Cell Permanence Time and Mobility Analysis in Infrastructure Networks: Analytical/Statistical Approaches and their Applications

Peppino Fazio[1], Floriano De Rango[1], Mauro Tropea[1], Miroslav Voznak[2]
[1]D.I.M.E.S. Department, University of Calabria, [2] VSB Technical University of Ostrava
[1]87036, Rende, Italy, [2]17 Listopadu 15/2172,70833 Ostrava-Poruba
e-mail: [1]{pfazio, derango, mtropea}, [2]miroslav.voznak@vsb.cz

Abstract— **Given the recent computation technologies and dedicated systems able to analyze huge amounts of "real-world" data and to extract information from it, the application possibilities reached in wireless mobile systems have significantly increased. The availability of mobile information, directly extracted from mobility traces, can be used to enhance the quality of the services offered, especially in a vehicular environment, where mobility is one of the main challenges. We present an in-depth study of mobility information management in infrastructure networks, where the covering devices can have a complete local visibility of what happens in terms of covered nodes. We focus our attention on the ways the gathered information can be used to set the coverage radius, to enhance a call admission control algorithm and to predict future movements. The paper introduces also a detailed description of how the obtained data can be statistically processed, and the obtained results confirm the benefits of the proposed scheme, especially in terms of call dropping probability (a minimum gain of 0.35 is obtained), which is a dominant parameter in infrastructure networks.**

Keywords - Realistic mobility model, Data analysis, Mobility prediction, Cell Permanence Time, Trace analysis, Hand-over.

I. Introduction

In recent years, mobile computing has become one of the main ways of information dissemination, due to its numerous advantages: users can be located anywhere without needing cables and they can move around (on feet, by cars, by trains, etc.), while maintaining their active flows. Analyzing mobility traces generated by real measurements is becoming a notable advantage in the enhancement of services offered in wireless networks [1], [2], [3]. The access to real datasets is very important in modern systems, because it allows for the investigation of on movement components, snatching detailed information about user behaviors. On the basis of the analyzed data, it is also possible to derive some analytical/statistical models: for this aim, it is very important to analyze real measurements, in order to obtain realistic models. In this work, attention is focused on the ways that real trace-files can be analyzed to extract the required information about users' behaviors. In particular, as stated previously, we carried out an in-depth analysis of mobility traces of wireless nodes in infrastructure networks, gathering different data, both in a *qualitative* and a *quantitative* way. A better knowledge of how the mobile nodes act during their flow lifetimes is obtained. We show the importance of the Cell Permanence Time (CPT) (also called Cell Dwell Time - CDT) for a mobile communication system, as well as the Hand-over Directional Probability matrix (HDP). We show how to build these structures and describe how they can be used in mobile systems to enhance the overall performance in terms of Quality of Service (QoS). Different additional features can be added to infrastructure wireless systems: mobility prediction (for in-advance reserving of future resources), bandwidth multiplexing (for increasing the number of admitted users, which can benefit from an enhanced transmission quality), enhanced coverage range and Signal-to-Noise Ratio (SNR, by detecting the most crowded places during the different moments of the day), discovering time-of-day periodicity and strong location preference, etc. The considered scenario, in this paper, regards a cellular infrastructure network, as illustrated in fig. 1. As can be seen, a particular geographical region (a part of London city in fig. 1) can be fully covered by the infrastructure network, which consists of a certain number of coverage cells (Access Points, Base Stations, LTE microcells, etc.), connected to a backbone, from which the service can be distributed. Each cell has a limited amount of resources (typically radio channels) which are used to satisfy mobile host requests.

Fig.1. The considered scenario, consisting of a mobility geographical region, covered by a certain number of service cells.

In brief, the amount of "mobility regularity" extracted from real traces provides the opportunity to observe high predictability of the patterns [4]. Nowadays, the interest for these features is migrating also to new research fields, such as opportunistic networks [5]: many research activities have focused on inter-contact and contact durations, obtaining different trends for their probability density functions (power-law, exponential, etc.). In addition, in opportunistic ad-hoc networking it is important to understand the opportunities for user devices to interact when users pass close to each other. In general, mobility brings an excellent level of comfort to Mobile Hosts (MHs) but, at the same time, it could introduce serious service degradations if coverage cells are not planned adequately. One of the main issues for Mobile Cellular Networks (MCNs) is represented by the consequence of hand-overs: once an MH leaves a coverage area, bandwidth resources are not available "for sure" in the new cell, because it is dependent on the congestion level of the new location. When a user pays for the perceived service, it is not admissible to have call dropping events, as well as a low level of QoS, in terms of throughput, end-to-end delay and jitter. A detailed analysis of mobility traces can bring sufficient knowledge for predicting future movements and user habits to the system, even if with a certain error probability. The main aim of this work is to show a possible way of considering the content of mobility traces in order to extract some knowledge about user habits. In addition, through the deployment of such information, different additional features can be added to infrastructure wireless systems, such as:

- mobility prediction (for in-advance reserving of future resources);
- bandwidth multiplexing (for increasing the number of admitted users, which can benefit from an enhanced transmission quality);
- enhanced coverage range and Signal-to-Noise Ratio (SNR, by detecting the most crowded places during the different moments of the day);
- discovering time-of-day periodicity and strong location preference.

The content of the paper helps the reader to know how data can be gathered from available resources (log-files mainly), with the possibility to obtain qualitative (related to the CPT and the number of future visited cells) and quantitative (related to the HDP and the direction behavior of mobile hosts) information. Differently from our previous works, here we would like to emphasize the way the previously obtained results can be applied in a real environment, not in a simulated nor theoretical one. In fact, we show how to obtain a log-file, analyze it and derive some interesting information. The other realistic approach regards the real coverage of the west part of Calabria (south Italy). The coverage map has been imported in our own simulator and, via C4R (a realistic mobility generator), real mobility traces have been analyzed (real existing roads have been considered) through the OpenStreetMap core engine. This work is not aimed at proposing a new prediction algorithm or reservation protocol, but rather seeks to demonstrate how a log-file (about user mobility) can be derived, analyzed and used to make predictions (with some examples of basic predictors, such as breadth-first, etc.). This paper is structured as follows: section II provides an overview of existing works, section III describes the structure of the trace-files and introduces different ways to analyze them, in order to enhance the quality of the services offered by mobile wireless networks. Section IV provides a summary of the main reachable results, while section V concludes the paper.

II. STATE OF THE ART

Recently, analyzing mobility traces in wireless networks has attracted the attention of the scientific community, also because people's behaviors are changing, in relation to mobile technology progress. Many papers exist, addressing the human behavior issue in different wireless network environments, such as GSM, WLAN, VANET, etc. The authors of [6] have based their study on collecting and analyzing GSM call data, stored by telecommunication operators. They try to create a user profiling by aggregating information on which data mining algorithms are performed. The process can automatically label the different call data into user behaviors. They have evaluated the considered users, classifying them into residents, commuters and visitors. Another work based on data collection is [7], where users' mobility information is gathered for analysis. The authors consider how a user behaves compared to its "friend" nodes. The interesting thing noticed by authors is the dependence of the users' location around the places of their friends and, in particular, where friends are denser. The considered location prediction model makes full use of this social information. The authors of the work proposed in [8] developed a new method to analyze the spatio-temporal activities of humans, extracted from smart phones information. Specifically, GSM and Wi-Fi network observations, collected by several users, are gathered to collaboratively build a symbolic base map of the logical structure of the geography. Their idea, called Proximity Map, is used to provide some spatial context to the individual mobility maps. This information is intended to be used for the analysis of transportation efficiency. The authors of [9] analyzed people's collective behavior, using a Big Data approach. They have investigated data coming from mobile operators that could be a support in producing reliable and timely estimates of intra-city mobility flows. The proposal consists of defining an estimation method based on calling data and of characterizing the mobility behavior at the level of a single municipality. In [10] the authors, on the basis of some literature conclusions that state that predictions based only on the movement history of a user are limited to a natural threshold, propose to gather and analyze information from other sources. In particular, they propose using information from calls between users in order to analyze the social aspects, in order to obtain greater reliability of the approach. They conclude that their location prediction can become more accurate by introducing social analysis methods. Another study of mobility prediction based on Markov chains is presented in [11]. The authors developed a model that predicts the user's trajectory in terms of handover sequences, in order to reduce the interruption time and signaling cost in the handover process. Another paper on the characterization of mobile user's traffic pattern for a better QoS user's experience is presented in [12]. In order to reserve the needed resources at the base stations, a method based on machine learning is proposed, in order to be able to forecast mobile user traffic pattern and guarantee better user QoS. They conclude that the proposal can improve system capacity and reduce system blocks. In [13] the authors propose a pattern prediction scheme associated with a bandwidth management in order to satisfy user requirements in term of QoS. The work is based on the proposal of two integrated schemes based on Markovian and statistical theories respectively. Their proposal has the task of minimizing bandwidth waste and the approach is completely independent of the considered technology, mobility model and vehicular environment. Another wireless environment, studied and analyzed by researchers, is the mobile ad-hoc network and, in particular, vehicular networks. The work presented in [3] presents an approach to extract mobile patterns from vehicular trace data in an urban environment, in order to perform path predictions. The authors, based on the gathered data, have studied and proposed a routing algorithm. They have simulated real vehicle traces, showing system performance, in terms of delivery ratio, also making a comparison with existing algorithms. In the work proposed in [2], the multicast issues in vehicular environment are taken into account. The authors' idea is represented by a new method, called TMC, which can use vehicle trajectories for an efficient multicast in vehicular networks. The novelty of the approach includes a message forwarding metric.

