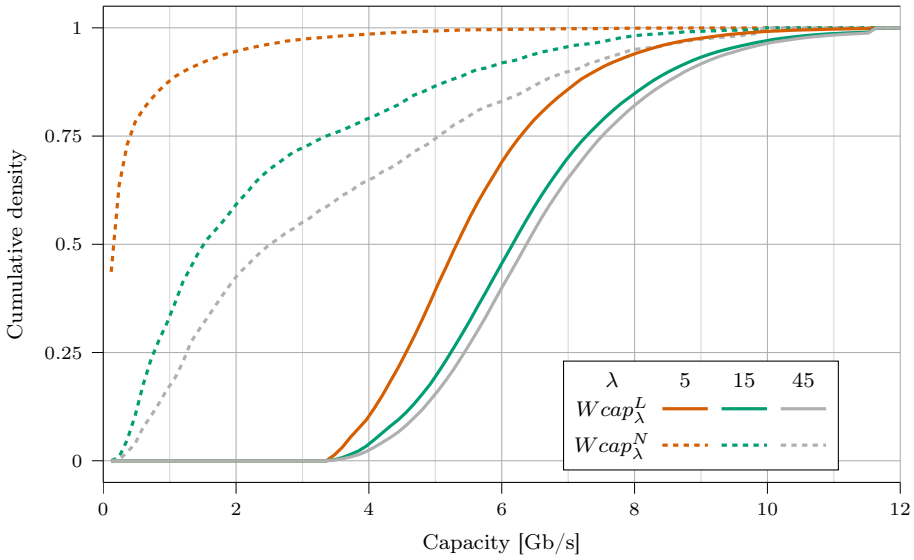
As averages hide information about the distribution of the capacities, Fig. 9a and 9b show the empirical Cumulative Density Function (eCDF) of the $Dcap_{ve}$ and $Wcap_{pe}$ sets. Each graph shows the distribution for $\lambda = 5$, the density to cover 95 % of the area ($\lambda = 35$ for pedestrians, $\lambda = 15$ for vehicles), and for $\lambda = 45$, for both LoS and NLoS links.

By looking at LoS capacities, regardless of the optimization strategy and the density, for the smallest possible density, we can see that all distributions have a reasonable shape. Most importantly we can see that no capacity is smaller than 3 Gbit/s. This means that, regardless of the objective, deploying gNBs focusing on LoS will result in very good performance.

With respect to NLoS, the quality of links is instead highly dependent on density. At low densities, a vast majority of links experience capacities that are very close to zero. For the densities providing 95 % coverage, instead, the distribution depends on the optimization strategy. When optimizing for vehicles (Fig. 9a), roughly 20 % of the links have close to zero capacity, and 75 % have a capacity lower than 1 Gbit/s. This means that, in high density zones, vehicles will experience very good network conditions, but in less dense areas communication might be at risk and some roads might be left completely without coverage or with poor communication. For the pedestrian strategy, there are fewer links with this problem, as 50 % of them have an available capacity of



a) Drivable areas



b) Walkable areas

Figure 9: empirical Cumulative Density Function (eCDF) of the capacity for different areas, LoS conditions and gNB densities. At the top, the metrics are displayed for drivable areas, while at the bottom for walkable areas. The gNB densities displayed, corresponds to the minimum considered ($\lambda = 5$), the maximum ($\lambda = 45$) and the density that guarantees a LoS coverage of 95 % of the areas ($\lambda = 15$ for drivable areas and $\lambda = 35$ for walkable areas).

