On the Effect of Coverage Range Extent on Next-Cell Prediction Error for Vehicular Mobility in 5G/6G Networks: a Novel Theoretic Model

Peppino Fazio, *† *Member, IEEE*
* DSMN, Ca' Foscari University of Venice, Via Torino 155, 30172 Mestre (VE), Italy
† Department of Telecommunications, VSB – Technical University of Ostrava, 17. listopadu 2172/15, 708 00 Ostrava, Czechia
*†peppino.fazio@{unive.it, vsb.cz}

***Abstract*—Over the last decade, 5G and forthcoming 6G architectures have undergone extensive standardization and preparations for the future. The literature in this field is saturated with studies on predicting mobile trajectories in cellular systems and guaranteeing quality of service and an adequate user experience in these environments. The current study aims to bridge mobility prediction and 5G/6G predictive approaches and demonstrate that the intrinsic paradigm of femto-cell and nano-cell deployment (based on very small radio coverage radii) for 5G provides the means to obtain more accurate time series data on user mobility and thus enable predictive models (e.g., machine learning) as suitable technologies for integration with 6G standards. This field is therefore an important avenue of research.**

***Index Terms*—6G, Mobility Prediction, Machine Learning, Deep Learning, Femto-cells, Nano-Cells, Vehicular Networks.**

I. Introduction

Since the advent of first and second generation (2G) cellular systems [1] several decades ago, researchers have increasingly studied mobility management in cellular networks, and as the latest system (5G) [2], [3] is deployed, this trend is likely to accelerate. Foreknowledge of host behavior allows researchers and engineers to enhance overall system and network quality in a range of key aspects, for example service continuity [4], Quality of Service (QoS) and Quality of Experience (QoE) [5], resource management [6], cellular and vehicular network integration [7], and others [8].The evolution of cellular systems has been prodigious, and by 2030, it is expected that 6G technology will be commercially available and provide users with an unprecedented telecommunications experience [9], [10]. The current study examines 6G architecture access layers in relation to currently proposed 6G wireless and cellular communications technologies (e.g., global coverage, full applications, all digital, all spectra, all senses, strong security) and aims to demonstrate that it is already easier to predict host movements between coverage cells in 5G and beyond architectures, given that their coverage radii are much smaller than in the past. The current study attempts to demonstrate that, in theory, the smaller the radii of the system cell coverage the greater the prediction accuracy (independently of the predictive model). At first glance this may seem intuitive, but given the varying character of areas covered by cellular networks, (urban, suburban, rural), irregularity in coverage shape, and the range of reasons for which users travel roads with active mobile devices (e.g., holidaying, commuting, walking), predicting movements as a function of the size of a coverage cell presents a complex task. The current study therefore investigates how the evolution of wireless cellular system generations and coverage radii potentially impact the performance of mobility prediction approaches. The remainder of the paper is organised as follows: Section II describes recent works on mobility prediction and the latest research results for beyond-5G technologies; Section III introduces a predictive model and its related assumptions; Section IV discusses numerical results in relation to the proposed coverage structure; Section V summarizes conclusions and outlines future work.

II. Related Work and Cellular Coverage Trend

This section reviews related work on current predictive models for mobile cellular networks and describes the main features of coverage cells in 5G and beyond environments.

A. Predictive Models for 5G/6G Cellular Networks in 2023

Continual advances in wireless technologies have raised consumer expectations for seamless connectivity in cellular networks and high QoE levels from Internet Service Providers (ISPs). A long-term study [8] attempts to solve mobility issues in wireless networks (e.g., resource and handover management, location updating, channel degradation) by predicting user movement either along a trajectory or at the cell level. Several recent works based on mobility prediction (both trajectory and cell patterns) explore enhancements to system performance at several levels, for example traffic throughput and mitigation of interference, spectrum optimization in Cognitive Radio Networks (CRNs) according to user preference and demand [11], optimization of handover procedures to reduce latency [12], prediction of traffic to optimize Base Station (BS) resource management [13], energy-aware mobility prediction to manage Remote Radio Head (RRH) activation in 5G Cloud Radio Access Networks (CRANs) [14], and optimization of Cache Hit Ratio (CHR) using knowledge of future trajectories [15]. It is clear that knowledge of the host's future position (or conditions, more generally) affords

architectural benefits for a wide range of applications; referred to as Native Intelligence (NI), this is one of the main objectives behind future 6G networks [16]. If architecture provides a separation between the control plane and the data plane (e.g., from 3G/4G to 6G), mobility prediction can be exploited to manage handover optimization and obtain huge savings in signaling overhead and Call Blocking Probability (CBP) [17]. In [18], the authors propose an accurate trajectory prediction algorithm for pre-configuring radio resources in cells that experience a future mobile host trajectory, resulting in a enhanced communication efficiency. The studies in [19], [20] implement Markov Chain (MC) predictive models. In the first study, the authors introduce the concept of a transition matrix for an MC, where spatial and temporal information are separated for outdoor environments to reduce correlation loss. In the second study, the authors use an MC for indoor mobility prediction, where mmWave coverage (e.g., 5G) is assumed, and attempt to reduce handover latency. Another work [21] extends the classic MC model using the Hidden Markov Model (HMM) to optimize power consumption while maintaining QoS. It also demonstrates that HMMs enhance prediction accuracy while reducing information loss, albeit at the cost of greater computational complexity. In [22], [23], the authors propose prediction algorithms based on Machine Learning (ML) approaches. In the first study, the authors compare four ML approaches and evaluate the benefits of these models in terms of prediction accuracy; the XGBoost ML scheme achieved a greater than 98% accuracy, the highest among the proposed models. In the second paper, the authors investigate the Support Vector Machine (SVM) to predict the trajectories of mobile vehicular nodes and ensure service continuity. In [24], [25], other ML models are applied in Long Short Term Memory (LSTM) networks to solve mobility prediction issues and to investigate extensions of LSTM networks for correct handover prediction management in 5G networks. In a previous work [26], the present author analyzes the relationship between next-cell prediction errors and the number of in-advance reserved coverage areas, demonstrating theoretically how to detect the trade-off between prediction error and resource wastage.

B. From kbps to Gbps: Reduction in Cellular Coverage Extent

This sub-section examines the evolution of digital cellular architectures in relation to coverage radius. Since the early 1990s, BS coverage has decreased for several reasons, which include spectrum efficiency, bandwidth/channel limitations, frequency reuse, QoS/QoE demand, and greater required throughput. Referring to the work in [8] concerning high throughput connections, it is clear when digital cellular communications (2G) were deployed, the coverage radius of a BS would reach a maximum extent of 35 km, with a channel capacity of around 200 kHz. With 3G technology, the maximum radio coverage reduced to 8km, with a channel capacity of 5 MHz. Starting with the fourth generation (4G), the coverage radius reduced drastically to a maximum of 2–3km. With 5G, the coverage radius depends on the carrier frequency which is used: for the low and mid-band spectrum (below the 3.7 GHz C-band), the achievable coverage radius is 4–5km; for the mmWave spectrum band (24 GHz to 47 GHz), with a good Line-of-Sight (LoS) connection, the range is instead 0.0015–0.6 km. Figure 1 illustrates the trend in average radius per cell over the years. The concept of the

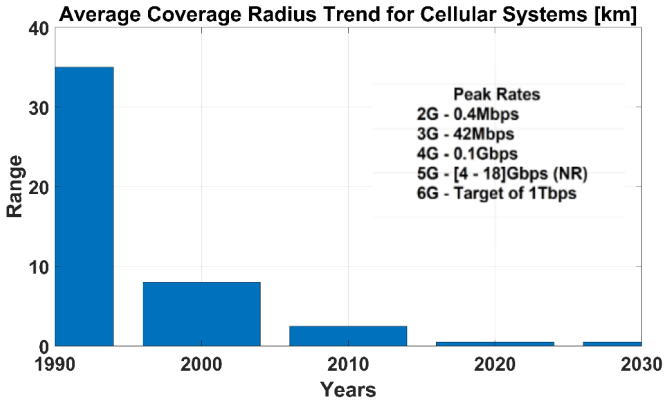

Fig. 1: Trend in average coverage radius per cell over the years/network generations (uplink and downlink).

femtocell introduced with 5G and IEEE802.11ac/ax (WiFi6, 6E) [27] will still be still applied in 6G, however each user will also be served simultaneously by more than one cell [28]. The results of the present study are therefore discussed with relevance to 6G cellular technology, but not the proposed cell-less (CL) architecture. The use of small cells in 6G networks relates to carrier frequencies, especially THz communications [29], where interactions with the atmosphere will have a heavy affect on transmitted signals. The spectrum band at 600–800 GHz experiences environmental attenuation of up to 200 dB/km, therefore it is not feasible, whereas at distances up to 0.1 km the typical attenuation is between 10 dB and 20 dB. The work in [29] describes several experiments in which THz transmissions achieve a coverage radius of about 0.02 km (up to 10 THz in the spectrum band). The present study offers a deep theoretical analysis of the relationship between coverage radius and the trajectory or next-cell prediction error in future-generation cellular systems; to the best of the present author's knowledge, this analysis is novel. The study also attempts to demonstrate the suitability of ML approaches for use with 6G networks by heavily exploiting the the small cell concept. In summary, a very general model is applied to demonstrate that mobility prediction errors relate to the coverage radii of cells and can be decreased in smaller areas of coverage.

III. System Model, Concepts and Proposal

This section discusses the proposed model for the relationship between the coverage area and the underlying road topology as a function of the coverage radius. Table I reviews the notation used in the article.

A. Basic Definitions, Geometric/Analytic Road Modeling

This sub-section presents the basic concepts and analytical definitions in the study.

TABLE I: Notation used in this work.

Symbol or notation	Meaning
Basic definitions	
p_i	generic geometric point i in R^2 expressed as a coordinate pair x_i and y_i
$s(p_i, p_j)$	segment between two points p_i and p_j
y_s	algebraic equation associated with segment s, defined in the domain $D_s = [x_i, x_j]$
$s_1 < s_2$	segment s_1 is a sub-segment of s_2 (same y_s but smaller domain)
GA	geographic area (GA) described by R^2
$G_{GA} =< V, V^*, E, E^*, Q^* >$	graph associated with the roads in the GA, where V is the set of vertices, E is the set of edges
V^*	set of vertices expressed as points $\implies v_i \equiv p_i$
E^*	set of edges expressed as segments $\implies e_j \equiv s(p_{j1}, p_{j2})$
Q^*	set of equations y_s associated with each edge in E^*
RD	set of roads r_i, each road being expressed as a set of edges in G_{GA}
E_{RD}	set of edges associated with the roads in RD
Geometric definitions	
C	circumference of radius r covering a part of the GA
$s \subseteq C$	segment s is fully contained in C
$s \subset\sim C$	segment s is partially contained in C
$s \overset{\text{ch}}{\subseteq} C$	segment s is a chord on C
$s \nsubseteq C$	segment s is external to C
$r_i \overset{\text{C}}{\otimes} r_j$	two roads r_i and r_j are connected in C (by definition, roads are connected if they share at least one point)
$\lvert r_j \rvert_C$	length of road r_j in C
$\lvert RD \rvert_C$	average length of roads in RD with respect to C
Vehicle model definitions	
L_C	number of roads in RD covered by C
$\theta_i^C(t)$	road r_i occupancy (in terms of % of space occupied by vehicles) at time t with respect to the coverage of C
$x_i^C(t)$	number of vehicles in r_i at time t with respect to the coverage of C
$\overline{\Gamma}$	average vehicle length
$M^C(t)$	mobility time variant matrix, with L_C rows and L_C columns
$m_{i,j}(t)$	(i,j)-th element of $M^C(t)$, i.e., the probability of moving from road r_i to road r_j at t under the coverage of C
$t_i^{abs}(t)$	absorption time for road r_i, i.e., the time spent by a user to travel the entire road r_i
$v_i^{avg}(t)$	average speed on road r_i at time t
v_i^{min}, v_i^{max}	minimum observed speed and maximum permitted speed on road r_i
$\Delta road_i$	length increase of covered road r_i after extension of the coverage radius of C from r to R, where R=$r + \Delta r$

Def. 1: Given two points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$, where $p_1, p_2 \in \mathcal{R}^2$, the length of the line segment connecting p_1 and p_2, $s(p_1, p_2)$ represents the Euclidean distance between p_1 and p_2, expressed as:

$$s(p_1, p_2) \overset{\text{def}}{=} \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \tag{1}$$

where $s(p_1, p_2) \in \mathcal{R}$. It can also be expressed in compact form as s, where the two endpoints p_1 and p_2 are not required.

Def. 2: A line segment $s(p_1, p_2)$ analytically described by the function $y_s(x) = m_s \cdot x + q_s$ is expressed as:

$$\begin{cases} y_s(x_1) = m_s \cdot x_1 + q_s = y_1 \\ y_s(x_2) = m_s \cdot x_2 + q_s = y_2, \end{cases} \tag{2}$$

where $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$ and $x \in [x_1, x_2]$.

Def. 3: Given the line segment $s(p_1, p_2)$, its function $y_s(x)$ and the range $[x_1, x_2]$, then $y_s(x)$ is said to be defined in its domain $D_s \overset{\text{def}}{=} [x_1, x_2]$.