TABLE I. MAIN ABBREVIATIONS AND NOTATIONS.

AF	*Admitted Flow*
BU	*Bandwidth Utilization*
CDT	*Cell Dwell Time*
CDP	*Call Dropping Probability*
CEHT	*Cell Holding Time*
CHT	*Call Holding Time*
CID	*Cell ID*
CPT	*Cell Permanence Time*
CST	*Cell Stay Time*
HDP	*Hand-over Directional Probabilities*
HOD	*Hand-over Directions*

Table I shows the main terms used in the paper, while the next section introduces the proposed scheme.

III. MOBILITY ANALYSIS

In this section, the knowledge of user mobility, which could be derived from real traces, is described. Clearly, there are many classes of analyses that can be made and conducted on the trace-files, but in this paper, our attention focuses on the information which could improve infrastructure system performance, in terms of number of Admitted Flows (AFs), Call Dropping Probability (CDP) and/or Bandwidth Utilization (BU). Fig. 2 illustrates the typical fields of trace-files which could be obtained by the G-MoN application [14], available for Android Systems (client-side).

LCID	CID	LAC	NET	PSC	RNC	RXL	QUAL
RSRP	RSRQ	ECIO	TYPE	LAT	LON	SPEED	ALT
SNR	DATE	TIME	N1CID	N1LAC	N1RXL	N2CID	N2LAC
N2RXL	N3CID	N3LAC	N3RXL	N4CID	N4LAC	N4RXL	N5CID
N5LAC	N5RXL	N6CID	N6LAC				

Fig.2. Typical structure of trace files, obtained from the G-MoN application (client-side).

We are not interested in the number of bits reserved for each field, since we want to focus our attention on the concepts behind the collected data. In particular, during active sessions, G-MoN is able to store the dynamics of a mobile phone that is covered by a GSM, GPRS, UMTS or HSPA network into a trace-file (in .txt and .kml formats). As illustrated in fig. 2, a table with different fields is filled-up during time: the Cell ID (CID) is used to identify each Base Transceiver Station (BTS), while the UTRAN Cell ID (LCID) is used for UMTS connections, constituted by the concatenation of the Radio Network Controller (RNC) and CID fields. The Local Area Code (LAC) represents an identification for a location, the NETwork (NET) field is used to identify the network, PSC represents the used Primary Scrambling Codes (PHY layer), while the RX-Level (RXL) and QUALity (QUAL) fields are related to the received signal quality.

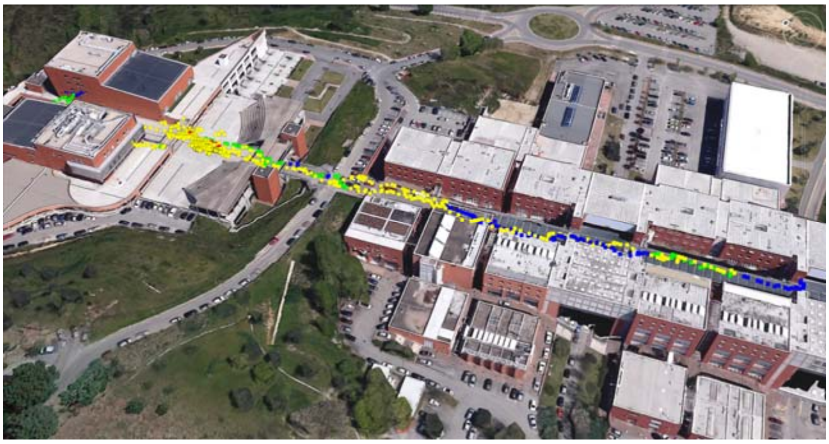


Fig. 3. A snapshot of the Google-Earth application, with the mobility traces obtained by G-MoN (colored points).

The Reference Signal Received Power (RSRP) and Reference Signal Received Quality (RSRQ) represent the power of the LTE Reference Signals spread over the full bandwidth and its quality, taking into account also the received signal strength and the number of used resource blocks (the RSRQ provides additional information when RSRP is not sufficient to make a reliable handover or cell reselection decision). The Ec/Io (ECIO) is the ratio of the received energy per chip and the interference level, usually given in dB, TYPE regards the kind of connection (EDGE, UMTS, etc.), LAT and LON are the LATitude and LONgitude of the current point, SPEED represents the current moving speed of the vehicle, while ALT is the ALTitude of the moving node.

CALL-ID	C-DATE	C-HOUR
C-DURATION	CALLER-ID	CALLED-ID
STARTING-BTS	ENDING-BTS	START-CID
START-CID-ADDR	START-CID-PLACE	ENDING-CID
ENDING-CID-ADDR	ENDING-CID-PLACE	CALLER-IMSI
CALLED-IMSI	IMEI-CTE/S.N.	CURRENT-CID
BEARER-CODE	TRAFFIC-TYPE	CARRIER-SELECTION

Fig. 4. Typical structure of trace files, obtained from supplier's data-base (server-side).

The SNR represents the signal to noise ratio, while the DATE and TIME fields are related to the time instant at which the measurements are made. The NiCID, NiLAC, NiRXL are the values of CID, LAC and RXL referred to the i-th coverage cell. Fig. 3 shows a mobility pattern generated by G-MoN for pedestrian mobility on our University Campus (UNICAL, north-side). Different colors represent the GSM coverage of the available coverage cells. Regarding the trace-files which can be obtained by service providers (server-side), the typical structure is illustrated in fig. 4.

TABLE II. THE MAIN VARIABLES USED IN THE PROPOSAL AND THEIR DESCRIPTION.

f_{CPT}, F_{CPT}	*Pdf/cdf of the Cell Permanence Time (CPT)*
α, β	*Parameters of a generic Gamma distribution*
α_{CPT}, β_{CPT}	*Parameters of the CPT Gamma distribution*
$\Gamma(\alpha)$	*Gamma function (eq. 2)*
$\gamma(\alpha, \beta t)$	*Support function for the evaluation of* F_{CPT} *(eq. 3)*
R	*Coverage radius of infrastructure cells (in meters)*
v_{max}	*Maximum speed of mobile hosts*
$E[CPT]$, $V[CPT]$	*Expected Value and Standard Deviation of the CPT variable*
A, B	*Matrices of the* a_{ij}, b_{ij} *coefficients of the regression analysis for* α_{CPT}, β_{CPT}
R_a, R_b	*Vectors of powers of R (from 0 to 3 and from 0 to 4 respectively)*
V	*Vector of powers of V (from 0 to 3)*
f_{CHT}	*Pdf of the CHT variable*
μ_{CHT}, σ_{CHT}	*Mean and standard deviation for the CHT distribution*
C	*Set of coverage cells (‖C‖=n)*
$MEANS_C$, $SIGMAS_C$	*Vectors of CPT means and standard deviations for the n coverage cells*
CPT_i	*Cell Permanence Time realization generated for the i-th cells, by the Marsaglia-Tsang approach [20]*
CHT_j	*Call Holding Time realization for the j-th user*
$N_{CELLS}(j)$	*Number of probably visited cells by the j-th user*
HOD	*Hand-over directions set for the possible hand-over directions (‖HOD‖=m)*
$p_i(j,k)$	*Probability of handing-out to direction k after CPTi amount of time, given the hand-in event from direction j*
HDP_i	*Hand-over probability matrix for the i-th cell*

There are different obvious fields, such as CALL-IDentification (CALL-ID), Call-DATE (C-DATE), Call-HOUR (C-HOUR), Call-DURATION (C-DURATION). Some other fields regard the information regarding the calling user and the called one: CALLER/CALLED-IDentification (CALLER/CALLED-ID), CALLER/CALLED-International Mobile Subscriber Identity (CALLER/CALLED-IMSI). The movement history of the moving nodes is also traced through the STARTING/ENDING-Base Transceiver Station (STARTING/ENDING-BTS), STARTing/ENDING-Cell IDentificator (START/ENDING-CID), STARTING/ENDING-Cell IDentificator-PLACE/ADDRess (STARTING/ENDING-CID-PLACE/ADDR) and CURRENT-CID. The other information regards internal managements, such as: International Mobile Equipment Identity-Cellular Telephone Entry System (IMEI-CTE), a code related to the carrier services (BEARER-CODE) and other information, such as TRAFFIC-TYPE (Mobile Originated - MO, or Mobile Terminated - MT, SMS, etc.) and CARRIER-SELECTION (when the connection should be switched from a carrier to another one). From the discussion above, it can be seen how a variety of interesting fields can be considered for our purposes. In the next subsections, different classes of analyses are introduced, each one aimed at optimizing some particular network profiles. Before starting the statistical/analytical description, a table with the main variables and parameters used in the proposal is shown (table II).

