Performance of Data Traffic in Small Cells Networks with Inter-Cell Mobility *

Philippe Olivier
Orange Labs
44 avenue de la République
92320 Châtillon, France
phil.olivier@orange.com

Alain Simonian
Orange Labs
44 avenue de la République
92320 Châtillon, France
alain.simonian@orange.com

ABSTRACT

We analyze the impact of inter-cell mobility on data traffic performance in dense networks with small cells. To this end, a Markovian queuing model with impatience is proposed as a generic tool that captures mobility through the sojourn time of users in the cell. An approximate model, based on Quasi-Stationary assumptions, is developed to speed up computation. We then show how the performance of a network of small cells is amenable to the application of the generic model to each individual cell; the analysis of a homogeneous network is further reduced to that of a single representative cell. Our approach is applied to derive the throughput of both mobile and static users, along with the probability of handover. Numerical evaluation and simulation results are provided to assess the accuracy of the approach and we show, in particular, that both classes of users benefit from a throughput gain induced by the "opportunistic" displacement of mobile users among cells.

CCS Concepts

•Mathematics of computing → Markov processes; •Networks → Network performance analysis; Mobile networks;

Keywords

Cellular networks; mobility; traffic; performance evaluation

1. INTRODUCTION

To address the permanent increase of mobile traffic, the capacity of networks can be upgraded by a massive deployment of small cells. This is a solution envisaged by network operators in the framework of LTE-A heterogeneous networks [8] or Ultra Dense Networks scenarios for future 5G networks [13]. In such dense networks, however, the amount of handover generated by users mobility will significantly increase with a notable impact on signaling overhead, thus counterbalancing possible throughput gains.

In this context, the present paper aims at evaluating the impact of *inter-cell mobility* on the performance of data traffic in dense networks. Specifically, considering small cells enables us to neglect the possible spatial variations of intra-cell capacities and thus to focus on the impact of inter-cell mobility itself. Furthermore, we decouple the performance evaluation problem from the modeling of the user displacement in the plane, the latter topic being out of scope of the present paper (see [11, 14] for current displacement models). Mobility is here supposed to be captured through the distribution of the users residual sojourn time in a cell, that is, the time a mobile user is physically present in the cell once its transmission has started; this distribution can be seen as the output of user displacement models.

Given this distribution, we construct a flow-level queuing model that allows us to derive the essential performance metrics in each cell, namely the mean throughput and the handover probability. The generic tool of our model is a multi-class Processor Sharing (PS) queue with "impatient" customers; the impatience here accounts for the mobility of customers from cell to cell. An approximate model is derived for this queue in order to replace the exact Markovian model when the latter requires exceeding computation time. This generic model can then be applied to each individual cell to solve the set of flow equations, which characterize the handover rates between cells, and compute the performance indicators.

The PS queue with impatience and one customer class has been early addressed in [7], where several asymptotic regimes are given for the reneging probability. For the multi-class PS queue in overload condition (with identical impatience rates but distinct service rates), an asymptotics for the reneging probability has been provided in [9]. The analysis of the stable multi-class PS queue with distinct impatience rates, however, has not received so far a significant contribution.

In the literature, throughput gains induced by mobility in cellular networks are generally related to the spatial variations of capacity inside the cells, which permits an opportunistic use of favorable transmission conditions by mobile users [1, 3, 4, 6, 10]. These papers also base their evaluation on flow-level traffic modeling but address mobility through a spatial Markov process where users jump between distinct capacity zones in the cells. In [1], throughput gains are ob-

*Part of this work has been carried out in the framework of IDEFIX project, funded by the ANR under the contract number ANR-13-INFR-0006

VALUETOOLS, October 26–28, 2016, Taormina, Italy
© 2016 ACM. ISBN 123-4567-24-567/08/06…$15.00
DOI: 10.475/123_4

VALUETOOLS 2016, October 25-28, Taormina, Italy
Copyright © 2017 FALSE
DOI 10.4108/eai.25-10-2016.2266520

served for the Round-Robin and opportunistic "max C/I" scheduling schemes with a main focus on intra-cell mobility. In [3, 4], the average throughputs of different classes of users are estimated through upper and lower bounds associated with fluid limits and quasi-stationary regimes, respectively. By means of an approximate model, [10] computes the mean flow throughput of a single class of users and highlights a throughput gain created by intra- and inter-cell mobility altogether. In a complementary way, [6] focuses on the characterization of network capacity with different scenarios of intra-cell and inter-cell mobility.