Def. 4: Given two line segments $s_1(p_1, p_2)$ and $s_2(q_1, q_2)$, s_1 is said to be a *sub-segment* of s_2 ($s_1 < s_2$) such that

$$\begin{cases} y_{s_1} = y_{s_2} \Rightarrow (m_{s_1} = m_{s_1}) \wedge (q_{s_1} = q_{s_1}) \\ D_{s_1} \subset D_{s_2}. \end{cases} \tag{3}$$

Given these definitions, we describe the first part of the modeling approach for the considered mobility scheme. First, a graph must be associated with a set of interconnected roads belonging to a generic geographic area (GA). Thus, given a graph $G_{GA} =< V, E >$, then V is the set of nodes (where $|V|$=n), E is the set of edges (where $|E|$=m), and $e_i \in E$ is a generic edge composed of nodes v_{i_1}, $v_{i_2} \in V$, where $e_i = (v_{i_1}, v_{i_2})$. The edges have no direction, and given a set of interconnected roads $RD = r_1, ..., r_k$, a generic road $r_j \in RD$ can be modeled as a sequence of edges $E_{r_j} \in E$, each including a number of nodes. Thus, the complete road structure can be modeled as $E_{RD} = \{E_{r_1}, ..., E_{r_n}\}$, where $\bigcup_{j=1}^{k} E_{r_j} = E$ and $E_{r_j} \bigcap E_{r_l} = \emptyset$, $\forall E_{r_j}, E_{r_l} \in E$, with $j \neq l$.

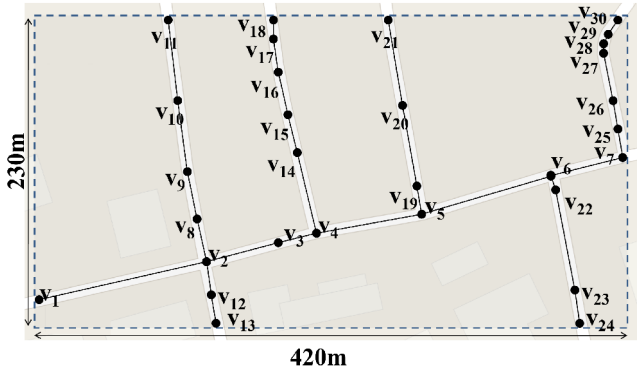

Fig. 2: Example of a graphic approximation of a GA.

Figure 2 represents a small area of the city of Rome (approximately 420 x 230 m^2). We observe that

- $V = \{v_1, .., v_{30}\}$, $|V| = n = 30$;
- $E = \{(v_1, v_2),\ ...,(v_6, v_7),\ (v_2, v_8),\ ...\ ,\ (v_{10}, v_{11}),\ (v_2, v_{12}),\ v_{12}, v_{13}),\ (v_4, v_{14}),\ ...,\ (v_{17}, v_{18}), (v_5, v_{19}),\ ...,\ (v_{20}, v_{21}),\ (v_6, v_{22}),\ ...,\ (v_{23}, v_{24}),\ (v_7, v_{25}),\ ...,\ (v_{29}, v_{30})\}$=$\{e_1, ..., e_{29}\}$, $|E| = m = 29$;
- $RD = \{r_1, ..., r_7\}$;
- E_{r_1}=$\{e_1, e_2, e_3, e_4, e_5, e_6\}$ = $\{(v_1, v_2),\ (v_2, v_3),\ (v_3, v_4),\ (v_4, v_5),\ (v_5, v_6),\ (v_6, v_7)\}$;
- E_{r_2}=$\{e_7,\ e_8,\ e_9,\ e_{10}\}$=$\{(v_2, v_8),\ (v_8, v_9),\ (v_9, v_{10}),\ (v_{10}, v_{11})\}$;
- $E_{r_3} = \{e_{11}, e_{12}\}$=$\{(v_2, v_{12}), (v_{12}, v_{13})\}$;
- $E_{r_4} = \{e_{13}, e_{14}, e_{15}, e_{16}, e_{17}\}$=$\{(v_4, v_{14}), (v_{14}, v_{15}),\ (v_{15}, v_{16}),\ (v_{16}, v_{17}), (v_{17}, v_{18})\}$;
- $E_{r_5} = \{e_{18}, e_{19}, e_{20}\}$ = $\{(v_5, v_{19}),\ (v_{19}, v_{20}),\ (v_{20}, v_{21})\}$;
- $E_{r_6} = \{e_{21}, e_{22}, e_{23}\}$=$\{(v_6, v_{22}), (v_{22}, v_{23}), (v_{23}, v_{24})\}$
- E_{r_7} = $\{e_{24},\ e_{25},\ e_{26},\ e_{27},\ e_{28},\ e_{29}\}$=$\{(v_7, v_{25}), (v_{25}, v_{26}),\ (v_{26}, v_{27}), (v_{27}, v_{28}),\ (v_{28}, v_{29}),\ (v_{29}, v_{30})\}$.

For the road topology in Figure 2, it is simple to verify that $\bigcup_{j=1}^{7} E_{r_j} = E$ and $E_{r_j} \bigcap E_{r_l} = \emptyset\ \forall E_{r_j}, E_{r_l} \in E,\ i \neq j$. At this point, we extend the classic graph modeling approach by adding three extra sets:

- V^* = $\{v_j^* = (v_{j_x}, v_{j_y})\}$, where v_{j_x} and v_{j_y} are the Cartesian coordinates of $v_j \in V$ and $v_{j_x}, v_{j_y} \in GA \subset \mathcal{R}^2$, i.e., each $v_j^* \in V^*$ represents the point in $GA \subset \mathcal{R}^2$ where $v_j \in V$ is situated;
- $E^* = \{s_i^*(v_{i_1}^*, v_{i_2}^*)/e_i = (v_{i_1}, v_{i_2}) \in E\}$, i.e., each $s_i^* \in E^*$ represents the segment associated with $e_i \in E$, connecting nodes $v_{i_1}, v_{i_2} \in V$, situated at in $v_{i_1}^*$ and $v_{i_2}^*$;
- $Q^* = \{y_{s_i}^*(x)/s_i^* \in E^*\}$, i.e., each element of Q^* represents the equation of the associated edge $e_i \in E$.

Therefore, based on the above definitions, a physical length and an equation can be assigned to each edge composing the graph G associated with GA, G_{GA}, since for each $s_i^* \in E^*$ the related length and $y_{s_i}^*(x)$ are easily evaluated.

B. Geometric/Analytic Modeling of Roads and Coverages

The shape of a generic cell is assumed to be circular, without loss of generality; the present study does not consider real shapes but demonstrates the result of changing the coverage area.

Def. 5: Given a circumference C of radius r centered at (x_0, y_0), then a segment $s(p_1, p_2)$ defined in $[x_1, x_2]$ is completely contained in C, $s \overset{\text{def}}{\subseteq} C$, if and only if the system

$$\begin{cases} (x - x_0)^2 + (y - y_0)^2 \leq r^2 \\ y = m_s \cdot x + q_s \\ x_1 \leq x_2 \end{cases} \tag{4}$$

can be solved; for $x \in [x_1, x_2]$, it is equivalent to

$$(1 + m_s^2) \cdot x^2 + 2 \cdot m_s \cdot q_s \cdot x + q_s^2 \leq r^2 \tag{5}$$

The length of a segment that is completely contained in C ($s \subseteq C$) is easily solved with Equation (1).

Def. 6: Given a circumference C of radius r centered at (x_0, y_0), then a segment $s(p_1, p_2)$ defined in $[x_1, x_2]$ is partially contained in C, $s \overset{\text{def}}{\subset\sim} C$, if and only if a value $\overline{x} \in [x_1, x_2]$ exists for which the relation expressed in Equation (5) is verified only for $x \in [x_1, \overline{x}]$ or for $x \in [\overline{x}, x_2]$. In order to evaluate the length of the segment partially contained in C, Equation (1) can be modified to

$$\overline{s} = \sqrt{(\overline{x} - x_c)^2 + [y_s(\overline{x}) - y_c]^2}, \tag{6}$$

where (x_c, y_c) are the coordinates of the end-point of s contained in C, i.e., (x_1, y_1) or (x_2, y_2).

Def. 7: Given a circumference C of radius r centered at (x_0, y_0), then a segment $s(p_1, p_2)$ defined in $[x_1, x_2]$ is a chord in C, $s \overset{\text{ch}}{\subseteq} C$, if and only if two values $\overline{x}_1$, $\overline{x}_2 \in [x_1, x_2]$ exist such that $x_1 < \overline{x}_1 < \overline{x}_2 < x_2$ for which the relation expressed in Equation (5) is verified only for $x \in [\overline{x}_1, \overline{x}_2]$. In order to evaluate the length of part of the segment contained in C, Equation (1) can be modified to

$$\overline{s} = \sqrt{(\overline{x}_1 - \overline{x}_2)^2 + [y_s(\overline{x}_1) - y_s(\overline{x}_2)]^2}, \tag{7}$$

Def. 8: Given a circumference C of radius r centered at (x_0, y_0), if a segment $s(p_1, p_2)$ defined in $[x_1, x_2]$ is neither completely/partially contained in C nor a chord, then s is said to be external to C: $s \overset{\text{def}}{\notin} C$.

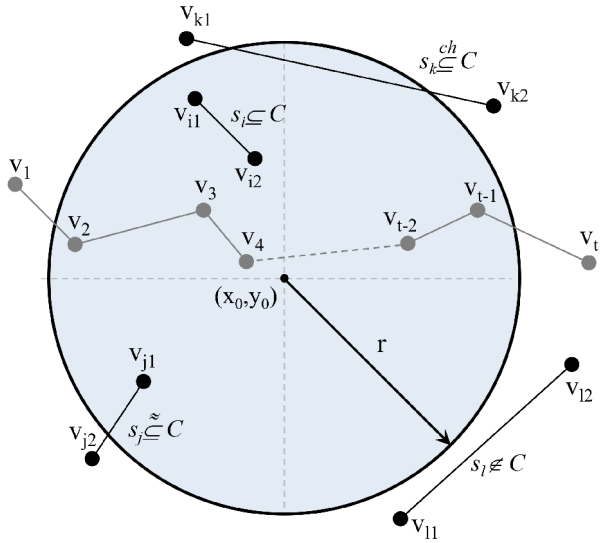

Fig. 3: Possible configurations for a road and several segments with respect to C.

Figure 3 illustrates all the conditions expressed by Definitions 5–8 for a segment s with respect to circumference C. Figure 3 also shows (in grey) that, for given a generic road $r_j \in RD$, we can write $r_j \subseteq C$ if $s_j^* \subseteq C \forall s_j^* \in E_{r_j}^*$, where $j = 1..|E_{r_j}^*|$, i.e., a road $r_j \in RD$ is completely contained in C if and only if all the segments belonging to r_j are completely contained in C. Similarly, $r_j \notin C$ (r_j is external to C) if $s_j^* \notin C\ \forall s_j^* \in E_{r_j}^*$, where $j = 1..|E_{r_j}^*|$. The road

$r_j \in RD$ is partially contained in C, if $\exists s_j^* \in E_{r_j}^*$ / $s_j^* \subset\sim C$ or $\exists s_j^* \in E_{r_j}^*$ / $s_j^* \overset{\text{ch}}{\subseteq} C$.

Def. 9: Given two roads r_i, $r_j \in RD$, they are connected in C ($r_i \overset{\text{C}}{\otimes} r_j$) if $\exists s_i^* \in E_{r_i} \wedge \exists s_j^* \in E_{r_j}$ / $e_i = (v_{i_1}, v_{i_2})$, $e_j = (v_{j_1}, v_{j_2})$ and $(v_{i_1} = v_{j_1} = v_c \vee v_{i_2} = v_{j_2} = v_c) \wedge v_c = (x_c, y_c)$, where $(x_c - x_0)^2 + (y_c - y_0)^2 \leq r^2$, i.e., two roads are connected in C if they share at least one node that is contained in the area covered by C. Figure 4 shows an example of four road intersections under the coverage of C, respecting Definition 9.

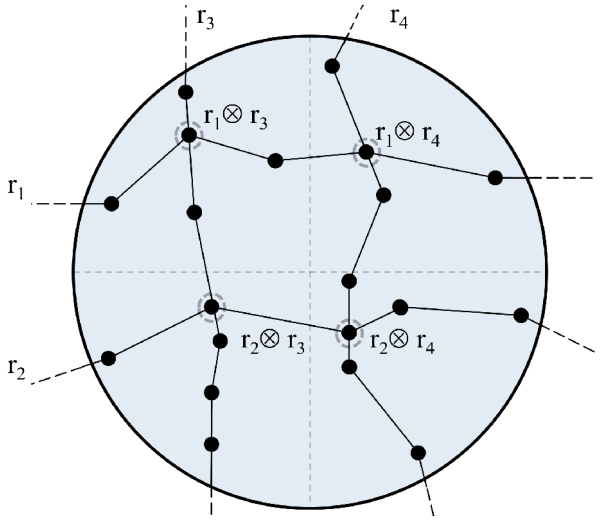

Fig. 4: Example of road intersections inside C.