A) CPT models and statistical analysis

In this subsection, a statistical study of the CPT (also called Cell Stay Time - CST or CEll Holding Time CEHT) in infrastructure environments is described. The CPT is a very important value in wireless networks, because through a deep analysis of a source trace-file, different significant statistics can be derived in order to estimate how long a mobile device will stay in a particular cell and/or how many cells it will probably visit during its flow duration. These features are very useful in resource reservations for the environments that support node mobility. Independently of the type of available trace, this kind of analysis can be made by considering the available time-stamp values, such as TIME and CID/LCID fields for the client-side or the C-HOUR, START/ENDING-CID for the server-side. Nevertheless, the idea is to collect all the samples related to the permanence of a mobile host in a particular coverage cell. Therefore, given the trace file, it is easy to derive all the values of the time a mobile node, during a mobile call, spent under the coverage of the same cell (with a given CID). There are many big-data analysis approaches which could be used [15], [16], in order to retrieve hidden mobility behaviors, such as CPT distribution. With the availability of a thousand of CPT samples (from client or server side trace-files), the distribution can be well observed and, after a results analysis, a CPT pdf can be obtained. Many studies in literature [17] have demonstrated that, considering CPT as a random variable, its pdf can be well approached with a Gamma distribution [17]. Fig. 5 shows the CPT values for 800 samples (of 1200), obtained by the G-MoN application installed on two ANDROID Zenfone2 smart-phones. The logs were collected in three days, on board our Campus bus (dedicated to the transportation of students from one side of the campus to the other), following the itinerary illustrated in fig. 3, of a length of approximately 2km. They have been statistically analyzed through the MATLAB application. It is possible to see that, for the given coverage, the average CPT value is 29.975s.

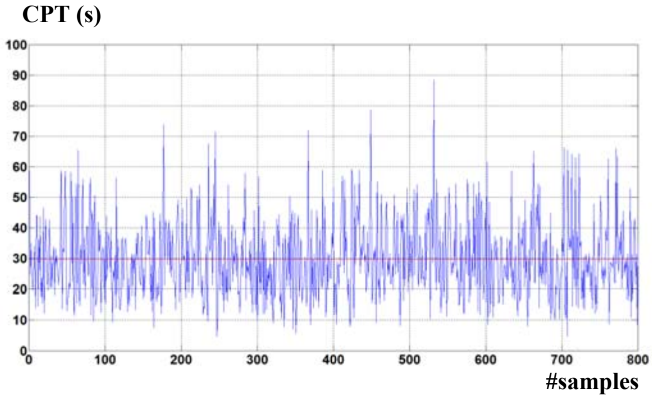


Fig. 5. CPT samples (s) extracted from a client-side trace-file, captured with G-MoN (the red line represents the average value).

Fig. 6 shows the fitting of the Gamma distribution for the obtained 800 samples (mean 29.4187 and variance 181.515), with parameters α=4.76798 (std. error 0.18816) and β=6.18005 (std. error 0.256778). Differently from [17], we consider instantaneous values, not the mean values in a certain amount of time.

So, the general expression of the CPT pdf can be considered to be:

$$f_{CPT}(t) = \frac{\beta^{\alpha} x^{\alpha-1} e^{-x\beta}}{\Gamma(\alpha)}, \tag{1}$$

where

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \tag{2}$$

is defined as the Gamma function. From eq. 1, the cdf of the instantaneous CPT is therefore:

$$F_{CPT}(t) = \int_{-\infty}^{t} f_{CPT}(z)dz = \frac{\gamma(\alpha,\beta t)}{\Gamma(\alpha)}, \quad and \quad \gamma(\alpha,\beta t) = \int_0^{\beta t} z^{\alpha-1} e^{-z} dz\,. \tag{3}$$

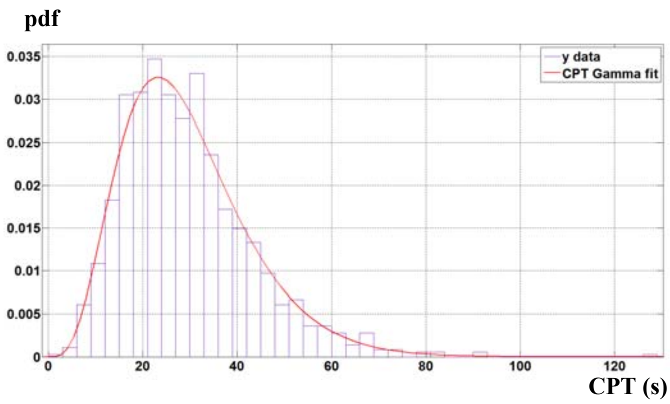


Fig. 6. CPT histogram (y data) and its Gamma approximation.

As in previous studies [17], where a normality test has been conducted, the assumption of a Gamma pdf can be easily verified through the Kolmogorov-Smirnov (KS) test [18]. MATLAB gives the opportunity to simply analyze the considered data, using the *kstest* function. In this way, supposing that we have a sample population $X_1, \ldots, X_n$ with an unknown distribution D, with the KS test we would like to verify the hypothesis that D is equal to a particular distribution D_0. In order to make an analytical analysis of the CPT trend as a function of some system parameters and to give more effectiveness to the practical approach mentioned before, a simulation approach was considered during our research activity. As stated in some of our previous works [17], the CPT statistic has a primary impact on the performance of a predictive system. In order to characterize its distribution parameters, we employed a realistic mobility model through C4R [19], able to reproduce the driving behavior of mobile hosts in a microscopic scale (more details about C4R will be given in the performance evaluation section); then, with our own simulator, a set of coverage clusters have been added to the considered map. In this way, we collected many statistics regarding user mobility, while varying the coverage radius extension of the single cell and mobility parameters (such as the acceleration, or average/max speed). We were, thus, able to evaluate the relation of the CPT distribution parameters with the coverage radius of the considered cells and the maximum speed of mobile hosts. Table III and table IV represent the obtained values of the CPT expected value and standard deviation for different values of coverage radius R [m] and mobile hosts max speed v_{max} [m/s].

TABLE III. VALUES OF E[CPT] (EXPECTED VALUE) FOR DIFFERENT R AND V_{MAX} VALUES .

v_{max} \ R	150	175	200	225	250	275	300
9.72	31.490982	36.563258	42.0307954	46.9125752	51.9393468	57.59939622	61.61854
11.11	30.590142	34.955985	39.6494136	44.42279253	49.01377209	53.24175677	57.71187
12.50	28.886033	33.634773	37.4530184	41.81122149	46.49551751	50.29318296	54.63064
13.80	27.367960	31.601184	35.3712918	39.28303761	43.32502283	47.05555509	51.53074

TABLE IV. VALUES OF √V[CPT] (STANDARD DEVIATION) FOR DIFFERENT R AND V_{MAX} VALUES.

v_{max} \ R	150	175	200	225	250	275	300
9.72	0.47242	0.51223	0.703947764	0.991345135	1.144869985	1.578785376	1.791512
11.11	0.59663	0.629521062	0.879106296	1.216987405	1.378359107	1.776885137	1.842804
12.50	0.65454	0.75654	0.983907056	1.260561831	1.534610323	1.856321	2.071542
13.80	0.735678	0.79552	1.069073383	1.240856118	1.756758088	1.905441902	2.2645175

At this point, it is easy to derive the statistical parameters, α and β, of the gamma distribution, recalling that [18]:

$$\alpha_{CPT} = \frac{E^2[CPT]}{\sqrt{V[CPT]}} \quad and \quad \beta_{CPT} = \frac{V[CPT]}{E[CPT]} \ . \tag{4}$$

So, given the values above in the tables, the following trends for the statistical parameters are obtained:

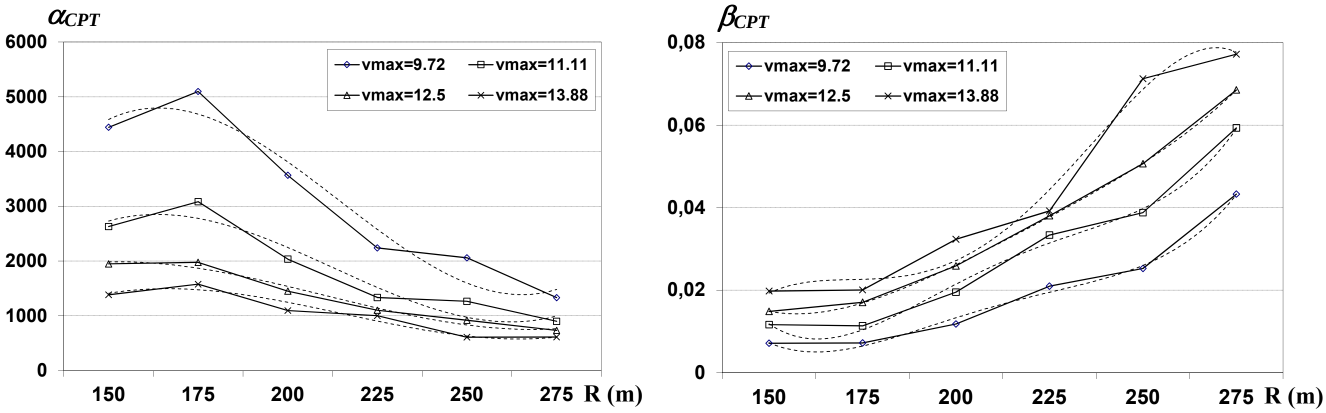


Fig. 7. Trend of α (left) and β (right) for different values of coverage radius R [m] and maximum mobility speed v_{max} [m/s].