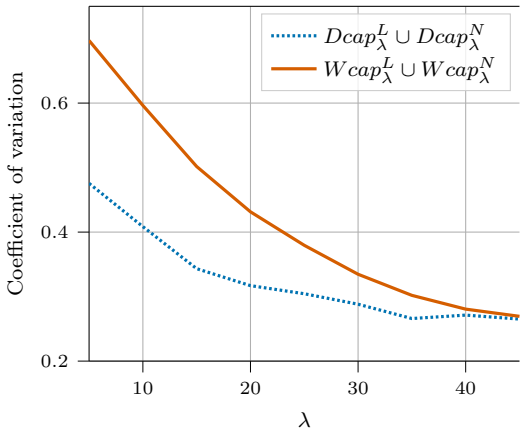


Figure 10: Coefficient of variation (ratio between standard deviation and average) of the capacity for drivable areas ($Dcap^L_\lambda \cup Dcap^N_\lambda$) and walkable areas ($Wcap^L_\lambda \cup Wcap^N_\lambda$).

at least 1.5 Gbit/s, but this comes at the cost of more than doubling the density of gNBs. At high density, the vehicular strategy results in very good performance, with just 20 % of links experiencing a capacity smaller than 3 Gbit/s and 75 % of them experiencing at least 4 Gbit/s. On the contrary, the eCDF for the pedestrian strategy grows at a much slower rate, with almost 50 % of the links having a capacity smaller than 2 Gbit/s.

Finally, Fig. 10 shows the coefficient of variation (i.e., the ratio between standard deviation and the average) of the capacity for different values of λ. Regardless of the optimization strategy, the coefficient of variation decreases with the density of gNBs. This indicates that, as we increase the density of gNBs the change in the experienced capacity between different locations reduces, indicating a more even distribution of resources among users, but also in the trajectory of one single user. Increasing the coverage not only improves the average performance but makes it more stable.

7. Conclusions

The foreseen densification of gNBs and the advancements in vehicular communications are playing a pivotal role in the deployment of XG access networks in ultradense urban areas. This paper proposes a novel data-driven method to optimize the placement of gNBs and provide crucial insights to the network operators to understand how the two coverages, for vehicular communication and pedestrians, are intertwined. We show that at least for the vehicular case a reasonably low density of gNBs is sufficient to provide 95 % coverage in urban areas, but that at the same time optimizing the coverage only towards the roads with the most traffic will not provide sufficient coverage for pedestrians. Our data numerically confirm the intuition that if operators want to have good and stable performance they must focus on the deployment of LoS links, and we

provide evidence on the expected LoS coverage for reasonable values of gNB density.

References

[1] Small Cell Forum, Hyperdense HetNets: Definition, drivers and barriers, Tech. rep. (Feb 2017).
URL `https://scf.io/en/documents/180_Hyperdense_HetNets_Definition_drivers_and_barriers.php`

[2] G. Gemmi, R. Lo Cigno, L. Maccari, On cost-effective, reliable coverage for los communications in urban areas, IEEE Transactions on Network and Service Management 19 (3) (2022) 2767–2779. doi:10.1109/TNSM.2022.3190634.

[3] S. Aoki, T. Higuchi, O. Altintas, Cooperative Perception with Deep Reinforcement Learning for Connected Vehicles, in: IEEE Intelligent Vehicles Symposium (IV 2020), IEEE, Virtual Conference, 2020. doi:10.1109/iv47402.2020.9304570.

[4] R. Lo Cigno, M. Segata, Cooperative driving: A comprehensive perspective, the role of communications, and its potential development, Elsevier Computer Communications 193 (2022) 82–93. doi:10.1016/j.comcom.2022.06.034.

[5] B. B. Haile, E. Mutafungwa, J. Hämäläinen, A Data-Driven Multiobjective Optimization Framework for Hyperdense 5G Network Planning, IEEE Access 8 (2020) 169423–169443.

[6] G. Gemmi, M. Segata, L. Maccari, Vehicles or Pedestrians: On the gNB Placement in Ultradense Urban Areas, in: 18th IEEE/IFIP Conference on Wireless On demand Network Systems and Services (WONS 2023), IEEE, Madonna di Campiglio, Italy, 2023.

[7] M. Wright, Optimization methods for base station placement in wireless applications, in: IEEE Vehicular Technology Conference, Vol. 1, 1998, pp. 387–391.

[8] N. Palizban, S. Szyszkowicz, H. Yanikomeroglu, Automation of millimeter wave network planning for outdoor coverage in dense urban areas using wall-mounted base stations, IEEE Wireless Communications Letters 6 (2) (2017) 206–209.