Applying the concepts introduced in the previous sub-sections, it is now possible to evaluate the length of a road for only the part that belongs to a circular generic coverage cell C. Given a road $r_j \in RD$, it is described by the sets E_{r_j} (edges), $E_{r_j}^*$ (segments) and V_{r_j} (nodes). Examining Figure 3 again, it is clear that the relations $|E_{r_j}| = |E_{r_j}^*| = |V_{r_j}| - 1 = t_j - 1$ are always verified $\forall j = 1, ..., k$. If $|r_j|$ represents the length of r_j, we write

$$|r_j| = \sum_{i=1}^{t_j - 1} \overline{s}_i^* \Rightarrow |r_j|_C = \sum_{i=1}^{I_C} \overline{s}_i^* + \sum_{j=1}^{J_C} \overline{s}_k^* + \sum_{k=1}^{K_C} \overline{s}_k^*, \quad (8)$$

where the quantity $|r_j|_C$ represents the length of r_j composed of segments that are completely/partially contained, external, or represent a chord on circumference C. For Equation (8), the following relations remain valid: $s_i^* \subseteq C, \forall i = 1, ..., I_C$, $s_j^* \subset\sim C, \forall j = 1, ..., J_C$, $s_k^* \overset{\text{ch}}{\subseteq} C, \forall k = 1, ..., K_C$, $s_i^*, s_j^*, s_k^* \in E_{r_j}^*$, $I_C + J_C + K_C \leq t_j - 1$. Equation (8) thus allows us to evaluate only the completely/partially contained segments and chords, where the external segments contribute nothing to the length of r_j inside the given shape. Therefore, the value $I_C + J_C + K_C$ (total completely/partially contained segments and chords) could be also less than the total number of edges $t_j - 1$ (potential external segments are ignored).

Def. 10: For the set of roads RD that can be modeled with $G_{GA} =< V, E >$, the average length of the roads belonging to the coverage area C can be expressed as

$$\overline{|RD|}_C = \frac{1}{L_C} \sum_{l=1}^{L_C} \left(\sum_{i=1}^{I_{l_C}} \overline{s}_{l_i}^* + \sum_{j=1}^{J_{l_C}} \overline{s}_{l_j}^* + \sum_{k=1}^{K_{l_C}} \overline{s}_{l_k}^* \right), \quad (9)$$

where L_C is the total number of roads that share at least one segment with C, s_{l_i} is the i-th segment of r_l completely contained in C, s_{l_j} is the j-th segment of r_l partially contained in C and s_{l_k} is the k-th chord of r_l.

C. Modeling Vehicle Traffic Evolution

This sub-section presents a model of the vehicular environment and the derived probabilistic approach for calculating the probability of traveling from road r_i to road r_j under the coverage of C. By observing the number of vehicles passing along a road, traffic evolution can be incorporated by using a discrete time dynamic system to analyze mobile host behaviors. In particular, we define a square $L_C x L_C$ time variant matrix $M^C(t)$, for which each element $m_{i,j}^C(t)$ represents the probability of passing from road r_i to road r_j in the time interval $[t, t+1]$. It is important to recall that $L_C = |RD|$ and that each road is completely/partially contained or a chord on the coverage circumference C.

The main aim now is to observe the trend of the average number of mobile hosts on different roads and to derive an expression for the L_C^2 elements $m_{i,j}^C(t)$. Many optimization approaches that describe the state equation introduced above can be applied [30], but for the aims of the current study, the system is modeled according to typical traffic/user parameters. A mobile user changes its current road according to local rules (e.g., the congestion level of adjacent roads), without involving system variables, and is distant from the current mobile host's position. In the proposed model, the non-diagonal elements $m_{i,j}^C(t)$, $i \neq j$, are non-zero if and only if $r_i \overset{\text{C}}{\otimes} r_j$ (obviously, a mobile host cannot choose a road not connected to the current road unless they intersect, i.e. a crossroad). To evaluate the expression $m_{i,j}(t)$, $i \neq j$, we introduce a probabilistic approach that examines mobile host behavior in the proximity of crossroads and determines the probability of the mobile host changing roads. Generally, drivers prefer roads that are more convenient and less congested. If $\theta_i^C(t)$ represents the road occupancy on $r_i^C \in RD$, it can be expressed as

$$\theta_i^C(t) = \frac{x_i^C(t) \cdot \overline{\Gamma}}{|r_i|_C}, \quad (10)$$

where $x_i^C(t)$ is the number of vehicles on r_i under the coverage of C and $\overline{\Gamma}$ is the average length of a vehicle (the numerator is always less than the denominator due to curves, minimum distance between vehicles, etc.), and the possible values belong to [0,1).

Figure 5 graphs the effect of road length ($|r_i|_C$) and the number of vehicles (the value of x_i^C) at a particular time on the occupancy of a general road $r_i \in RD$, in strict relation to Equation (10) such that $|r_i|_C >> \overline{\Gamma}$. For an average vehicle

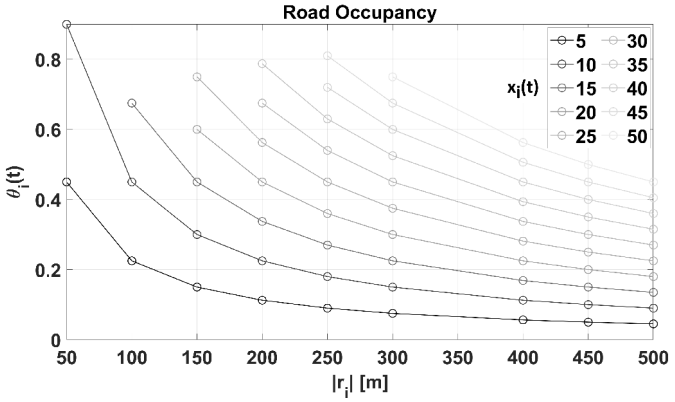

Fig. 5: Graphical representation of the trend of $\theta_i(t)$ for different road lengths and vehicle numbers, for an avergae vehicle length $\overline{\Gamma}$ = 4.5 m.

length $\overline{\Gamma}$ = 4.5, road occupancy demonstrates an increasing trend for greater numbers of vehicles and a decreasing trend for greater lengths.

Therefore, the probability of changing to road r_j from road r_i, $r_i, r_j \in RD$, $i \neq j$, is related to $1 - \theta_i^C(t)$ and expressed as

$$m_{i,j}(t) \stackrel{\text{def}}{=} \frac{1 - \theta_i^C(t)}{\sum_{l=1}^{K_i}[1 - \theta_l^C(t)]},\ i \neq j, \tag{11}$$

where K_i is the number of roads $r_j \in RD$, such that $r_i \overset{\text{C}}{\otimes} r_j$ and $i \neq j$.

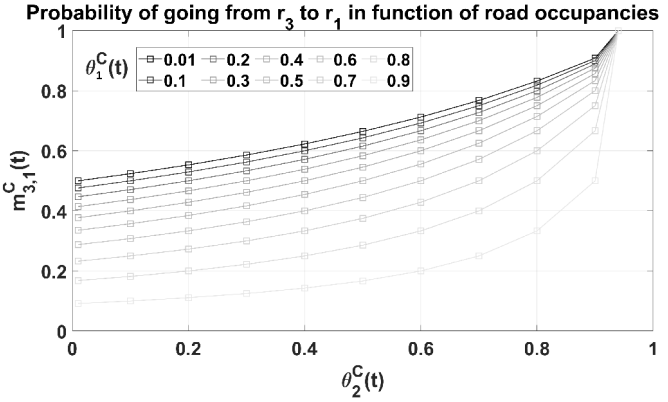

Fig. 6: Trend of $m_{3,1}^C(t)$ versus the function of $\theta_1^C(t)$ and $\theta_2^C(t)$ for the set of roads described in Figure 4 (fixed time).

For the four roads described in Figure 4, Figure 6 shows the probability of changing from road r_3 to road r_1, where r_1, r_2, r_3, $r_4 \in RD$, $r3 \overset{\text{C}}{\otimes} r1$, and $r3 \overset{\text{C}}{\otimes} r2$, versus the function $\theta_1^C(t)$ and $\theta_2^C(t)$ for traffic occupancy. It is assumed that $|r_1| = |r_2| = |r_3| = 200$ m. The illustrated curves reflect the behavior of the mobile hosts always selecting a road with lower density; therefore when r_1 is under-loaded, M^C assumes high values (a transition from r_3 to r_1 is more probable). By contrast, the probability of changing to road r_1 is low and increases only if r_2 also becomes denser. Once $m_{i,j}^C(t)$, $i \neq j$, is calculated, then we examine the diagonal elements of $M^C(t)$, with $m_{i,i}(t)$, which indicates the probability of remaining on the same road at each discrete observation time. These values depend on the length of the road $r_i \in RD$ within the circumference $|r_i|^C$ and on the average mobile host traveling speed $\overline{v}_i(t)$ [31]. The resulting quantities are inversely proportional to the average of the defined absorption time $t_i^{abs}(t)$ spent by a mobile host traveling along r_i, taking into account traffic conditions and congestion [31] (extreme situations when a road $r_i \in RD$ is completely congested, i.e., $t_i^{abs}(t) \to \infty$, are not considered). It follows that [31]:

$$m_{i,i}^C(t) \stackrel{\text{def}}{=} \frac{v_i^{avg}(t)}{|r_i|_C} \stackrel{\text{def}}{=} \frac{1}{t_i^{abs}(t)}, \tag{12}$$

since the greater the traveling time, the greater the probability that a mobile host will change to a more viable road (having a lower probability of staying on the current road). It also follows that, in the case of the $m_{i,i}^C(t)$ elements, the road occupancy $\theta_i^C(t)$ and traveling speed can be expressed as a convex combination of the minimum and maximum road speeds, v_i^{min} and v_i^{max}, respectively, which depend on the main structural road parameters [30], [32]:

$$\begin{aligned} v_i^{avg}(t) =& [1 - \theta_i^C(t)] \cdot v_i^{max} + \theta_i^C(t) \cdot v_i^{min} = \\ =& v_i^{max} + \theta_i^C(t) \cdot [v_i^{min} - v_i^{max}]. \end{aligned} \tag{13}$$

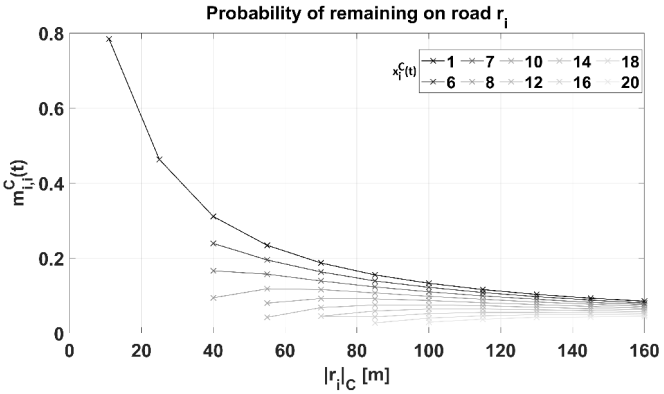

Fig. 7: Typical trend of $m_{i,i}^C$ versus $|r_i|_C$ and $x_i^C(t)$.

Figure 7 plots the probability of remaining on the same road $r_i \in RD$, on the basis of road length and the number of vehicles on the road (fixed time). Recalling the explanation for the previous curves, in this case we observe the higher tendency of the mobile host to change roads if the current road is congested (higher values of $x_i^C(t)$); by contrast, the probability of remaining on the same road is also greater if the road is shorter. This last concept is also theoretically demonstrated below. At this point, for a final definition of the elements of $M^C(t)$, we recall the expressions from Equations (10) to (13), and define the following:

$$m_{i,j}^{C}(t) = \begin{cases} \frac{1}{|r_i|_C} \cdot \left[v_i^{max} - \theta_i^C(t) \cdot (v_i^{max} - v_i^{min})\right], i = j \\ \left[(1 - m_{i,i}^C(t)\right] \cdot \frac{1 - \theta_j^C(t)}{\sum_{k=1}^{K_i} 1 - \theta_i^C(t)}, i \neq j, r_i \overset{\text{C}}{\otimes} r_j \\ 0 \;\; else. \end{cases} \tag{14}$$

Figure 8 plots the relationship between the maximum permitted speed v_i^{max}, road density θ_i^C and road length $|r_i|_C$. As expressed by the previous equations, it is clear that the lower the road density, the higher the probability of remaining on the same road, whereas for longer roads the probability of remaining on them decreases (higher probability of crossroads). We observe that the maximum permitted speed (and, hence, the average speed) plays an important role in mobility dynamics: when a host is able to move at a higher speeds, it is less willing to execute a turn and change roads, therefore the term $m_{ii}^C(t)$ increases.

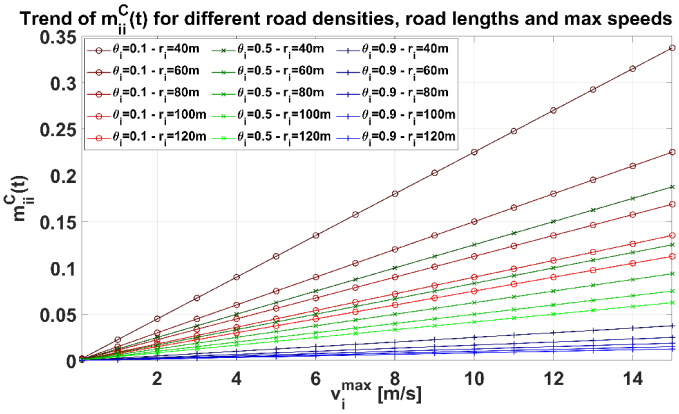

Fig. 8: Trend of $m_{i,i}^C(t)$ for different values of $|r_i|_C$, $\theta_i^C(t)$ and v_i^{max}.