In fig. 7, the solid curves represent the obtained values, while the dotted ones are their polynomial regression [20]: we used the MATLAB application to analyze the collected data and to evaluate the *determination coefficient* for each approximation. We noticed that it maintains above the value 0.9823, which is sufficient to confirm that different polynomial functions can be considered for describing the trends of α and β. In particular, a 3rd order polynomial is considered for the trend of α, while a 4th order polynomial is sufficient to describe the trend of β:

$$\alpha_{CPT} = a_3 R^3 + a_2 R^2 + a_1 R + a_0 \quad and \quad \beta_{CPT} = b_4 R^4 + b_3 R^3 + b_2 R^2 + b_1 R + b_0 \quad . \tag{5}$$

Relating the expressions in eq. (5) with the other considered parameter (v_{max}), it is possible to notice that the a_i and b_j coefficients, i=0,..,3 and j=0,..,4, are functions of v_{max}. Also in this case, a 3rd order polynomial regression can describe the trend of those coefficients in function of v_{max}, so eq. (5) can be re-written as:

$\alpha_{CPT} = (a_{33}v_{\max}^3 + a_{32}v_{\max}^2 + a_{31}v_{\max} + a_{30}) \cdot R^3 +$ $+(a_{23}v_{\max}^3 + a_{22}v_{\max}^2 + a_{21}v_{\max} + a_{20}) \cdot R^2 + (a_{13}v_{\max}^3 + a_{12}v_{\max}^2 + a_{11}v_{\max} + a_{10}) \cdot R +$ $+(a_{03}v_{\max}^3 + a_{02}v_{\max}^2 + a_{01}v_{\max} + a_{00})$	$\beta_{CPT} = (b_{43}v_{\max}^3 + b_{42}v_{\max}^2 + b_{41}v_{\max} + b_{40}) \cdot R^4 + (b_{33}v_{\max}^3 + b_{32}v_{\max}^2 + b_{31}v_{\max} + b_{30}) \cdot R^3 +$ $+(b_{23}v_{\max}^3 + b_{22}v_{\max}^2 + b_{21}v_{\max} + b_{20}) \cdot R^2 + (b_{13}v_{\max}^3 + b_{12}v_{\max}^2 + b_{11}v_{\max} + b_{10}) \cdot R +$ $+(b_{03}v_{\max}^3 + b_{02}v_{\max}^2 + b_{01}v_{\max} + b_{00})$

(6)

which, in compact form, corresponds to the expressions $\alpha_{CPT}= \mathbf{R_\alpha A V}$ and $\beta_{CPT}= \mathbf{R_\beta B V}$, if the vectors $\mathbf{R_\alpha}=[1\ R\ R^2\ R^3]$, $\mathbf{R_\beta}=[1\ R\ R^2\ R^3\ R^4]$ and $\mathbf{V}=[1\ v_{max}\ v^2_{max}\ v^3_{max}]$ are defined and $\boldsymbol{A}$, $\boldsymbol{B}$ are respectively:

$$A = \begin{bmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{bmatrix} \quad and \quad B = \begin{bmatrix} b_{00} & b_{01} & b_{02} & b_{03} \\ b_{10} & b_{11} & b_{12} & b_{13} \\ b_{20} & b_{21} & b_{22} & b_{23} \\ b_{30} & b_{31} & b_{32} & b_{33} \\ b_{40} & b_{41} & b_{42} & b_{43} \end{bmatrix} \tag{7}$$

The following figure summarizes the regression coefficients for the considered curves.

$$A=\begin{bmatrix} 7.297 & -43.927 & 42.596 & 95.884 \\ -78.403 & 464.62 & -421.39 & -1061 \\ 260.34 & -1600.7 & 1926.6 & 2086.8 \\ -344.68 & 2666.4 & -6773.7 & 7357.5 \end{bmatrix} \qquad B=\begin{bmatrix} 3e-5 & -0.0005 & 0.0022 & -0.0024 \\ 0.0003 & 0.0013 & -0.0174 & 0.0247 \\ -0.0005 & -0.0097 & 0.0824 & -0.1077 \\ 0.0058 & -0.0162 & -0.076 & 0.1441 \\ -0.0094 & 0.048 & -0.0381 & -0.0117 \end{bmatrix}$$

Fig. 8. Regression coefficients for matrixes ***A*** and ***B*** of eq. (7).

It should be noticed that, alternatively to the considered system parameters, such as R and v_{max}, the same analysis can be made by considering different variables (e.g. average speed, transmission power, road topology, acceleration/deceleration, etc.): the main aim of this subsection is to show how the CPT distribution parameters can be analytically described.

B) The importance of CPT statistics for cellular networks

The approach described in the previous sub-section is based on the extraction of a pdf (Gamma or Gaussian) from a trace-file, in order to describe the main parameters (such as mean and standard variation, for example) of the CPT for a given cell. As shown, these parameters are closely affected by the coverage radius and the geographical morphology since, for a given coverage range, the average speed of the vehicles influences the obtained results. In this subsection, the importance of knowledge of the CPT is highlighted. First, it is mandatory to introduce the concept of Call Holding Time (CHT), considered as the time duration of an active session (for voice or data). From many recent works in literature, dedicated to the study of its parameters, we can conclude that the classical hypothesis of exponential distribution is no longer valid [21]. User behaviors have changed in time and all the studies on real traces have demonstrated that the obtained recent distributions are different from the exponential one. Nowadays, people are becoming technologically experienced and the impact of new communication paradigms heavily affects human behavior, which reflects on CHT statistics. In order to validate the log-normal assumption [21] for the CHT pdf, we carried out a preliminary set of calls: 20 students, each one equipped with an ANDROID Lollipop 5.0 smart-phone, logged their inbound/outbound conversations through G-MoN for one week. After a proper trace-files filtering activity, a number of 1600 samples of CHT were obtained.

Fig. 9 shows the trend, expressed in seconds, of the obtained samples, with a mean value of 218.8799s (red line), while fig. 10 shows its cdf and the log-normal fitting. The considered expression for the CHT pdf is:

$$f_{CHT}(t)=\frac{1}{t\sqrt{2\pi}\sigma_{CHT}}e^{-\frac{(\ln t-\mu_{CHT})^2}{2\sigma^2{}_{CHT}}}, \tag{8}$$

and its parameter values for the trend of fig. 7 are μ_{CHT}=4.0138 and σ_{CHT}=1.69438.

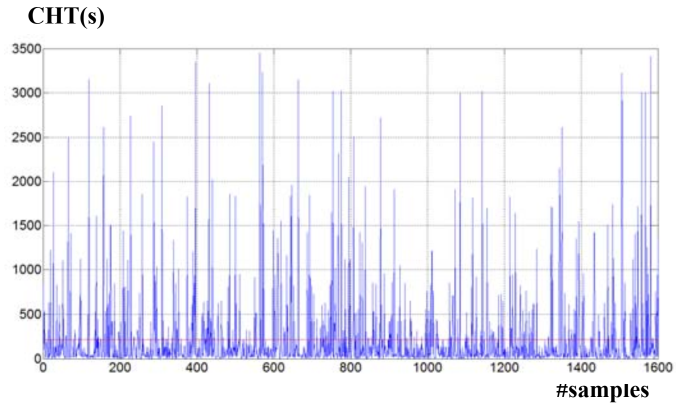


Fig. 9. CHT values obtained from the analysis of 20 students' trace files for one week (1600 samples).

Further details on CHT fitting can be found in [21], where a filtering method is proposed, in order to find different log-normal distributions in the same trace-file: in this sense, the desired pdf is said to be constituted by a *mixture* of *k* log-normal distributions.

At this point, an important consideration about CPT and CHT can be made. Given an infrastructure wireless network with mobile nodes, formed by a set $C=\{c_1,...,c_n\}$ of coverage cells, after an in-depth analysis of the possible trace-files (client or server side), it is possible to characterize each coverage cell by its typical CPT distribution. In fact, two arrays can be associated to C: $MEANS_C=\{\mu_1,..., \mu_n\}$ and $SIGMAS_C=\{\sigma_1,..., \sigma_n\}$, containing, respectively, the means and the standard deviations of the CPT distributions for the *i-th* cell, $i=1,..,n$.

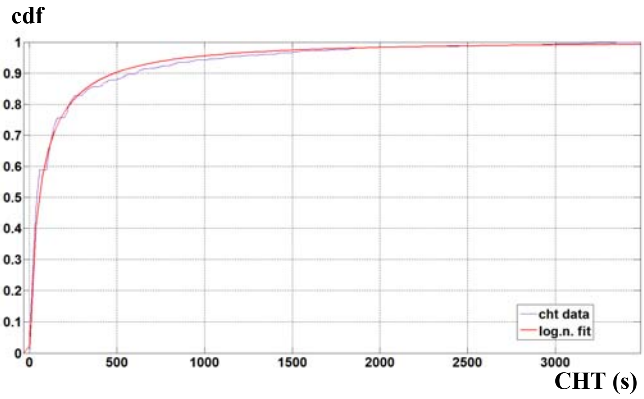


Fig. 10. CHT cdf fitting with a log-normal distribution.