[9] Y. Zhang, L. Dai, E. W. M. Wong, Optimal BS Deployment and User Association for 5G Millimeter Wave Communication Networks, IEEE Trans. on Wireless Communications 20 (5) (2021) 2776–2791.

[10] P. Seda, M. Seda, J. Hosek, On mathematical modelling of automated coverage optimization in wireless 5g and beyond deployments, Applied Sciences 10 (24) (2020).

[11] C. K. Anjinappa, F. Erden, I. Guvenc, Base Station and Passive Reflectors Placement for Urban mmWave Networks, IEEE Transactions on Vehicular Technology (TVT) 70 (4) (2021) 3525–3539. doi:10.1109/tvt.2021.3065221.

[12] C. Sommer, D. Eckhoff, A. Brummer, D. S. Buse, F. Hagenauer, S. Joerer, M. Segata, Veins: The open source vehicular network simulation framework, in: Recent Advances in Network Simulation: The OMNeT++ Environment and its Ecosystem, Springer International Publishing, 2019.

[13] A. Brummer, R. German, A. Djanatliev, On the Necessity of Three-Dimensional Considerations in Vehicular Network Simulation, in: Wireless On-demand Network Systems and Services Conference (WONS), 2018.

[14] A. Brummer, R. German, A. Djanatliev, Methodology and Performance Assessment of Three-Dimensional Vehicular Ad-hoc Network Simulation, IEEE Access 11 (2023).

[15] E. Zanotto, L. Maccari, Evaluating the Impact of a 3D Simulation Model on the Performance of Vehicular Networks, in: Wireless On-Demand Network Systems and Services Conference (WONS), 2024.

[16] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, V. Koltun, CARLA: An open urban driving simulator, in: Annual Conference on Robot Learning, 2017.

[17] P. Jacquet, D. Popescu, B. Mans, Connecting flying backhauls of drones to enhance vehicular networks with fixed 5g nr infrastructure, in: IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops, 2020, pp. 472–477. doi:10.1109/INFOCOMWKSHPS50562.2020.9162670.

[18] G. Gemmi, R. L. Cigno, L. Maccari, On the properties of next generation wireless backhaul, IEEE Transactions on Network Science and Engineering (2022) 1–12doi:10.1109/TNSE.2022.3205864.

[19] K. Haneda, et al., 5G 3GPP-Like Channel Models for Outdoor Urban Microcellular and Macrocellular Environments, in: 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), 2016, pp. 1–7.

[20] P. A. Lopez, E. Wiessner, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y.-P. Flotterod, R. Hilbrich, L. Lucken, J. Rummel, P. Wagner, Microscopic Traffic Simulation using SUMO, in: 21st IEEE International Conference on Intelligent Transportation Systems (ITSC 2018), IEEE, Maui, HI, 2018, pp. 2575–2582. doi:10.1109/itsc.2018.8569938.

[21] L. Codecá, R. Frank, S. Faye, T. Engel, Luxembourg SUMO Traffic (LuST) Scenario: Traffic Demand Evaluation, IEEE Intelligent Transportation Systems Magazine 9 (2) (2017) 52–63. doi:10.1109/mits.2017.2666585.

[22] M. Rapelli, C. E. Casetti, G. Gagliardi, TuST: from Raw Data to Vehicular Traffic Simulation in Turin, in: 23rd IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications (DS-RT 2019), IEEE, Cosenza, Italy, 2019. doi:10.1109/ds-rt47707.2019.8958652.

[23] L. Codecá, J. Härri, Towards multimodal mobility simulation of C-ITS: The Monaco SUMO traffic scenario, in: 9th IEEE Vehicular Networking Conference (VNC 2017), IEEE, Turin, Italy, 2017, pp. 97–100. doi:10.1109/VNC.2017.8275627.

[24] European Telecommunications Standards Institute, Study on channel model for frequencies from 0.5 to 100 GHz, TR 138 901 V14.0.0, European Telecommunications Standards Institute (5 2017).