D. Geometric Relationships Between Coverage Areas, Road Lengths and Crossroads Probabilities

Although an intuitive concept, let us now demonstrate that for each road $r_l \in RD$ (partially contained in C), a variation of the coverage radius from r to R, where $R = r + \Delta r$, results in an increase $\Delta road$ of the road length inside the coverage region C, where $\Delta road > 0$. For simplicity, we refer to Figure 9 and the road segment $s_{l_i}(p_1, p_2)$ of road r_l and its related equation $y_{s_{l_i}} = m_{s_{l_i}} x + q_{s_{l_i}}$. Without loss of generality, only the first quadrant of the Cartesian plane is considered, and hence, $\alpha \in (0, \pi/2]$, where $\alpha = tan^{-1}(m_{s_{l_i}})$. Beginning with a coverage area equal to C_r (the notation indicates a coverage area C with radius r), it is clear that the sub-segment (p_1, B) represents the contained part of the segment s_{l_i}. When the coverage radius is extended to R, then the sub-segment contained in C_R is (p_1, D). In this case, the obtained increase in the segment length is $|(p_1, D)| - |(p_1, B)| = \overline{BD} = \Delta segment$. In a demonstration of this concept, it is important to note that:

$$\Delta segment \cdot \sin(\alpha) + r \cdot \sin(\alpha_1) = R \cdot \sin(\alpha_2), \tag{15}$$

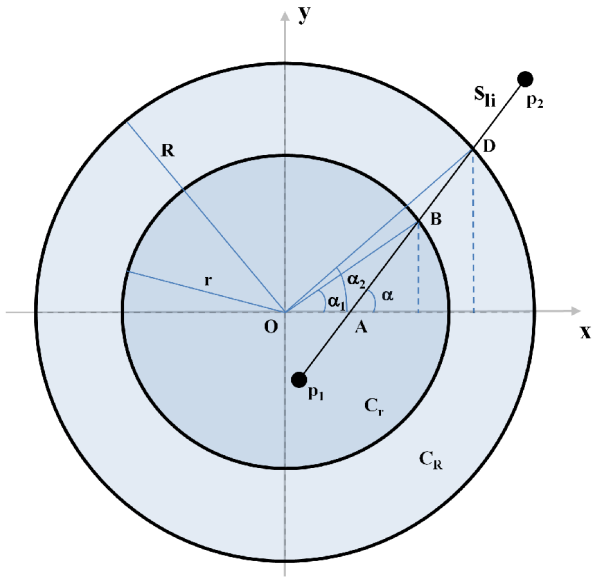

Fig. 9: Geometrical representation of the main relations between coverage region and segment length.

where α_1 and α_2 are the angles with the x-axis determined by segments $\overline{OB}$ and $\overline{OD}$, respectively. Given that $\alpha \in (0, \pi/2]$, the relation $\alpha_2 \geq \alpha_1$ is always satisfied, and subsequently, $\alpha_2 = \alpha_1 + \Delta\alpha$, where $\Delta\alpha \geq 0$. Therefore,

$$\Delta segment = \frac{R \cdot \sin(\alpha_1 + \Delta\alpha) - r \cdot \sin(\alpha_1)}{\sin(\alpha)}. \tag{16}$$

Given that the quantity $\sin(\alpha)$ is greater than or equal to 0 in $(0, \pi/2]$, $\Delta segment$ is positive if and only if the following relation is satisfied:

$$\frac{\sin(\alpha_1 + \Delta\alpha)}{\sin \alpha_1} > \frac{r}{R}. \tag{17}$$

However, the first ratio is always greater than 1 since the sin function increases in the range $(0, \pi/2]$, while the second ratio is always less than 1 since $r < R$. The relation in Equation (17) is therefore always satisfied, and the term $\Delta segment$ is always positive. The result obtained in Equation (17) is demonstrated for one segment only, but it can be generalized to an entire road or a set of roads. Clearly, considering all the combinations illustrated in Figure 3, a coverage radius that increases from r to R may not necessarily affect all segments of the roads on the example map. In fact, if a road r_i is completely contained in C, the variation of the coverage range will not affect $|r_i|_C$. In other cases (r_i is partially contained or is a chord), the variation might not affect all road segments, but for at least one segment the relation expressed in Equation (17) will be valid. In the latter case, where r_i is external to C, the variation may result in adding r_i to the coverage of C. This is valid for all roads contained within the coverage area

C. Therefore, for a given road $r_i \in RD$, we use the following the notation in Equations (8) and (9):

$$\Delta road_i \stackrel{\text{def}}{=} \sum_{j=1}^{J_C} \Delta segment_j + \sum_{k=1}^{K_C} \Delta segment_k, \tag{18}$$

and

$$\overline{\Delta road}_C \stackrel{\text{def}}{=} \frac{1}{L_C} \sum_{l=1}^{L_C} \Delta road_l, \tag{19}$$

where, $\Delta road_i$ represents the increase in $|r_i|_C$ when the coverage radius increases by δr, J_C is the number of segments of r_l partially contained in C_r, and K_C is the number of segments of r_i that are chords on C_r. Equation (19) calculates the average increase for all roads. At this point, beginning with the expression in Equation (14), we examine the effects of increasing the coverage radius from r to $R = r + \Delta r$ on $m_{i,i}^{C}(t)$ elements and the consequent increase in length $\Delta road_i$ for r_i. Examining $m_{i,i}^{C}(t)$ from Equation (14), it follows that for any road r_i covered by C_r, it is valid to write (neglecting the modulus symbol to indicate the length of r_i):

$$\begin{aligned} m_{i,i}^{C_r}(t) - m_{i,i}^{C_R}(t) = \frac{1}{|r_i|_{C_r}} \cdot \\ \cdot \left[v_i^{max} - \frac{x_i^{C_r}(t) \cdot \overline{\Gamma} \cdot (v_i^{max} - v_i^{min})}{|r_i|_{C_r}} \right] + \\ - \frac{1}{|r_i|_{C_R}} \cdot \left[v_i^{max} - \frac{x_i^{C_R}(t) \cdot \overline{\Gamma} \cdot (v_i^{max} - v_i^{min})}{|r_i|_{C_R}} \right]. \end{aligned} \tag{20}$$

Consequently, where $\Delta v_i = v_i^{max} - v_i^{min}$ (independently of the extended coverage radius), $R = r + \Delta r$ (from Equations (16), (18)), $x_i^{C_R}(t) = \gamma \cdot x_i^{C_r}(t)$ (where $\gamma \geq 1$ is a magnifying factor), and $|r_i|_{C_r} = r_i^{C_r}$ (for simplicity of notation), Equation (20) becomes:

$$\begin{aligned} \frac{v_i^{max}}{r_i^{C_r}} - \frac{x_i^{C_r}(t) \cdot \overline{\Gamma} \cdot \Delta v_i}{(r_i^{C_r})^2} - \frac{v_i^{max}}{r_i^{C_r} + \Delta r} + \frac{\gamma \cdot x_i^{C_r}(t) \cdot \overline{\Gamma} \cdot \Delta v_i}{(r_i^{C_r} + \Delta r)^2} = \\ = v_i^{max} \cdot \left(\frac{1}{r_i^{C_r}} - \frac{1}{r_i^{C_r} + \Delta r} \right) + x_i^{C_r}(t) \cdot \overline{\Gamma} \cdot \Delta v_i \cdot \\ \cdot \left[\frac{\gamma}{(r_i^{C_r} + \Delta r)^2} - \frac{1}{(r_i^{C_r})^2} \right]. \end{aligned} \tag{21}$$

All the quantities in Equation (21) are positive, except for the difference in the square parentheses; therefore, under the above-mentioned assumptions and notations, $m_{i,i}^{C_r}(t) - m_{i,i}^{C_R}(t) > 0$ if and only if, from Equation (21):

$$\frac{v_i^{max}}{x_i^{C_r}(t) \cdot \overline{\Gamma} \cdot \Delta v_i} > \frac{2 \cdot r_i^{C_r} + \Delta r}{r_i^{C_r} \cdot (r_i^{C_r} + \Delta r)} - \frac{(r_i^{C_r})^2 \cdot (\gamma - 1)}{\Delta r \cdot (r_i^{C_r} + \Delta r) \cdot r_i^{C_r}}. \tag{22}$$

To simplify the expression in the inequality (22), the increase in radius can be expressed as a percentage of the starting coverage radius r, i.e, $\Delta r = \beta \cdot r$, with $0 < \beta < 1$. Substituting into (22) and simplifying the second member, we write

$$\frac{v_i^{max} \cdot r_i^{C_r}}{x_i^{C_r}(t) \cdot \overline{\Gamma} \cdot \Delta v_i} > \frac{\beta + 2}{\beta + 1} + \frac{1 - \gamma}{\beta \cdot (\beta + 1)}. \tag{23}$$

One of the most important observations from the present study is that if this relationship is satisfied, then the probability of remaining on the same road r_i ($m_{i,i}^{C}(t)$) decreases if the coverage is extended. By contrast, if the coverage radius is reduced (i.e., shortening the length of the roads in question), then the evaluated probability of remaining on the current road under the coverage of C increases. Before it is assumed that the inequality (23) is always satisfied, some analysis must be done. Recalling that $\Delta r = \beta \cdot r_i^{C_r}$, where $(0 < \beta < 1)$, and $x_i^{C_R}(t) = \gamma \cdot x_i^{C_r}(t)$, where $\gamma \geq 1 \implies \Delta x_i(t) = x_i^{C_r}(t) \cdot (\gamma - 1)$, it is easy to verify that the relationship between γ and β derives from

$$\frac{\Delta r}{r} \cdot \left[\Delta r \cdot \theta_i^{C_r}(t) \right] = (\gamma - 1) \cdot x_i^{C_r}(t), \tag{24}$$

where the ratio $(\Delta r)/r$ is the probability of finding new cars on the new road segment under coverage, $\Delta r \cdot \theta_i^{C_r}(t)$ is the number of probable vehicles on the new road segment, and $(\gamma - 1) \cdot x_i^{C_r}(t)$ is the increase in vehicle numbers. The latter expression can be simplified to

$$\beta^2 = \frac{\gamma - 1}{\gamma}, \; or \; \gamma = \frac{1}{1 - \beta^2}, \tag{25}$$

which concurs with the definitions of α and β. Substituting the expression of γ into (23), it becomes

$$\frac{v_i^{max} \cdot r_i^{C_r}}{x_i^{C_r}(t) \cdot \overline{\Gamma} \cdot \Delta v_i} > \frac{1}{\beta + 1} \cdot \left[\frac{\beta^2 \cdot (\beta + 2) - 2}{\beta^2 - 1} \right]. \tag{26}$$

By indicating the second member of inequality (26) with the Lower Bound (LB(β)), we observe that $\lim_{\beta \to 0^+} LB(\beta) = 2$ and $\lim_{\beta \to 1^-} LB(\beta) = -\infty$. Figure 10 illustrates the trend

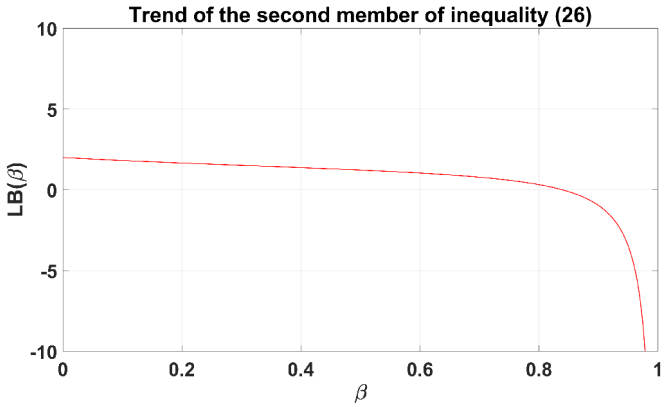

Fig. 10: Trend of the second member in inequality (26).

of $LB(\beta)$, where the left and right limits are respected, the function is concave, and the trend is decreasing (zero is reached for $\beta = 0.839282$, then the function assumes negatives values). Now, the first member of inequality (26) must be evaluated. A preliminary view of the effect on the trend of

the first member of the inequality (26) is given in Figure 11, and a deeper analysis is presented in Section IV. Figure 11

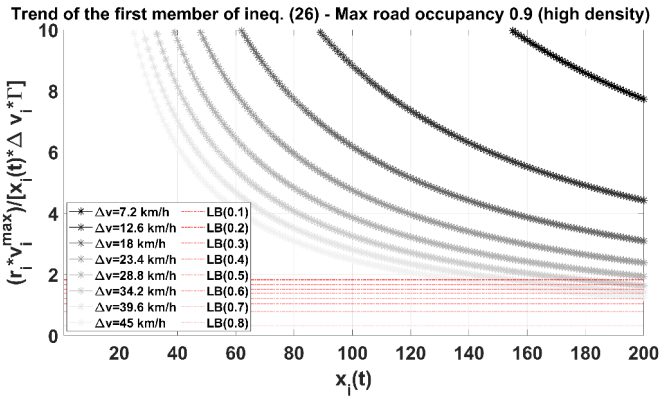

Fig. 11: Trend of the first member of the inequality (26), for $r_i^{C_r}$=1km and $max\{\theta_i^{C_r} = 0.9\}$.

shows the values of $LB(\beta)$ in the function of β (red dotted lines) and the values of the first member of (26) for different values of $x_i^{C_r}$ and Δv_i (the graph is magnified on the y axis for easier readability), for a fixed value of ($r_i^{C_r}$) of 1 km ($\overline{\Gamma}$ = 4.5 m). At a lower Δv_i, the first member is always greater than $LB(\beta)$, independently of the increasing radius parameter β. For denser roads, the trend of the first member is decreasing, becoming significant at a higher Δv_i (e.g., r_i is not properly straight). When the road is highly populated with vehicles and Δv_i is high, the probability of changing roads increases (independently of the considered parameters).