Then, following the Marsaglia-Tsang approach [20], [21] able to create Gamma deviates, it is possible to obtain a realization of the CPT distribution for cell i as follows. First, two constant values c_1 and c_2 should be set to α-1/3 and $1/(9c_1)^{1/2}$ respectively. Then two numbers u_1 and u_2 are generated, where u_1 follows a normal distribution (null mean and a standard deviation equal to 1), while u_2 follows a uniform distribution, in the range (0,1]. If $u_1 > (-1/c_2)$ and $log(u_2) < 0.5u_1^2+c_1-c_1u_3+c_1ln(u_3)$, with $u_3=(1+c_2u_1)^3$ then:

$$CPT_i = c_1 \cdot u_3 \ . \tag{9}$$

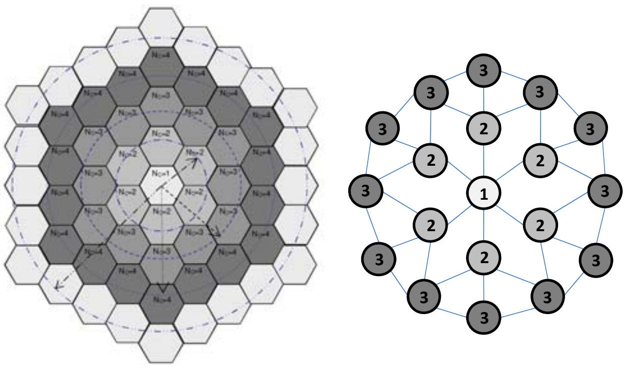
If the conditions are not satisfied, the values u_1 and u_2 are extracted again and the conditions are checked another time. This approach is valid for β=1 and α≥1. For the general approach, please refer to [21], [22]. The Box-Muller approach [24] is able to generate normal deviates, so a realization of the CHT for the *j-th* user can be obtained as:

$$CHT_j = \exp[stdev_{CHT} \cdot \sqrt{-2\ln u_1} \sin(2\pi u_2) + mean_{CHT}] \ , \tag{10}$$

where u_1 and u_2 are two random numbers, uniformly distributed in (0,1] and the exponential term is added to obtain a log-normal realization, starting from a Gaussian one. The terms $mean_{CHT}$ and $stdev_{CHT}$ represent the mean and standard deviation of the log-normal distribution, evaluated from its parameters μ_{CHT} and σ_{CHT} as follows:

$$mean_{CHT} = \log(\mu_{CHT}) - \frac{1}{2}\log\left[1+\left(\frac{\sigma_{CHT}}{\mu_{CHT}}\right)^2\right] \qquad stdev_{CHT} = \sqrt{\log\left[1+\left(\frac{\sigma_{CHT}}{\mu_{CHT}}\right)^2\right]} \tag{11}$$

The parameters μ_{CHT} and σ_{CHT} are not related to a particular mobile host, but they describe the CHT behavior for the whole system. When an infrastructure network (such as a GSM, UMTS, HSPA, WLAN with handover management, etc.) needs to guarantee QoS to mobile nodes, it has to reserve the adequate amount of bandwidth for mobile hosts, also in future locations, not only in the current ones. In this way, the CDP can be minimized. We do not focus attention on a particular in-advance resource reservation protocol or algorithm, however, the importance of the CPT and CHT statistics for this issue should be underlined.



N_{CELLS}	1	2	3	4	5	6	7	8	...
$N_{RESERV.}$	1	7	19	37	61	91	127	169	...

Fig. 11. A typical structure of a cluster of coverage cells, with regular shapes (on the left) and the associated graph (on the right).

Without a specific pattern prediction algorithm, the network can derive only *quantitative* knowledge about user behaviors: that is to say CPT and CHT. Fig. 11 helps us to understand the concept. In fig. 11, we suppose that a mobile host starts its call admission request in the center of the cluster (light colored cell). A fully connected graph can be easily associated to the map, in which each node represents a cell, while the edges are related to the adjacencies. Assuming that the coverage structure manages a mobility map (urban, sub-urban, rural, highway, etc.), and that the infrastructure network wants to minimize the CDP (no droppings during hand-overs, so a pre-reservation policy is mandatory) with the trace-files analyses described until now, it is possible to make only a *quantitative* prediction. The network can calculate a predicted amount of probably visited cells, but it cannot have a probabilistic description of the CIDs related to the cells which will be visited. In other words, we can know the number of future visited cells (*quantitative* prediction), but, at the moment, we cannot identify them (*qualitative* prediction). So, if the CPT_i is the same (in terms of μ_{CPT} and σ_{CPT}) for each coverage cell c_i, let us indicate it with CPT^*, $i=1,...,n$, the number of probably visited cells by the *j-th* user $N_{CELLS}(j)$ can be derived as:

$$N_{CELLS}(j) = \left\lceil \frac{CHT_j}{CPT^*} \right\rceil. \tag{12}$$

The assumptions of identical CPT_i and regular coverages are not suitable for real environments. Regarding the latter, the network can be modeled by a Voronoi tessellation [25], but we do not focus our attention on that issue.

```
//set initial variables
//maximum reached level
NCMAX=0;
//residual time
res_time=CHTj;
//add the starting cell and its attributes to the queue
r_cells.push(starting_cell, NCMAX, residual_time);
//while visiting all the cells
while (reached_cells.size()>0) {
    //dequeue the current item
    (curr_ci, curr_NC, res_time)=r_cells.pull();
    //check the current breadth
    if (curr_NC>NCMAX) NCMAX++;
    //evaluate if there is residual time for another hand-off
    if (res_time>CPTcurr_ci) {
            //visit all the possible next cells
            for_each (Adjacent(curr_ci) adj_ci) {
                    r_cells.push(adj_ci, NCMAX+1, res_time-CPTcurr_ci);
            }
    }
}
return NCMAX;
```

Fig. 12. The pseudo-code related to the breadth-first prediction of future visited cells.

Regarding the first assumption, the value of $N_{CELLS}(j)$ can be simply evaluated through the pseudo-code above (fig. 12), considering that the CPT_i changes as a function of the shape of c_i and the covered area morphology. If we consider the graph associated to the cluster of cells, the process of evaluating the maximum number of probable visited cells (N_{CMAX}) is resumed in fig. 12. The pseudo-code above simply evaluates the maximum breadth that can be reached in a time CHT_j, considering the different CPT_i and the adjacencies map. The functions *push()* and *pull()* enqueue and dequeue the elements to/from the residual_cells (r_cells) array. So, in the real case, $N_{CELLS}(j)=N_{CMAX}$. Given the predicted value of $N_{CELLS}(j)$, the effective number of cells on which the resources should be reserved (considering the current one) is:

$$N_{RESERVATIONS}(j)=3 \cdot N_{CELLS}(j) \cdot [N_{CELLS}(j)-1]+1 \tag{13}$$

because only the maximum breadth of the mobile host in the network is known, without any information about the CIDs (it increases in a polynomial way, as illustrated in fig. 11). The possibility of extracting *quantitative* information from trace-files gives the opportunity of enhancing the CDP by pre-reserving the adequate amount of bandwidth [17] at the cost of a significant number of unused channels.

C) HDP and its importance for statistical analyses in infrastructure networks

The resource wastage is an undesired effect, especially in wireless networks. The information which can be obtained by analyzing the trace-files from a *quantitative* point of view (CPT and CHT) needs to be integrated with a *qualitative* analysis, that is to say an approach able to predict (with a negligible amount of error) not only the number of visited cells, but the CIDs of the cells which will accommodate a mobile user during its active session. Also in this case, independently from the type of available trace, this kind of analysis can be made by considering the available fields, related to the movement history of mobile hosts: CIDs, and DATE/TIME fields for the client-side or the START-CID, CURRENT-CID, ENDING-CID and C-DATE/C-HOUR for the server side.

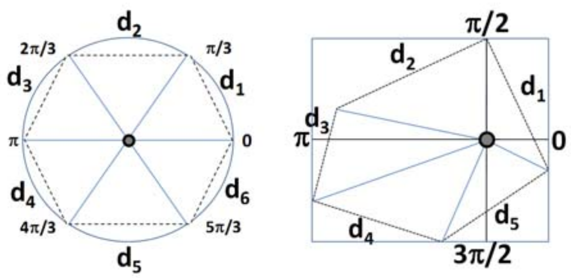


Fig. 13. Cellular shape approximation through a regular polygon (on the left) and a Voronoi cell (on the right).

For a *qualitative* approach, the idea is to collect all the sequences of visited cells, and derive the directional hand-over probabilities, on the basis of the adjacencies of the cluster map. Therefore, given the trace file, for each stored call, it is easy to derive the historical movements, in terms of cell sequences. Once the *qualitative* information has been collected, many different approaches can be used to predict the set of future visited cells: neural networks [26], Markov chains [27], [28], Kalman filters [29], or data mining [30] for example. Independently from the chosen predictive approach, the procedure is the same: the predictor needs to learn user behavior from trace-files (by the historical sequence of actually visited cells). Clearly, our attention does not focus on a particular unconventional approach, but only the general idea is now illustrated. A generic coverage area can be described by a regular polygon, or a Voronoi cell [25], for which a finite set of possible Hand-Over Directions (HOD) can be defined, fig. 13. If the polygon has m edges, then $||HOD||=m$. Without loss of generality, a value of m=6 can be considered and $HOD=\{d_1, d_2, d_3, d_4, d_5, d_6\}$.

```
//set initial variables
for_each(j,k) {
  HDPi(j,k)=0;
}
//extract information from client-side trace-files
if (trace-files are client-side) {
  for_each(trace-file tfm) {
    update_probability(tfm, client);
  }
}
//extract information from a server-side trace-file tfm
else {
  update_probability(tfm, server);
}
return HDPi;

void update_probability(trace-file tfm, type) {
  for_each(hand-over event in tfm) {
    if (type==client) use the CIDs and TIMEs fields;
    else use CURRENT-CIDs and C-DATEs/HOURs fields
    extract the ordered sequence of visited CIDs;
    //analysis of mobile host hand-over sequences
    //determine when an hand-in to ci occurred
    if (current cell is ci and previous cell is ch) {
      //evaluate hand-in direction
      din=evaluate_hand_in_direction(ch,ci);
      //evaluate hand-out direction
      cl=evaluate_next_visited_cell(tfm);
      dout=evaluate_hand_out_direction(ci,cl);
    }
    update HDPi(in,out);
  }
}
```

Fig. 14. The pseudo-code related to the HDP extraction from trace-files.