E. Next Cell Prediction Error and Coverage Radius: What Happens in Small Cells?

From the above, we now derive the final and possibly most useful relationship regarding the qualitative analysis of mobility prediction errors in cellular networks. Using the equations above, we derive the $L_{C_r} x L_{C_r}$ matrix $M^{C_r}(t^*)$ for each time instant t^* and where each element $m_{i,j}^{C_r}(t^*)$ represents the probability of passing from road r_i to road r_j in the time interval $[t^*, t^* + \Delta t]$ ($\Delta t << 0$). Let us examine the i-th row of $M^{C_r}(t^*)$, $M_i^{C_r}(t^*)$, at the time instant t^*.

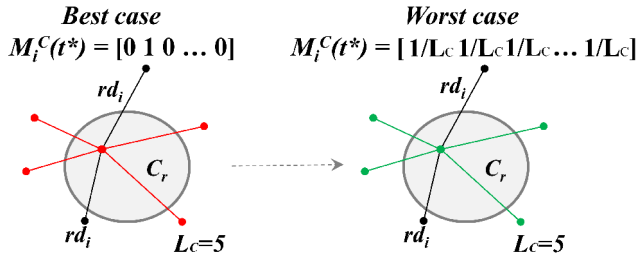

Fig. 12: Example of the best/worst prediction scenario.

Figure 12 illustrates the best and worst cases of the composition of $ri \in RD$ in relation to the probability of changing roads. In the best (ideal) case, only one element of the i-th row is different from zero (ideally, only $m_{i,i}^{C_r}(t^*) = 1 \neq 0$), namely each vehicle remains on a particular road (i.e., roads are not reachable from r_i). In the worst case, there is no preferred direction of movement, and the choice of road is random (i.e., a discrete uniform distribution is obtained for each element of the row; L_{C_r}=5 in the illustration). To calculate the probability of an error in predicting the next road/cell, independently of any predictor for determining the future directions of a moving host inside a cell [8], [9], we recall the general expression of the variance σ_i^2 for the elements of $M_i^{C_r}(t^*)$:

$$\sigma_i^2(t^*) = \frac{1}{L_{C_r} - 1} \cdot \sum_{j=1}^{L_{C_r}} \left[\mu_i(t^*) - m_{i,j}^{C_r}(t^*)\right]^2, \tag{27}$$

where

$$\mu_i(t^*) = \frac{1}{L_{C_r}} \cdot \sum_{j=1}^{L_{C_r}} m_{i,j}^{C_r}(t^*) = \frac{1}{L_{C_r}}, \tag{28}$$

since $\sum_{j=1}^{L_{C_r}} m_{i,j}^{C_r}(t^*)$ = 1 by definition, $\forall i = 1, ..., L_{C_r}$ (Eq. (14)). It is clear that in the best case, $\sigma_i^2 = 1/L_{C_r}$, while in the worst case, $\sigma_i^2 = 0$ (the row elements are all the same). In the best case, the prediction error is null (i.e, only one possibility to exit the coverage cell on the same road). In the worst case, the prediction error reaches its maximum since all possible events have the same probability (i.e., uniform distribution). To derive and analyze the trend of σ_i^2 (in addition to the prediction error), we assume that (among all the possible combinations)

$$M_i^{C_r}(t^*) = \left[\frac{\Delta L_C}{L_{C_r} - 1}, \frac{\Delta L_C}{L_{C_r} - 1}, ..., 1 - \Delta L_C, \frac{\Delta L_{C_r}}{L_C - 1}, ..., \frac{\Delta L_C}{L_{C_r} - 1}\right], \tag{29}$$

where $0 \leq \Delta L_C \leq 1 - 1/L_{C_r}$. Therefore, when $\Delta L_C \to 0^+, \sigma_i^2(t^*) \to 1/L_{C_r}$ and when $\Delta L_C \to (1 - 1/L_{C_r})^-, \sigma_i^2(t^*) \to 0$. Using the structure of $M_i^{C_r}(t^*)$ in (29), Equation (27) becomes

$$\sigma_i^2(t^*) = \frac{1}{L_{C_r} \cdot (L_{C_r} - 1)} \cdot \left\{[1 - L_{C_r} \cdot (1 - \Delta L_C)]^2 + \right.$$
$$\left. + \frac{1}{L_{C_r} \cdot (L_{C_r} - 1)} \cdot [L_{C_r} \cdot (1 - \Delta L_C) - 1]^2 \right\}. \tag{30}$$

Figure 13 illustrates the trend as a function of ΔL_{C_r} and L_C, and notably, the shape is always the same (independent of L_{C_r}) and the limits are reached, as with Equation (30). At this point, the qualitative trend of the next-cell prediction error can be defined as a function of $\sigma_i^2(t^*)$:

$$P_{err}\left[\sigma_i^2(t^*)\right] \stackrel{\text{def}}{=} \frac{L_{C_r} - 1}{L_{C_r}} \cdot \left[1 - L_{C_r} \cdot \sigma_i^2(t^*)\right]. \tag{31}$$

Figure 14 illustrates the results for different values of $\sigma_i^2(t^*)$ and L_{C_r}.

For the second main contribution of the current study, it is demonstrated that if the radius of C_r extends to R, where $R = r + \Delta R$, then the prediction error under certain conditions also increases.

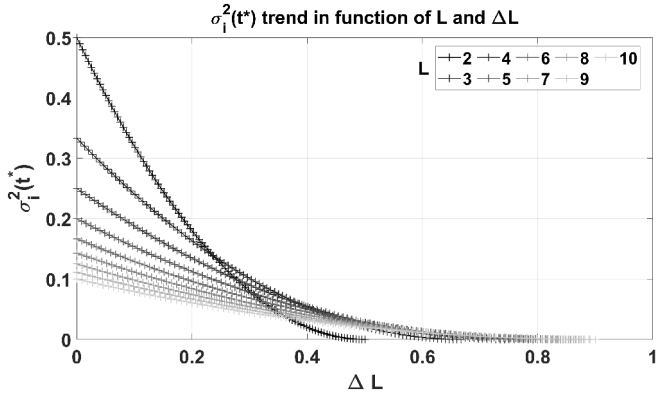

Fig. 13: Trend of $M_i^{C_r}(t^*)$ variance as a function of L_{C_r} and ΔL_C (best case to worst case).

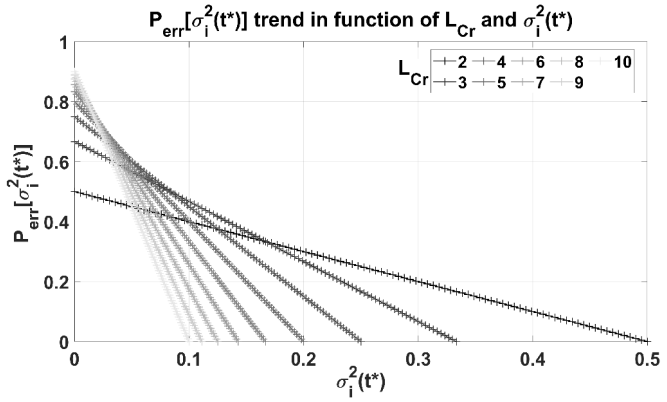

Fig. 14: Next-cell prediction error trend versus $\sigma_i^2(t^*)$ and L_{C_r}.

Theorem 1 Let us generalize the results obtained from (29) and (30). Given a coverage cell C_r, the set of RD of $|RD| = L_{C_r}$ covered roads ($L_{C_r} \geq 1$), a time instant t^*, and the values of $P_{err}[\sigma_i^2(t^*)]$ defined in (31), where $i = 1, ..., L_{C_r}$, then an extension of the coverage radius from r to R, where $R = r + \Delta R$ and assuming that L_{C_r} does not change, implies a decrease in the diagonal elements $m_{i,i}^{C_r}(t^*)$ of $M^{C_r}(t^*)$ and a consequent increase of $P_{err}[\sigma_i^2(t^*)]$ if the relationship in (26) is respected and if

$$\Delta M_i < \frac{m_{i,i}^{C_r}(t^*)}{2} \quad \forall i = 1, ..., L_{C_r}, \tag{32}$$

where ΔM_i is the increase in the i-th diagonal element of $M^{C_r}(t^*)$ on row $M_i^{C_r}(t^*)$.

Demonstration of Theorem 1: Satisfaction of the inequality (26) implies that $m_{i,i}^{C_r}(t^*) > m_{i,i}^{C_R}(t^*)$, i.e.,

$$m_{i,i}^{C_r}(t^*) = m_{i,i}^{C_R}(t^*) + \Delta M_i, \quad \forall i = 1, ..., L_{C_r}, \tag{33}$$

$0 < \Delta M_i < 1$; but $\sum_{j=1}^{L_{C_r}} m_{i,j}^{C_r}(t^*) = 1$ from (14), therefore if

$$M_i^{C_r}(t^*) = [m_{i,1}^{C_r}(t^*), m_{i,2}^{C_r}(t^*), .., m_{i,i}^{C_r}(t^*), .., m_{i,L_{C_r}}^{C_r}(t^*)], \tag{34}$$

then

$$M_i^{C_R}(t^*) = [m_{i,1}^{C_r}(t^*) + f_{i,1}(\Delta M_i), m_{i,2}^{C_r}(t^*) + f_{i,2}(\Delta M_i), ... \\ ..., m_{i,i}^{C_r}(t^*) - \Delta M_i, ..., m_{i,L_{C_r}}^{C_r}(t^*) + f_{i,L_{C_r}}(\Delta M_i)], \tag{35}$$

and

$$\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i) - \Delta M_i = 0, \tag{36}$$

where the functions $f_{i,j}(\cdot)$: $(0,1) \rightarrow (0,1)$, $j = 1, ..., L_{C_r}, j \neq i$, are strictly related to the road topology under the coverage of C_r. Now, to evaluate the trend of variances before and after extension of the radius, referring to the definition in Equation (27) and assuming that the theorem is valid, we write

$$\sigma_i^2(t*)\big|_{C_r} - \sigma_i^2(t*)\big|_{C_R} > 0, \tag{37}$$

which translates into

$$\frac{1}{L_{C_r}-1} \cdot \left\{ \left[\frac{1}{L_{C_r}} - m_{i,i}^{C_r}(t^*)\right]^2 + \right. \\ \left. + \sum_{j=1, j\neq i}^{L_{C_r}} \left[\frac{1}{L_{C_r}} - m_{i,j}^{C_r}(t^*)\right]^2 \right\} > \frac{1}{L_{C_r}-1} \cdot \left\{ \left[\frac{1}{L_{C_r}} + \right. \right. \\ \left. - (m_{i,i}^{C_r}(t^*) - \Delta M_i)\right]^2 + \sum_{j=1, j\neq i}^{L_{C_r}} \left[\frac{1}{L_{C_r}} + \right. \\ \left. \left. - (m_{i,j}^{C_r}(t^*) + f_{i,j}(\Delta M_i))\right]^2 \right\}. \tag{38}$$

Simplifying the terms and expanding the power operations, we obtain

$$\frac{1}{L_{C_r}^2} - \frac{2}{L_{C_r}} \cdot m_{i,i}^{C_r}(t^*) + [m_{i,i}^{C_r}(t^*)]^2 + \sum_{j=1, j\neq i}^{L_{C_r}} \left\{ \frac{1}{L_{C_r}^2} + \right. \\ \left. - \frac{2}{L_{C_r}} \cdot m_{i,j}^{C_r}(t^*) + [m_{i,j}^{C_r}(t^*)]^2 \right\} > \frac{1}{L_{C_r}^2} - \frac{2}{L_{C_r}} \cdot [m_{i,i}^{C_r}(t^*) + \\ - \Delta M_i] + [m_{i,i}^{C_r}(t^*) - \Delta M_i]^2 + \sum_{j=1, j\neq i}^{L_{C_r}} \frac{1}{L_{C_r}^2} - \frac{2}{L_{C_r}} \cdot \\ [m_{i,j}^{C_r}(t^*) + f_{i,j}(\Delta M_i)] + [m_{i,j}^{C_r}(t^*) + f_{i,j}(\Delta M_i)]^2, \tag{39}$$

and again, simplifying the same terms and expanding the power operations, we have

$$\frac{2}{L_{C_r}} \cdot \Delta M_i - 2 \cdot m_{i,i}^{C_r}(t^*) \cdot \Delta M_i + \Delta^2 M_i - \sum_{j=1, j\neq i}^{L_{C_r}} \frac{2}{L_{C_r}} \cdot f_{i,j}(\Delta M_i) + 2 \cdot \sum_{j=1, j\neq i}^{L_{C_r}} m_{i,j}^{C_r}(t^*) \cdot f_{i,j}^2(\Delta M_i) + \sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}^2(\Delta M_i) < 0, \tag{40}$$

which by Equation (36) becomes

$$\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}^2(\Delta M_i) + 2 \cdot \sum_{j=1, j\neq i}^{L_{C_r}} m_{i,j}^{C_r}(t^*) \cdot f_{i,j}^2(\Delta M_i) < < 2 \cdot m_{i,i}^{C_r}(t^*) \cdot \Delta M_i - \Delta^2 M_i. \tag{41}$$

Equation (41) can be rewritten as

$$\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}^2(\Delta M_i) + 2 \cdot \sum_{j=1, j\neq i}^{L_{C_r}} m_{i,j}^{C_r}(t^*) \cdot f_{i,j}^2(\Delta M_i) + + \left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i) \right]^2 < 2 \cdot m_{i,i}^{C_r}(t^*) \cdot \Delta M_i, \tag{42}$$

and, because $0 < f_{i,j}(\Delta M_i) < 1$, $\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}^2(\Delta M_i) < \left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i)\right]^2$ and $\sum_{j=1, j\neq i}^{L_{C_r}} m_{i,j}^{C_r}(t^*) \cdot f_{i,j}^2(\Delta M_i) < \sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}^2(\Delta M_i) < \left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i)\right]^2$, the first two terms in Equation (42) can be substituted with $\left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i)\right]^2$, and we obtain

$$\left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i) \right]^2 + 2 \cdot \left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i) \right]^2 + + \left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i) \right]^2 < 2 \cdot m_{i,i}^{C_r}(t^*) \cdot \Delta M_i, \tag{43}$$

which leads to

$$4 \cdot \left[\sum_{j=1, j\neq i}^{L_{C_r}} f_{i,j}(\Delta M_i) \right]^2 = 4 \cdot \Delta^2 M_i < 2 \cdot m_{i,i}^{C_r}(t^*) \cdot \Delta M_i, \tag{44}$$

and hence,

$$\Delta M_i < \frac{m_{i,i}^{C_r}(t^*)}{2}. \tag{45}$$

IV. Extensive Numerical Analysis

In this section the theoretic proposal described above is validated and assessed. To obtain a detailed set of simulations, the following tasks were performed:

- Several maps were extracted from Open Street Map GUI [33], exported in XML format, and used in SUMO [34] to generate real traffic patterns;
- Using a dedicated MATLAB script, all maps were overlaid with a cellular map of circular cell shapes that can be configured in relation to the surrounding architecture. To simplify the readability of the maps and not show cell overlaps, Figures 15, 16, 17 depict hexagonal instead of circular shapes;
- The same MATLAB script was extended to evaluate the number of roads covered in each cell and to determine the appropriate statistics for road occupancy, variance, and theoretic prediction error;
- As examples of real predictive approaches, the Direction Aware Static Predictor (DASP) described in [36] and the User Mobility Profile with Trace Record Matrix (UMP-TRM) from [37] were applied to compare the theoretic and practical results.

The main results are presented below. First, several GAs were selected from maps of Rome (Italy), New York (USA), and Melbourne (Australia). Their dimensions were fixed to approximately 1050 m x 1600 m, which represents almost 1.7 km^2 in area. Each city represents a different road topology: urban (high road density), residential (medium road density), and suburban (low road density), as well as three coverage radii r of 50 m, 75 m and 100 m. The coverage system's origin is set to the bottom left hand corner.

Figures 15, 16, 17 illustrate the coverage maps (urban cases) for different values of r. Red cells were not evaluated for statistics because vehicle movements in these cells do not terminate correctly due to map boundaries. Table II summarizes the Number of Cells (NCs) and value of L_{C_r} for each cell in the three maps above.

The summary in Table II indicates that the average number of roads covered by a cell increases with larger radii: at a radius of 50 m (Rome), the average $L_{C_{50}}$ is 1.7682; at 75 m (New York), the average $L_{C_{75}}$ is 2.4583; at 100 m (Melbourne), the average $L_{C_{75}}$ is 2.925. Let us instead examine the mean L_{C_r} values for each map (Rome, New York, Melbourne) and all possible coverage radii; Figure 18 plots the L_{C_r} pdfs and indicates that the probability of finding an *empty* cell (no roads covered) decreases if the radius is increased (empty cells, $L_{C_r} = 0$, also evident in the results given in Table I for r = 50 m). The same trend can be observed in the probability for only one road covered, while $L_{C_r} = 2$ (independently of r) demonstrates an almost constant trend; by contrast, the probability of covering a greater number of roads (5 or 6) increases with larger radii. To observe the result when a certain number of vehicles is present in each GA, SUMO was configured to simulate a varying number of moving vehicles, using the $RandomTrips$ Python script

This article has been accepted for publication in IEEE Transactions on Vehicular Technology. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVT.2024.3453450

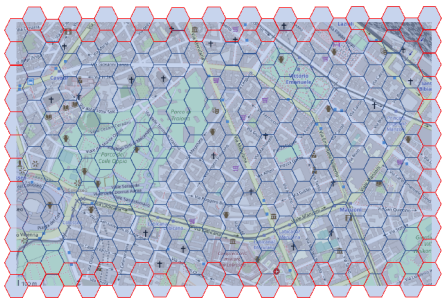

Fig. 15: Coverage map of Rome, r = 50 m.

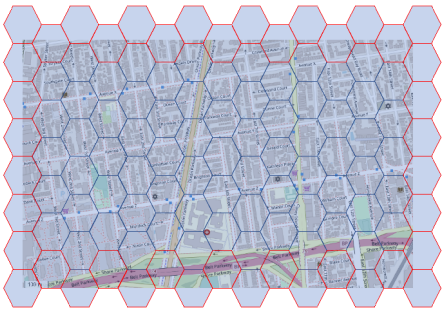

Fig. 16: Coverage map of New York, r = 75 m.

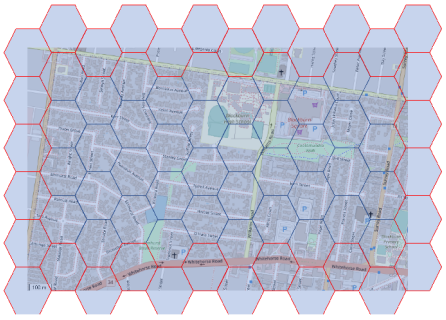

Fig. 17: Coverage map of Melbourne, r = 100 m.

TABLE II: Number of covered roads in each cell: Fig. 15 (Rome, dense, 220 cells, 50 m); Fig. 16 (New York, dense, 66 cells, 75 m); Fig. 17 (Melbourne, dense, 40 cells, 100 m).

1	5	2	2	1	1	2	2	3	3	2	2	2	2	1	2	1	2	1	1	1	2
3	3	3	2	1	2	6	2	4	2	2	1	2	1	0	1	2	2	2	2	2	3
1	1	1	2	1	1	0	2	2	0	1	3	2	1	1	1	1	3	1	2	2	1
2	4	3	1	1	1	3	1	0	0	2	2	2	2	1	2	2	2	2	2	2	2
1	2	0	1	2	1	2	3	1	2	1	1	2	2	2	2	1	1	2	1	2	4
3	2	2	1	0	1	5	1	3	3	2	1	2	1	1	1	2	3	1	1	2	1
1	1	2	2	3	3	2	3	2	3	1	2	2	3	1	1	2	2	1	1	1	1
0	0	1	1	2	2	3	4	4	3	2	3	1	2	2	3	2	2	3	1	4	2
1	1	3	3	3	2	2	2	1	1	1	3	2	1	1	1	0	1	1	1	2	0
2	2	2	2	1	3	2	3	2	2	3	2	0	1	1	1	1	2	1	2	1	0

a)

1	3	3	2	3	3	3	2	3	2	3	1
3	2	2	2	3	2	2	1	2	2	3	2
3	2	2	2	4	2	3	2	2	1	1	4
3	3	2	2	3	2	3	3	3	3	2	1
2	3	2	3	5	1	2	3	1	3	3	4
	1		2		4		3		3		3

b)

5	4	4	2	1	1	3	6	3	4
4	3	5	3	2	2	1	1	3	2
4	4	2	4	2	3	2	3	4	2
4	2	3	2	2	3	2	6	2	2

c)

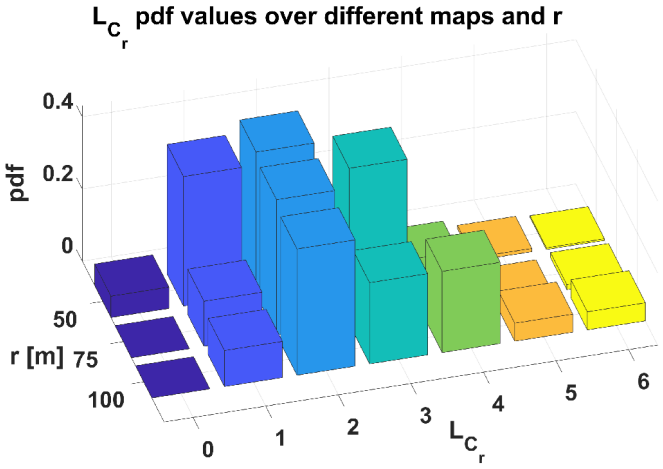

Fig. 18: Trend of the L_{C_r} pdf for different r, over the three considered GAs.

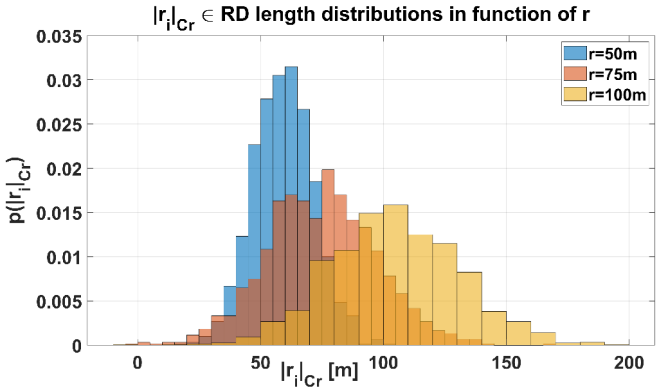

Fig. 19: Road length probability densities for different coverage extents r.

with the attribute *maxSpeed* equal to 13.9 m/s (50 km/h) for urban scenarios, 19.4 m/s (70 km/h) for residential scenarios, and 25 m/s (90 km/h) for suburban scenarios. Simulation time (per run) was fixed to 600 s (attribute e), and only cars were tagged as mobile nodes, with the standard attribute *vClass* equal to *passenger* (bitmask 6) and the *length* attribute set to Γ = 4.5 m. Figure 19 graphs the distributions of road lengths (not external to C_r, as in Def. 8, sub-section III-B), for the three values of r and independently of the mobile environment (urban, residential, suburban). The Kolmogorov-Smirnov (KS) test [35] was applied to test goodness of fit and evaluate the correctness of a Normal approximation of these distributions. In this case, road length samples were standardized and compared to a standard normal distribution, always satisfying the test against the null hypothesis. Notably, the higher the radius, the higher the variance and the lower the number of different roads covered by a cell, because if the GA is the same and the radius is increased, then the probability of finding segments belonging to the same road covered by different cells decreases. The mean and standard deviation values for the three Gaussian distributions are μ_{50} = 60.126 m, σ_{50} = 12.269 m, μ_{75} = 75.126 m,

σ_{75} = 21.877 m, μ_{100} = 105.11 m, σ_{100} = 25.367 m. The

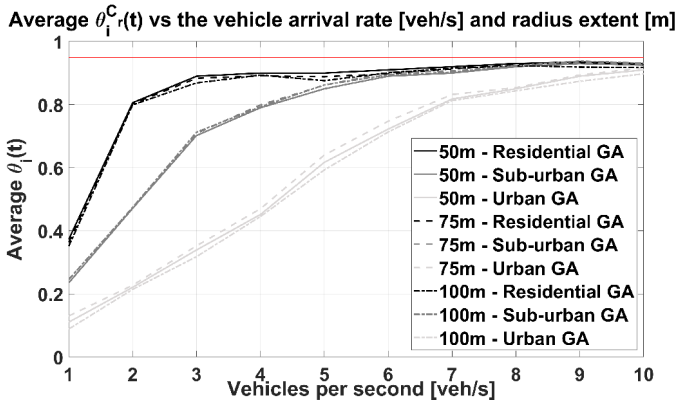

Fig. 20: Average road occupancy after a simulation time of 600 seconds for three different GA categories and different values of r.