Each direction, univocally, identifies the next adjacent coverage cell. Given a cell $c_i \in C$, the conditional probability that a mobile host will be handed-out to direction d_k after CPT_i amount of time, if it was handed into c_i from direction d_j is defined as $p_i(j,k)$:

$$p_i(j,k) = p(hand_{out} to \;\; d_j, t = t_0 + CPT_i \,/\, hand_{in} \;\; from \;\; d_i, \;\; t = t_0) \quad , \tag{14}$$

where t_0 is the hand-off arrival time in c_i. All the values of eq. 14, for a given coverage cell c_i, are suitable to be stored in a square matrix, since the value of m is a-priori established; let us indicate it with HDP_i. Independently from the nature of the trace-file (client or server side), the elements of the HDP_i can be calculated as illustrated in fig. 14. It shows the pseudo code related to the analysis of a trace-file, from which the structure of HDP_i can be gathered, for each c_i, i=1,...,m. In particular, if there is the availability of client-sided traces, the system will analyze mobility for all mobile hosts, each one providing a trace-file for its own pattern. On the other side, if the trace is managed from the server-side, all mobile patterns are contained in a single file. In both cases, the function *update_probability* is called. For each triplet (c_h, $\boldsymbol{c_i}$, c_l), where c_i is the cell for which HDP_i is being updated, the hand-in and hand-out directions can be evaluated, so the related element $HDP_i(d_{in},d_{out})$ can be updated.

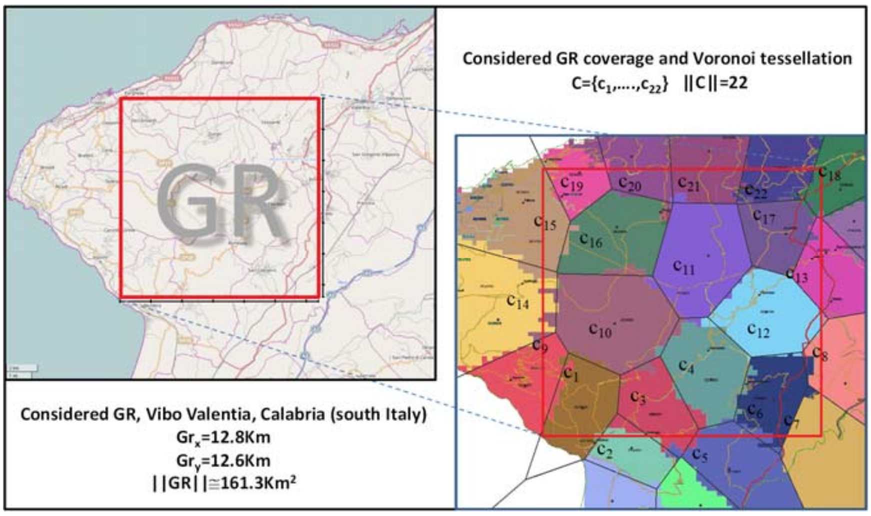


Fig. 15. The considered UMTS coverage map and the related Voronoi approximation.

Clearly, the pseudo-code contains only an example of the utilization of that function. When all the sequences of the trace-files have been considered, the elements of HDP_i will be formed as follows:

$$HDP_i(j,k) = \frac{Hand\_out_i(j,k)}{Hand\_out_i(j)} \quad , \tag{15}$$

where the term $Hand\_out_i(j,k)$ represents the number of total hand-over events from direction d_j to direction d_k and $Hand\_out_i(j)$ is the number of total handover events from direction d_i to any other one. Clearly:

$$\sum_{k=1}^{n} Hand\_out_i(j,k)=1 \qquad (16)$$

In order to obtain different numerical values for HDP_i, a Geographical Region (GR) of about 160km^2 has been considered, as illustrated in fig. 15.

CGI	t	CELL	Place	Latitude	Longitude	Irrad. Dir. (°)	Alt. (m)	Tilt (°)
222-01-22511-57310	1	ZC55D2	ROMBIOLO	38 33 54.08	15 56 46.37 E	320	328	0.00
222-01-22511-57005	2	CZ43D1	NICOTERA	38 33 26.15	15 56 19.95 E	98	236	4.00
222-01-22511-57309	3	ZC55D1	ROMBIOLO	38 33 54.08	15 56 46.37 E	60	328	0.00
222-01-21141-56718	4	CZ99D2	CALIMERA	38 33 54.99	16 01 04.03 E	339	276	0.30
222-01-22511-57165	5	ZC19D1	SCIORDELLA	38 31 24.07	16 00 37.38 E	20	75	0.00
222-01-21141-56717	6	CZ99D1	CALIMERA	38 33 54.94	16 01 04.19 E	58	276	2.60
222-01-21141-56630	7	CZ77D2	BIVIO MILETO	38 34 02.67	16 04 03.28 E	146	233	6.00
222-01-21141-56629	8	CZ77D1	BIVIO MILETO	38 34 02.74	16 04 03.41 E	45	233	5.00
222-01-22511-57053	9	CZ55D1	JOPPOLO	38 34 05.26	15 54 11.39 E	10	50	0.00
222-01-21141-57194	10	ZC26D2	ZUNGRI	38 38 08.36	15 57 36.15 E	200	610	1.60
222-01-21141-57193	11	ZC26D1	ZUNGRI	38 38 08.36	15 57 36.28 E	131	610	7.00
222-01-21141-57079	12	CZ61D3	MILETO	38 36 25.77	16 03 47.92 E	300	364	0.70
222-01-21141-57181	13	ZC23D1	IONADI	38 37 51.49	16 03 31.05 E	80	501	2.20
222-01-21141-57278	14	ZC47D2	SPILINGA	38 38 37.08	15 53 32.36 E	150	414	0.00
222-01-21141-64817	15	ZC76D1	BRATTIRO'	38 39 39.25	15 53 58.39 E	60	325	6.00
222-01-21141-57195	16	ZC26D3	ZUNGRI	38 38 08.35	15 57 36.02 E	279	610	4.00
222-01-21141-57183	17	ZC23D3	IONADI	38 37 51.57	16 03 30.96 E	347	501	1.00
222-01-21141-56859	18	CZ06D3	VIBO V.C.	38 40 24.72	16 06 33.90 E	238	561	6.00
222-01-22511-57009	19	CZ44D1	TROPEA BASSA	38 40 22.08	15 52 30.36 E	60	450	0.00
222-01-21141-56982	20	CZ37D2	ZAMBRONE	38 43 04.68	16 00 44.85 E	199	169	0.50
222-01-21141-57142	21	ZC13D2	BRIATICO	38 43 15.61	16 02 14.04 E	178	78	1.00
222-01-21141-64858	22	ZC86D2	VENA	38 39 45.20	16 03 35.74 E	326	484	9.10

Fig. 16. Main radio-mobile parameters of the considered cells.

The GR is formed by a total number of m=22 cells and their main radio-mobile parameters are illustrated in fig. 16. The map has been managed by a specific Voronoi tessellation [25], as depicted in fig.15 (right side). Fig. 17 shows the obtained values for HDP_3 (m=5), HDP_{10} (m=7) and HDP_{11} (m=6). The next section shows some of the obtainable results if the information gathered from trace-files is integrated into the system (with *qualitative*, *quantitative* or both approaches).

$$\mathbf{HDP_3}=\begin{array}{c|ccccc|} & 1 & 2 & 3 & 4 & 5 \\ 1 & 0.00125 & 0.2712 & 0.34556 & 0.2902 & 0.0909 \\ 2 & 0.53998 & 0.00443 & 0.20557 & 0.1566 & 0.09332 \\ 3 & 0.15879 & 0.3394 & 0.00556 & 0.31966 & 0.17677 \\ 4 & 0.3499 & 0.00899 & 0.10669 & 0.000743 & 0.53271 \\ 5 & 0 & 0.829 & 0.0035 & 0.15779 & 0 \end{array}$$

$$\mathbf{HDP_{10}}=\begin{array}{c|ccccccc|} & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 0 & 0 & 0.0675 & 0.12335 & 0.2268 & 0.0548 & 0.5218 \\ 2 & 0.55432 & 0.00023 & 0.35663 & 0.044 & 0.0033 & 0.0043 & 0.0375 \\ 3 & 0.0566 & 0.78332 & 0 & 0.07899 & 0.0007 & 0 & 0.081 \\ 4 & 0.00336 & 0.07559 & 0.25794 & 0 & 0.15336 & 0.2778 & 0.2289 \\ 5 & 0.78446 & 0.03364 & 0.00788 & 0.01856 & 0 & 0 & 0.1567 \\ 6 & 0.43566 & 0.00032 & 0.0079 & 0.055923 & 0.44561 & 0.0006 & 0.0511 \\ 7 & 0.00422 & 0 & 0.15699 & 0.18779 & 0.36997 & 0.2756 & 0.0061 \end{array}$$

$$\mathbf{HDP_{11}}=\begin{array}{c|cccccc|} & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 0.27107 & 0.14905 & 0.15559 & 0.119348 & 0.15452 & 0.1504 \\ 2 & 0.19993 & 0.23777 & 0.20825 & 0.124466 & 0.12232 & 0.1079 \\ 3 & 0.12915 & 0.20651 & 0.23885 & 0.198155 & 0.10908 & 0.1225 \\ 4 & 0.12158 & 0.15378 & 0.15179 & 0.26855 & 0.15054 & 0.1538 \\ 5 & 0.12452 & 0.1254 & 0.108 & 0.195957 & 0.24041 & 0.2057 \\ 6 & 0.1984 & 0.10828 & 0.12388 & 0.123982 & 0.20585 & 0.2396 \end{array}$$

Fig. 17. Hand-over Directional Matrices for cells c_3, c_{10}, c_{11}.