results in Figure 20 assist in determining which arrival frequency to select in SUMO to avoid road congestion, however this phenomenon is also very important in assessing the validity of the proposed theoretic model. The information in Figure 20 indicates when congestion may occur in a typical GA scenario and coverage extent: while the GA scenario affects the congestion slope (in terms of road occupancy), the coverage extent has an almost neutral effect. The *RandomTrips period* parameter in SUMO was set to a variable arrival rate, expressed in vehicles per second [veh/s]. Urban environments have a more complicated and denser road topology and can accommodate more vehicles; this capacity decreases in sparser topologies (e.g., residential or suburban GAs, although the maximum speed is higher). Using this data, the maximum arrival rates were set to 3 [veh/s] for residential GAs, 6 [veh/s] for suburban GAs, and 10 [veh/s] for urban GAs. To verify the theoretic trend of $m_{i,i}^{C_r}(t)$ as defined in Equation (14), several simulations were compiled for the arrival rate, and hence also the function of $\theta_i^{C_r}(t)$, and the results were averaged for the overall set of cells in the system; Figure 21 plots the main trends. It is clear from these results that the values of $m_{i,i}^{C_r}(t)$ are higher for sparser environments (suburban or residential), and the resulting trends confirm that travel along longer roads indicates a greater probability of vehicles changing from those roads. The curves in Figure 20 confirm the expected result that the probability of remaining on the same road decreases as congestion on the road increases (high arrival rates), independently of the specific GA. In addition, in scenarios with denser roads, the probability also decreases as a consequence of a greater number of connected roads (as in Definition 9 $r_i \overset{C}{\otimes} r_j$). Regarding the coverage range, the simulations confirm that $m_{i,i}^{C_r}(t)$ decreases at larger coverage radii; this suggests that under the selected simulation parameters, the inequality (26) should be always verified. In fact, the results in Figure 22 indicate that, by evaluating the first member of (26) and averaging it over the total cells of the system, the values of $LB(\beta)$ are always respected (β = 0.5

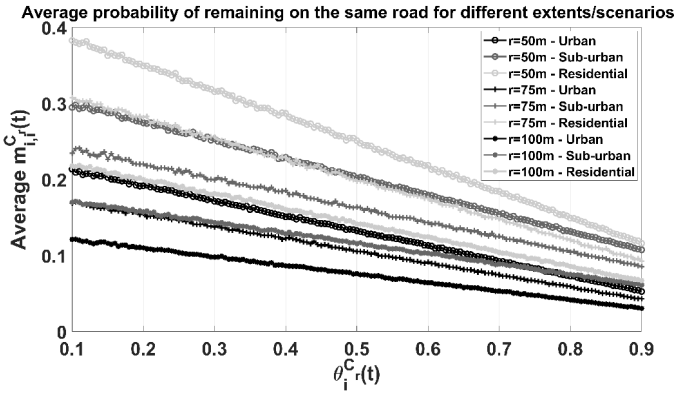

Fig. 21: Average $m_{i,i}^{C_r}(t)$ in a simulation time of 600 s for three different GA categories (urban, v_i^{max} = 13.9 m/s; suburban, v_i^{max} = 19.4 m/s; and residential, v_i^{max} = 25 m/s) and different radii

.

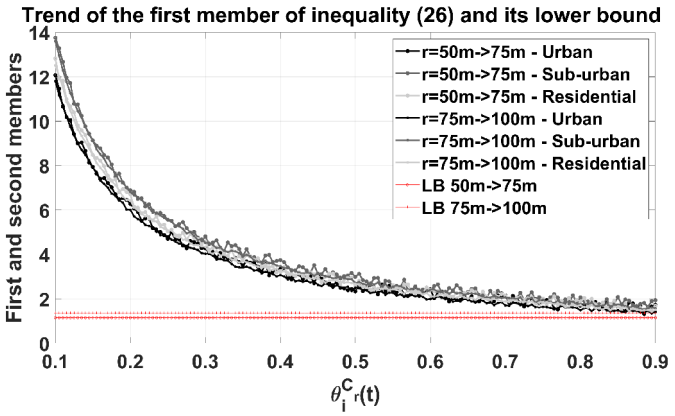

Fig. 22: Average trend of the first member of inequality (26) and related lower bounds (LBs).

for an increase in r from 50 m to 75 m, and β = 0.33 for an increase in r from 75 m to 100 m). Simulations for a varying arrival rate [veh/s] were performed based on the GA scenario and the limits derived from the results in Figure 20. The results were then normalized for the function $\theta_i^{C_r}(t)$ to adjust each scenario to a comparable level (according to the simulations, congestion occurs at different times as a result of the differing numbers of roads in each GA). Finally, simulations were performed to analyze the relationship in (45), providing an important result in the present study. We now examine the diagonal elements of $M^{C_r}(t)$. The averaged results (obtained from the dynamics of each cell) are clearly shown. In the graphs, ΔR is set to 15 m; for the GAs, the assumption of a constant L_{C_r} is valid if $\Delta R < 16.4$ m. Figure 23 graphs the quantities involved in Equation (45), with a change in the coverage extent from r = 50 m to R = 65 m, i.e., ΔR = 25 m. Only urban areas were analyzed, with an arrival rate of 1 veh/s ($\theta(t)$ at approx. 0.1 in each of the three GAs, as in the results in Figure 20). The statistics were classified according to the function of the number of roads covered (L_{C_r}) by each cell.

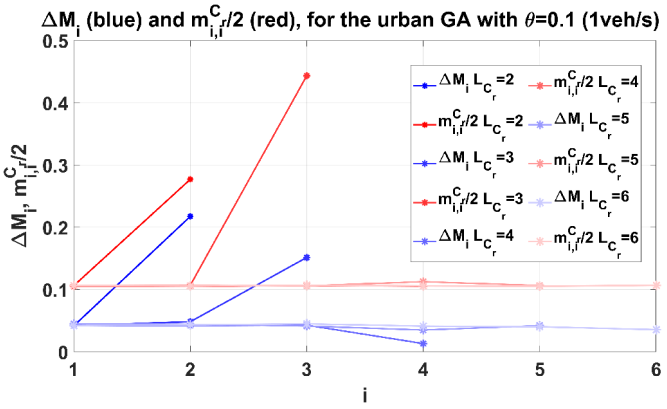

Fig. 23: Average trend of the terms in inequality (45), with a change in the coverage extent from 50 m to 65 m.

The horizontal axis represents the i-th index of the related diagonal element; for example, $L_{C_r} = 2$ $i = 1, 2$, for $L_{C_r} = 5$ $i = 1, 2, 3, 4, 5$. It can be seen that the red upper bounds ($m^{C_r}_{i,i}/2$) across the average are always respected, given that the roads are never congested during the simulation time. For reasons of limited space in the graph, the results for higher $\theta(t)$ are not shown; however, it was observed that if $\theta(t)$ increases to 0.6 (an arrival rate of 5 veh/s), the red upper-bound is not always respected in the average. A worse trend is observed when $\theta(t)$ increases to 0.85 (arrival rate of 8 veh/s). For a complete illustration, the worst cases for residential and suburban GAs are also shown. Figures 24 and 25 indicate

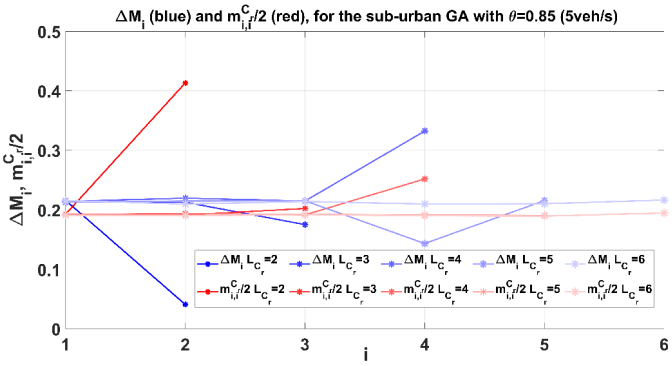

Fig. 24: Average trend of the terms in inequality (45), with a change in the coverage extent from 50 m to 65 m.

that in all cases the inequality (45) is not respected across the average. In order to compare the trend of the prediction error in a real next-cell predictive model with the trend from the proposed theoretic model, we take into account Equation (31), in particular, the DASP algorithm of [36], without considering the features of the NSIS protocol and the MIP service class [36] and implementing only the algorithm with the aim of evaluating the next-cell prediction for any active mobile host. The same simplification was applied to the UMP-TRM (UMP) of [37], which has implemented for the purpose of comparing the proposed idea. Only the first handover event is taken into

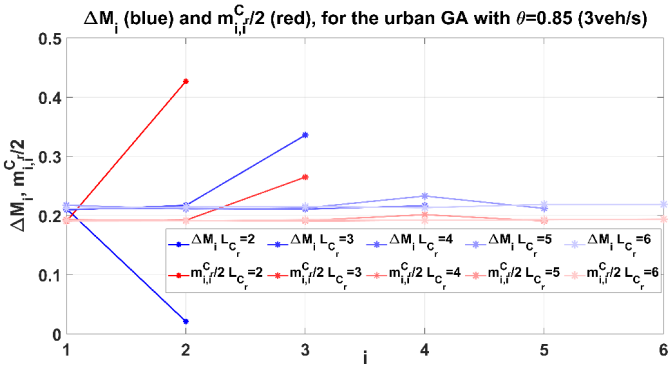

Fig. 25: Average trend of the terms in inequality (45), with a change in the coverage extent from 50 m to 65 m.

account, predicting only one cell (C_{i_1}=1 for DASP and UMP). Figure 26 shows the performance obtained with DASP, UMP and the proposed model, averaged for the three urban maps and cells (r = 50 m, 75 m, and 100 m), versus road occupancy $\theta(t)$ (which varies with the vehicle arrival rate from 1 veh/s to 10 veh/s). It is clear that the prediction error increases with larger coverage extents and for higher values of $\theta(t)$. The green and blue lines in the graph respectively represent the DASP and UMP prediction errors for the first handover of the active mobile hosts; the red lines represent an evaluation of the quantity P_{err} (Eq. (31)) during simulation. For complete

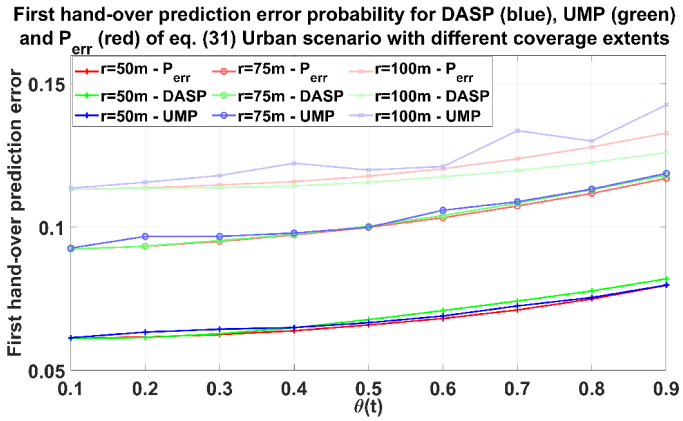

Fig. 26: Prediction error for DASP, UMP-TRM, and the proposed model (urban scenario) versus $\theta(t)$ and r.

results, the curves for residential and suburban GAs are also shown (Figs. 27, 28). The trends in these graphs are the same as in the graph in Figure 26, but the values generally decrease because of the sparser topologies of the residential and suburban scenarios.

V. Conclusions, Discussion and Future Works

The current study proposes a novel, general model for analyzing the trend of next-cell prediction errors as a function of coverage cell extent in 5G and beyond vehicular cellular networks. Using a theoretic model, the study demonstrated the relationship between radius extent and prediction error for

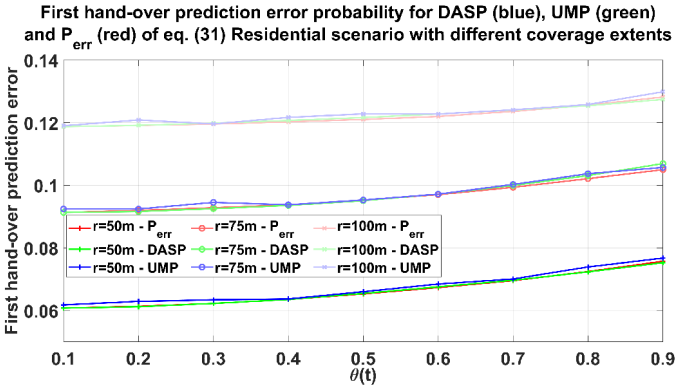

Fig. 27: Prediction error for DASP and the proposed model (residential scenario), versus $\theta(t)$ and r.