IV. MAIN REACHABLE RESULTS WITH QUANTITATIVE AND QUALITATIVE ANALYSES

In this section, some of the possible enhancements reachable with the proposed analyses are illustrated. The following figures are the results of different simulations, in which real coverage maps have been imported and real mobility patterns have been used, through the C4R [19] graphical interface. In particular, we considered the GR illustrated in fig. 13 and different simulation campaigns were conducted. In particular, the GR coverage map was approximated by a Voronoi tessellation, and a Java simulator, able to reproduce user mobility, reservations mechanisms and to apply *quantitative* (CPT-based), *qualitative* (HDP-based) or both prediction approaches, as described in section III was implemented. The main aim of the curves consists in showing the QoS enhancements which could be reached if the mobility traces are correctly analyzed. As performance parameters, we considered AFs, CDP and BU. The simulated system consists of $\|C\|$=22 wireless cells, each one covered by an Access Point (AP). We assumed a total capacity for each AP equal to 11Mb/s, and a total number of 11 channels of 1Mb/s. We assume that one user can reserve only one channel for each AP. The coverage area varies from one cell to another one, as well as the possible hand-over directions. A log-normal distribution for CHT has been assumed, with a mean value of 200s, while the inter-arrival process has been considered to be Poissonian, with a rate of 1 new request every 3 seconds. The Call Admission Control (CAC) and the Bandwidth Reallocation Algorithm (BRA) schemes are the same as those of [13], and the reservation protocol is the MRSVP [23]. The average speed was set to [5-75] km/h, and the simulation time to 2000s for each run. After a first simulation campaign, the values of CPT_i and the HDP_i matrices were collected for the entire network, i=1,...,22: once the Voronoi tessellations were set, real mobility traces were generated through C4R and the hand-over events were monitored and considered in the statistical analysis.

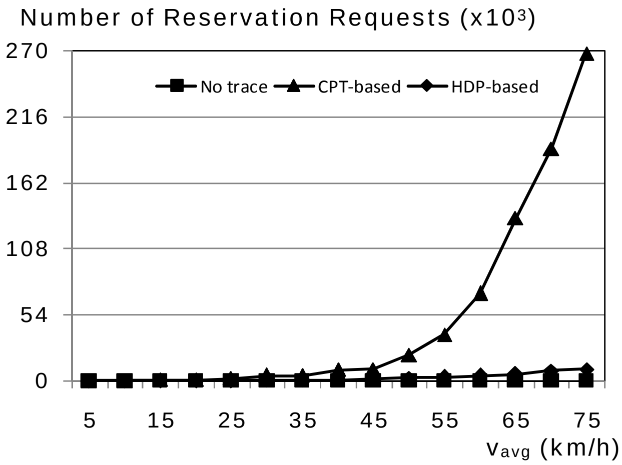


Fig. 18. Average Number of Reservation Requests (NRR) as a function of the average nodes speed.

The illustrated figures represent only an example of how the performance of an infrastructure network can be enhanced, by extracting information on user behavior. Fig. 18 shows the trend of the average number of reservation requests. As described in section III, for the *quantitative* trace-analysis (CPT-based), the number of requests increases in a polynomial way and it is not suitable for large networks, due to the high level of overhead introduced into the system. On the other side, the *qualitative* approach (HDP-based) maintains the request number within an acceptable level (it is not based on a circular reservation policy, and the requests are sent to a restricted and directional set of cells), while without the trace analysis (No-trace curve) the requests are sent only on the current cells (a constant trend is obtained, around two thousand requests).

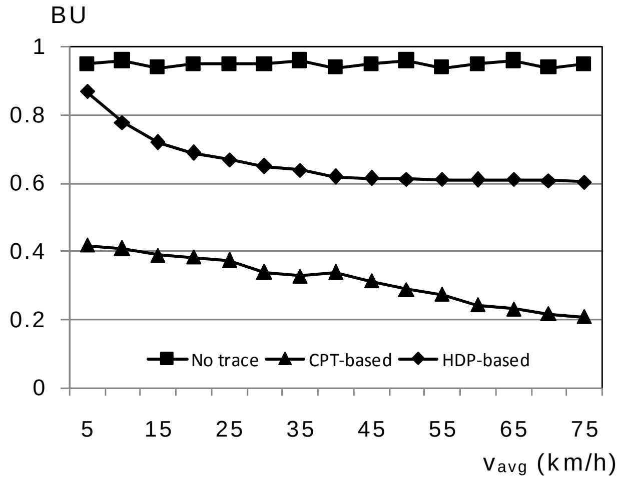


Fig. 19. Bandwidth Utilization for the whole network for different values of average speed.

In the CPT/HDP-based cases, the number of requests increases, because given a higher speed value, the mobile host will visit a larger number of cells. In fig. 19, the trend of the bandwidth utilization (BU) is illustrated, as a function of the average node speed. The term BU is evaluated as the ratio of used channels and the total available ones (passive reserved channels are considered unused).

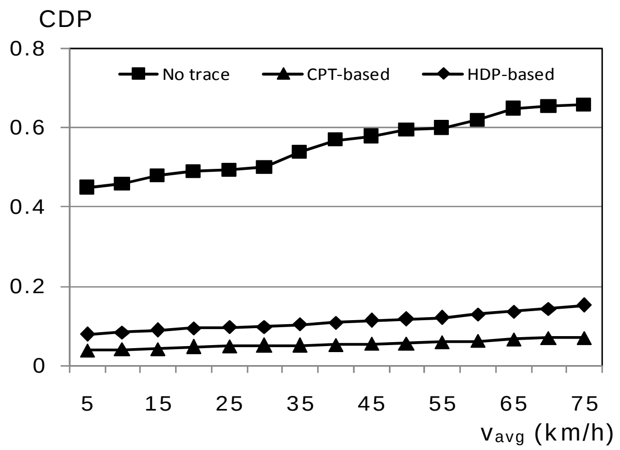


Fig. 20. Call Dropping Probability (CDP) for different values of average speed.

In the No-trace case, the trend is constant, since the reservations are made only on the current cell (without any prediction): the utilization is the maximum reachable, because there is neither additional overhead (for signaling to remote cells), nor in-advance reservations, which waste an enormous amount of bandwidth (it remains unused). It is possible to see, in addition, that the CPT-based case wastes many resources: there is a huge amount of passive (in-advance) reservations as well as a high overhead. The HDP-based approach is able to guarantee a good trade-off between the No-trace and the CPT-based approach. The utilization decreases as a function of the average speed, due to the higher number of visited cells (and thus a higher number of passive reservations).

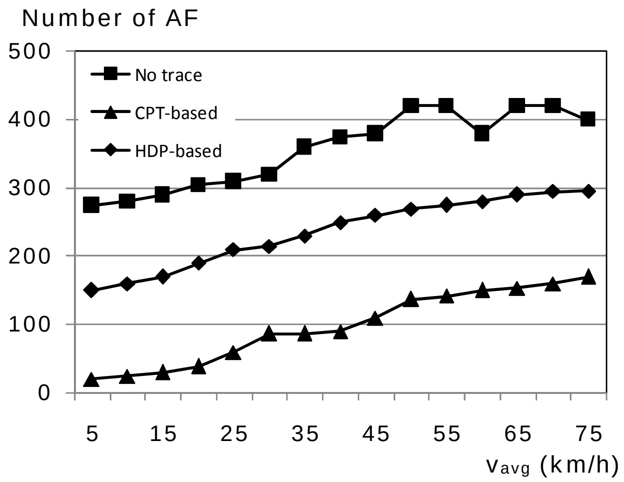


Fig. 21. Average number of admitted flows (AF) for different values of average speed.

Fig. 20 depicts the trend of the CDP, which is one of the main parameters which should be optimized in QoS infrastructure wireless networks. Regarding the No-trace approach, there are no passive reservations, so no bandwidth is guaranteed after a hand-over; in this way, the CDP is high and increases for larger speed values. Trace analyses are necessary, in order to minimize the CDP: with the CPT-based approach the maximum CDP is around 0.08, while for the HDP-based approach it is around 15. There is a gain of about 55%.

The last figure (fig. 21) illustrates the average number of admitted flows: in general, there is an increasing trend in as a function of the average speed. For the No-trace case, since the CAC is made only on the current cell, a higher number of flows can access the network. When a trace-analysis is made, there is a lower number of admitted flows, due to the high amount of passive in-advanced reservations.

V. CONCLUSIONS

In the proposed paper, an in-depth description of trace-file analysis in infrastructure networks has been introduced. Differently from the previous works, we have illustrated how the theoretical conclusions can be applied to a more practical approach, especially for the results which can be obtained when the information contained in a stored trace-file is used to enhance network performance. These features may be suitable, for example, for dimensioning the coverage areas of a cellular network or to supply some critical details to ICT company owners, when making important decisions. A statistical approach for deriving cell permanence time and directional knowledge has been illustrated, as well as the possible benefits which could be introduced into the system, if a proper knowledge is extracted from mobility traces. We have shown that such kinds of approaches bring new knowledge into the considered network, enhancing the overall quality of the perceived services. Our treatment is of a general application, and represents a starting point for future research and innovations in wireless infrastructure networks, where there is the possibility of optimizing the average system utilization, the CDP and the number of admitted requests, on the basis of the target QoS which should be offered by the network. This paper also represents a way for considering a practical approach to network systems analysis, especially when user habits need to be taken into account.