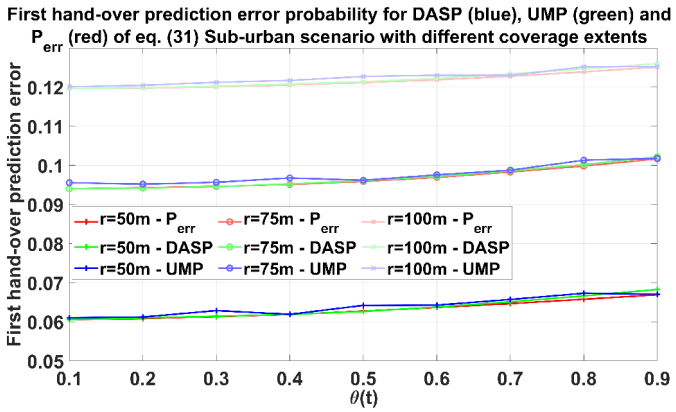

Fig. 28: Prediction error for DASP and the proposed model (suburban scenario) versus $\theta(t)$ and r.

certain conditions and analyzed simulations of real geographic areas of different character covered by small-cell network architectures. The relationship between increasing accuracy and shorter coverage extent was investigated and described in detail. The results of the study indicate that proposed model is suitable for both academic research (the application of machine learning, artificial intelligence and predictive models in 6G core architecture is a popular research area) and industrial application (ICT companies can obtain next-cell prediction accuracy data for road topologies in areas with network coverage to assist with creating appropriate link budgets for small-cell deployment and performance optimizations). Future work will involve modifying the model by removing some assumptions and incorporating additional traffic parameters, such as the number of channels allocated to mobile hosts or the time spent in a cell before handover.

VI. Acknowledgements

This work was supported by the European Union under the REFRESH project – Research Excellence For REgion Sustainability and High-tech Industries, ID No. CZ.10.03.01/00/22003/0000048, of the European Just Transition Fund. I would like to express deep gratitude to my brother Francesco for supporting me in the past and always being a source of inspiration, my wife Alessandra for supporting me everyday, and Professor Miroslav Voznak for always believing in me. Map data is under copyright of OpenStreetMap contributors and available at https://www.openstreetmap.org.

References

[1] T. Pögel and L. Wolf, "Optimization of GSM/UMTS Inter-System Handover times provided by Connectivity Maps," in Proceedings of *International Conference on Connected Vehicles and Expo (ICCVE)*, 2013, pp. 641-646, doi: 10.1109/ICCVE.2013.6799870.

[2] S.M. Asad, *et al.*, "Mobility Management in the Applications of 5G and Beyond: A Handover Skipping Topology Analysis," in Proceedings of *4th Global Power, Energy and Communication Conference (GPECOM)*, 2022, pp. 650-654, doi: 10.1109/GPECOM55404.2022.9815787.

[3] T. Liu, *et al.*, "Uplink and Downlink Decoupled 5G/B5G Vehicular Networks: A Federated Learning Assisted Client Selection Method," in *IEEE Transactions on Vehicular Technology*, 2023, Vol. 72, Issue 2, pp. 2280-2292, doi: 10.1109/TVT.2022.3207916

[4] M. Manalastas, *et al.*, "A Data-Driven Framework for Inter-Frequency Handover Failure Prediction and Mitigation," in *IEEE Transactions on Vehicular Technology*, 2022, Vol. 71, Issue. 6, pp. 6158-6172, doi: 10.1109/TVT.2022.3157802.

[5] Z. Liu, *et al.*, "Context-Aware and Adaptive QoS Prediction for Mobile Edge Computing Services," in *IEEE Transactions on Services Computing*, 2022, Vol. 15, Issue 1, pp. 400-413, doi: 10.1109/TSC.2019.2944596.

[6] J. Li, *et al.*, "Deep Reinforcement Learning-Based Mobility-Aware Robust Proactive Resource Allocation in Heterogeneous Networks," in *IEEE Transactions on Cognitive Communications and Networking*, 2020, Vol. 6, no. 1, pp. 408-421, doi: 10.1109/TCCN.2019.2954396.

[7] W. Liu and Y. Shoji, "Edge-Assisted Vehicle Mobility Prediction to Support V2X Communications," in *IEEE Transactions on Vehicular Technology*, 2019, Vol. 68, Issue 10, pp. 10227-10238, doi: 10.1109/TVT.2019.2937825.

[8] L. Huang, L. Lu and W. Hua, "A Survey on Next-Cell Prediction in Cellular Networks: Schemes and Applications," in *IEEE Access*, 2020, Vol. 8, pp. 201468-201485, 2020, doi: 10.1109/ACCESS.2020.3036070.

[9] C.X. Wang *et al.*, "On the Road to 6G: Visions, Requirements, Key Technologies, and Testbeds," in *IEEE Communications Surveys and Tutorials*, 2023, Vol. 25, Issue 2, pp. 905-974, doi: 10.1109/COMST.2023.3249835.

[10] P. Jain, A. Gupta, N. Kumar and M. Guizani, "Dynamic and Efficient Spectrum Utilization for 6G With THz, mmWave, and RF Band," in *IEEE Transactions on Vehicular Technology*, 2023, Vol. 72, Issue 3, pp. 3264-3273, doi: 10.1109/TVT.2022.3215487.

[11] Y. Zhao *et al.*, "Prediction-Based Spectrum Management in Cognitive Radio Networks," in *IEEE Systems Journal*, 2018, Vol. 12, Issue 4, pp. 3303-3314, doi: 10.1109/JSYST.2017.2741448.

[12] Y. Sun, Y. Chang, M. Hu and T. Zeng, "A Universal Predictive Mobility Management Scheme for Urban Ultra-Dense Networks With Control/Data Plane Separation," in *IEEE Access*, 2017, Vol. 5, pp. 6015-6026, doi: 10.1109/ACCESS.2017.2694863.

[13] H. Farooq, A. Asghar and A. Imran, "Mobility Prediction-Based Autonomous Proactive Energy Saving (AURORA) Framework for Emerging Ultra-Dense Networks," in *IEEE Transactions on Green Communications and Networking*, 2018, Vol. 2, Issue 4, pp. 958-971, doi: 10.1109/TGCN.2018.2858011.

[14] P. Hyebin, and Y. Lim., "Energy-Effective Power Control Algorithm with Mobility Prediction for 5G Heterogeneous Cloud Radio Access Network," in *Sensors*, 2018, Vol. 18, Issue 9, doi:10.3390/s18092904.

[15] Y. Ye, M. Xiao and M. Skoglund, "Mobility-Aware Content Preference Learning in Decentralized Caching Networks," in *IEEE Transactions on Cognitive Communications and Networking*, 2020, Vol. 6, Issue 1, pp. 62-73, doi: 10.1109/TCCN.2019.2937519.

[16] P. Li, Y. Xing and W. Li, "Distributed AI-native Architecture for 6G Networks," in Proceedings of *International Conference on Information Processing and Network Provisioning (ICIPNP)*, 2022, pp. 57-62, doi: 10.1109/ICIPNP57450.2022.00019.

[17] A. Mohamed *et al.*, "Mobility prediction for handover management in cellular networks with control/data separation," in Proceedings of *IEEE International Conference on Communications (ICC)*, 2022, pp. 3939-3944, doi: 10.1109/ICC.2015.7248939.

[18] A. Nadembega, A. Ha, and T. Taleb, "A destination and mobility path prediction scheme for mobile networks," in *IEEE Transactions on Vehicular Technology*, 2015, Vol. 64, Issue 6, pp. 2577-2590, doi: 10.1109/TVT.2014.2345263.

[19] L. Menz *et al.*, "An improved method for mobility prediction using a Markov model and density estimation," in Proceedings of *IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1-6, doi: 10.1109/WCNC.2018.8377086.

[20] A. Turkmen *et al.*, "Indoor Mobility Prediction for mmWave Communications using Markov Chain," in Proceedings of *IEEE Wireless Communications and Networking Conference (WCNC)*, 2021, pp. 1-5, doi: 10.1109/WCNC49053.2021.9417348.

[21] K.L. Yap and Y.W. Chong, "Optimized access point selection with mobility prediction using hidden Markov Model for wireless network," in Proceedings of *International Conference on Ubiquitous and Future Networks (ICUFN)*, 2017, pp. 38-42, doi: 10.1109/ICUFN.2017.7993744.

[22] H. Gebrie, H. Farooq and A. Imran, "What Machine Learning Predictor Performs Best for Mobility Prediction in Cellular Networks?," in Proceedings of *IEEE International Conference on Communications Workshops (ICC Workshops)*, 2019, pp. 1-6, doi: 10.1109/ICCW.2019.8756972.

[23] N. P. Kuruvatti *et al.*, "Mobility Awareness in Cellular Networks to Support Service Continuity in Vehicular Users," in Proceedings of *International Conference on Information and Communications Technology (ICOIACT)*, 2020, pp. 431-435, doi: 10.1109/ICOIACT50329.2020.9332125.

[24] A. Abebe *et al.*, "Mobility Prediction in Wireless Networks Using Deep Learning Algorithm," in Proceedings of *International Conference on Advances of Science and Technology (ICAST)*, 2020, pp. 454–461, doi: 10.1007/978-3-030-43690-2_31.

[25] M. Ozturk *et al.*, "A novel deep learning driven low-cost mobility prediction approach for 5G cellular networks: The case of the Control/Data Separation Architecture (CDSA)," in *Neurocomputing*, 2019, Vol. 358, doi:10.1016/j.neucom.2019.01.031.

[26] F. De Rango and P. Fazio, "A Stochastic Approach for Resource Prediction Error and Bandwidth Wastage Evaluation in Advanced Dynamic Reservation Strategies," in *IEEE Transactions on Mobile Computing*, 2023, Vol. 22, Issue 9, pp. 4986-5000, 1 Sept. 2023, doi: 10.1109/TMC.2022.3176046.

[27] A. Behara, T. G. Venkatesh, "Performance Analyses of Uplink MU-OFDMA Hybrid Access MAC in IEEE 802.11ax WLANs," in *IEEE Systems Journal*, 2022, Vol. 16, Issue 4, pp. 5108-5119, doi: 10.1109/JSYST.2022.3211860.

[28] F. Kooshki *et al.*, "Efficient Radio Resource Management for Future 6G Mobile Networks: A Cell-Less Approach," in *IEEE Networking Letters*, 2023, Vol. 5, Issue 2, pp. 95-99, doi: 10.1109/LNET.2023.3263926.

[29] D. Serghiou *et al.*, "Terahertz Channel Propagation Phenomena, Measurement Techniques and Modeling for 6G Wireless Communication Applications: A Survey, Open Challenges and Future Research Directions," in *IEEE Communications Surveys and Tutorials*, 2022, vol. 24, Issue 4, pp. 1957-1996, doi: 10.1109/COMST.2022.3205505.

[30] A. Bressan, K.T. Nguyen, "Conservation Law Models for Traffic Flow on a Network of Roads," in Networks and Heterogeneous Media, 2014, vol. 10, Issue 2, doi: 10.3934/nhm.2015.10.255.

[31] X.M. Chen, L.L. Qixin Shi, "Stochastic evolutions of dynamic traffic flow: Modeling and applications," in Stochastic Evolutions of Dynamic Traffic Flow: Modeling and Applications, 2015, pp. 1-193, doi: 10.1007/978-3-662-44572-3.

[32] D.Helbing, "Traffic and related self-driven many-particle systems", 2001, in Reviews of Modern Physics, Vol. 73, Issue 4, doi: 10.1103/RevModPhys.73.1067.

[33] Open Street Map Contributors, https://www.openstreetmap.org, 2017.

[34] Pablo Alvarez Lopez, *et al.*, "Microscopic Traffic Simulation using SUMO", in Proceedings of IEEE International Conference on Intelligent Transportation Systems, 2018, pp. 2575-2582, doi: 10.1109/ITSC.2018.8569938.

[35] D.C. Montgomery and G.C. Runger, "Applied Statistics and Probability for Engineers", 7th Edition, Wiley, 2018, ISBN: 978-1-119-40036-3

[36] P. Fazio, M. Tropea, "Advanced Resources Reservation in Mobile Cellular Networks: Static vs. Dynamic Approaches under Vehicular Mobility Model", in *MDPI Telecom*, 2021, Vol. 2, Issue 4, pp. 302-327, doi: 10.3390/telecom2040020.

[37] I. F. Akyildiz and W. Wang, "The predictive user mobility profile framework for wireless multimedia networks," in *IEEE/ACM Transactions on Networking*, 2004, Vol. 12, Issue 6, pp. 1021-1035, doi: 10.1109/TNET.2004.838604.

BIOGRAPHY

Peppino Fazio (*Member, IEEE*) received the Ph.D. degree in Electronics and Communications Engineering, University of Calabria (UNICAL - Italy) in January 2008 and completed his habilitation as Associate Professor in April 2017. After 12 years of Assistant Professor in UNICAL, he is Associate Professor at the Ca' Foscari University of Venice (UNIVE), in the Department of Molecular Sciences and Nanosystems (DSMN), from January 2021. From May 2008 to September 2008, he was a Visiting Research Fellow with the UPV of Valencia (Spain), GRC Research Group. He collaborates also, from 2017, with the VSB – Technical University of Ostrava (Czech Republic), as senior research fellow. He is co-author of more than 140 papers (60 in International Journals) and one book, all indexed in Scopus and/or WoS. He has been a member of various start-up companies and recipient of national and international awards (e.g. the Intel Business Challenge Award in 2013). He co-advised more than 70 BS and MS students and three Ph.D. students. His research interests include mobile communication networks, IP QoS architectures, wireless and wired networks, mobility modelling for WLAN environments, mobility analysis for prediction purposes, routing, vehicular networking, MANET, VANET and Quantum Key Distribution Networks.