VI. REFERENCES

[1] Riaz, Z.; Durr, F.; Rothermel, K., "Optimized location update protocols for secure and efficient position sharing", Networked Systems (NetSys), 2015 International Conference and Workshops on, pp. 1 - 8, DOI: 10.1109/NetSys.2015.7089083.

[2] R. Jiang; Y. Zhu; X. Wang; Ni, L.M., "TMC: Exploiting Trajectories for Multicast in Sparse Vehicular Networks", Parallel and Distributed Systems, IEEE Transactions on, 2015, Volume: 26, Issue: 1, pp. 262 - 271, DOI: 10.1109/TPDS.2014.2307852.

[3] Y. Zhu, Y. Wu, B. Li, "Trajectory Improves Data Delivery in Urban Vehicular Networks", Parallel and Distributed Systems, IEEE Transactions on, 2014, Volume: 25, Issue: 4, pp. 1089 - 1100, DOI: 10.1109/TPDS.2013.118.

[4] C. Song, Z. Qu, N. Blumm, and A.-L. Barabasi, “Limits of predictability in human mobility,” Science Magazine, Vol. 327(5968), pp. 1018-1021, 2010, DOI: 10.1126/science.1177170.

[5] Socievole, A., Yoneki, E., De Rango, F., Crowcroft, "J.ML-SOR: Message routing using multi-layer social networks in opportunistic communications", Computer Networks, vol. 81, pp.201-219, 2015.

[6] Furletti, B., Gabrielli, L., Renso, C., & Rinzivillo, S. (2013, October). Analysis of GSM calls data for understanding user mobility behavior. In Big Data, 2013 IEEE International Conference on (pp. 550-555). IEEE.

[7] Zhou, C., & Huang, B. (2015, December). Exploration of Collective Pattern to Improve Location Prediction of Mobile Phone Users. In 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity) (pp. 54-59). IEEE.

[8] Pérez-Penichet, C., Conde, Â., & Moreira, A. (2012). Human mobility analysis by collaborative radio landscape observation. In Multidisciplinary Research on Geographical Information in Europe and Beyond Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012.

[9] Gabrielli, L., Furletti, B., Giannotti, F., Nanni, M., & Rinzivillo, S. (2014). Use of Mobile Phone Data to Estimate Visitors Mobility Flows. In Software Engineering and Formal Methods (pp. 214-226). Springer International Publishing.

[10] Leca, C. L., Tută, L., Nicolaescu, I., & Rincu, C. I. (2015, November). Recent advances in location prediction methods for cellular communication networks. InTelecommunications Forum Telfor (TELFOR), 2015 23rd (pp. 898-901). IEEE.

[11] Mohamed, A., Onireti, O., Hoseinitabatabaei, S. A., Imran, M., Imran, A., & Tafazolli, R. (2015, June). Mobility prediction for handover management in cellular networks with control/data separation. In 2015 IEEE International Conference on Communications (ICC) (pp. 3939-3944). IEEE.

[12] Singh, R., Srinivasan, M., & Murthy, C. S. R. (2015, January). A learning based mobile user traffic characterization for efficient resource management in cellular networks. In 2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC) (pp. 304-309). IEEE.

[13] Tropea, M., & Marano, S. (2013, June). A Novel Pattern Prediction and Bandwidth Management Scheme in Wireless Networks with Mobile Hosts. InVehicular Technology Conference (VTC Spring), 2013 IEEE 77th (pp. 1-5). IEEE.

[14] The G-MoN application - https://play.google.com/store/apps/details?id=de.carknue.gmon2&hl=it

[15] Furletti, B.; Gabrielli, L.; Renso, C.; Rinzivillo, S., "Analysis of GSM calls data for understanding user mobility behavior", Big Data, 2013 IEEE International Conference on, pp. 550 - 555, DOI: 10.1109/BigData.2013.6691621.

[16] Boyle, C.W.; Kovvali, S.K., "Distributed Data Collection, Correlation and Data Reduction for Iterative Streaming Analytics in Wireless Mobile Networks", Big Data Computing Service and Applications (BigDataService), 2015 IEEE First International Conference on, pp. 384 - 391, DOI: 10.1109/BigDataService.2015.65.

[17] F. De Rango, P.Fazio, S.Marano, “Cell Stay Time Analysis under Random Way Point Mobility Model in WLAN Networks”, IEEE Communication Letters, Vol.10, Issue 11, pp.763-765, November 2006.

[18] C. Montgomery, “Applied statistics and probability for engineers”, Third Edition, Wiley, 2003.

[19] Martinez F.J., Cano J.C., Calafate C.T., Manzoni P., "CityMob: A Mobility Model Pattern Generator for VANETs", ICC Workshops 2008, pp. 370 – 374, Beijing.

[20] Norman R. Draper, Harry Smith, “Applied Regression Analysis”, Wiley Series in Probability and Statistics, 3rd edition, 1998.

[21] Marsaglia, G. and Tsang, W. W., “A simple method for generating gamma variables”, ACM transactions on mathematical software, 28(3):363–372.

[22] Buyukcorak, S.; Kurt, G.K.; Cengaver, O., "A Probabilistic Framework for Estimating Call Holding Time Distributions", Vehicular Technology, IEEE Transactions on, 2014, Volume: 63, Issue: 2, pp. 811 - 821, DOI: 10.1109/TVT.2013.2275081.

[23] Talukdar A.K., Badrinath B.R., Charya A.A., "MRSVP: A Resource Reservation Protocol for an Integrated Services Network with Mobile Hosts", Wireless Networks, Kluwer Journal, pp.5-19, 2001.

[24] G. E. P. Box and Mervin E. Muller, "A Note on the Generation of Random Normal Deviates", The Annals of Mathematical Statistics (1958), Vol. 29, No. 2 pp. 610-611.

[25] Landstrom, A., Jonsson, H., Simonsson, A., "Voronoi-Based ISD and Site Density Characteristics for Mobile Networks", in Proc IEEE VTC Fall, 2012, pp. 1-5.

[26] Meetei, K.P., George, A., "Handoff management in wireless networks using predictive modeling", Communications (NCC), 2011 National Conference on, Page(s): 1 - 5.

[27] Y. Chon, E. Talipov, H. Shin, H. Cha. (2014, Mar.). SmartDC: Mobility Prediction-Based Adaptive Duty Cycling for Everyday Location Monitoring. IEEE Trans. on Mobile Computing, 13(3), 2014, pp. 512-525.

[28] K. Li, "Analysis of Distance-Based Location Management in Wireless Communication Networks", Parallel and Distributed Systems, IEEE Transactions on, Volume: 24, Issue: 2, 2013 , Page(s): 225 - 238.

[29] H. Feng, C. Liu, Y. Shu, O. Yang, "Location Prediction of Vehicles in VANETs Using A Kalman Filter", Wireless Pers. Commun. (2015) vol. 80, pp.543–559.

[30] Yang Wang ; Liusheng Huang ; Tianbo Gu ; Hao Wei ; Kai Xing ; Junshan Zhang, "Data-driven traffic flow analysis for vehicular communications", INFOCOM, 2014 Proceedings IEEE, DOI: 10.1109/INFOCOM.2014.6848138, 2014, Page(s): 1977 - 1985.

BIOGRAPHIES

Peppino Fazio received the degree in computer science engineering in May 2004. Since November 2004 he has been a Ph.D student in Electronics and Communications Engineering at the University of Calabria and he has got the Ph.D. in January 2008; at the moment he is a research fellow at DEIS Department of University of Calabria, after many collaborations with Department of "Universidad Politecnica de Valencia". His research interests include mobile communication networks, QoS architectures and interworking wireless and wired networks, mobility modeling for WLAN environments and mobility analysis for prediction purposes.

Floriano De Rango graduated in Computer Science in October 2000, and a Ph.D. in electronics and telecommunications engineering in January 2005, both at the University of Calabria, Italy. From January 2000 to October 2000 he worked in the Telecom Research LAB C.S.E.L.T. in Turin with a scholarship. From March 2004 to November 2004 he was visiting researcher at the University of California at Los Angeles (UCLA). He is now Assistant Professor at D.E.I.S. Department, University of Calabria. He was recipient of the Young Researcher Award in 2007 and is reviewer TPC member for many International Conferences and reviewer for many journals such as IEEE Communication Letters, JSAC, WINET etc. His interests include Satellite networks, IP QoS architectures, Adaptive Wireless Networks, Ad Hoc Networks and Pervasive Computing.

Mauro Tropea was born in 1975 and graduated in Computer Science Engineering at the University of Calabria, Italy, in 2003. Since 2003 he has been with the telecommunications research group of D.E.I.S. in the University of Calabria. In 2004 he won a regional scholarship on Satellite and Terrestrial broadband digital telecommunication systems. Since November 2005 he has a Ph.D student in Electronics and Communications Engineering at University of Calabria. His research interests include satellite communication networks, QoS architectures and interworking wireless and wired networks, mobility models and vehicular issues.

Miroslav Voznak born in 1971 is an associate professor with the Department of Telecommunications, Technical University of Ostrava, Czech Republic and foreign professor with Ton Duc Thang University in Ho Chi Minh City, Vietnam. He received his Ph.D. degree in telecommunications in 2002 at the Technical University of Ostrava. He is a senior researcher in the Supercomputing center IT4Innovations in Ostrava, Czech Republic, a member of editorial boards of several journals and a member of many boards of conferences supported by IEEE such as TSP, INBIS, CN, etc. Topics of his research interests are IP telephony, wireless networks, speech quality and network security.