The present work is an attempt to analyze a queuing model to derive multi-class performance indicators, namely the throughput gains as well as the handover rates induced by inter-cell mobility alone, and to apply it to a whole network of small cells. In the following, a generic one-cell Markovian model is first constructed in Section 2 to derive such indicators; a heuristic Quasi-Stationary approximation is then proposed to alleviate computation. Section 3 presents our approach to model networks with mobility. Section 4 provides numerical results, including simulation experiments, to validate the approach. Finally, Section 5 draws conclusions and summarizes our main achievements.

2. GENERIC QUEUEING MODEL

As a first step, we consider a single cell whose performance model may be used as the generic tool to further analyze the impact of inter-cell mobility in a network. This section is an extension of the single cell model proposed in [2] and proposes a different approximation, the so-called *QS approximation model* developed in Section 2.2, which allows us to quantify the performance of various classes of customers.

2.1 A PS Queue with Impatience

The considered cell is supposed to be "small", i.e., of limited range so that its transmission capacity C can be assumed spatially constant; this capacity is viewed as an input parameter accounting for radio and interference conditions in the considered cellular network. We handle downlink traffic only and suppose that capacity C is equally shared among all active users present in its service area, as implemented by means of a Round-Robin scheduling. Following this fair sharing policy, the system occupancy (defined by the number of active transmissions at any time) can then be modeled by a Processor-Sharing (PS) queue [5].

We consider K classes of users which may have different traffic characteristics and mobility behaviours. The users generate requests for transmission in the cell according to Poisson processes with respective arrival rate λ_k, $k = 1, ..., K$. In this open queuing system, the arrival rates are supposed to be exogenous input parameters. How the arrival rates may account for the handover traffic flows from cell to cell will be addressed in Section 3. Class-k users have i.i.d. transmission requests of data volume Σ_k with mean σ_k, hence a service rate $\mu_k = C/\sigma_k$. Since customers may actually leave the cell during their communications, we call T_k the *remaining* sojourn time of a mobile user, i.e., the time duration he physically stays in the cell after the transmission has started. We finally denote by $\theta_k = 1/\mathbb{E}(T_k)$ the mean cell departure rate of class-k users, called class-k *mobility rate*; any class k where $\theta_k = 0$ will be called *static*.

Given this framework, the cell occupancy can be described by the K-dimensional process $\mathbf{N}(t) = (N_1(t), ..., N_K(t))$, $t \geq 0$, where $N_k(t)$ denotes the number of ongoing class-k data transfers in the cell at time t. The evolution of this process is represented by a Processor-Sharing queue with impatience, the "impatient" customers here corresponding to mobile users that may leave the system before their service completion within the given cell.

We assume that Σ_k and T_k are exponentially distributed with parameters $1/\sigma_k$ and θ_k, respectively. The process $(\mathbf{N}(t))_{t\geq 0}$ is then Markovian in the state space $\mathbb{N}^K$; from state $\mathbf{n} = (n_1, ..., n_K)$ and for $\mathbf{e}_k = (0, ..., 1, ...0)$ with 1 at the k-th component, we can reach state $\mathbf{n}+\mathbf{e}_k$ with transition rate λ_k, or state $\mathbf{n}-\mathbf{e}_k$ with transition rate $n_k\mu_k/L(\mathbf{n})+n_k\theta_k$ if $n_k > 0$, $L(\mathbf{n}) = \sum_{1\leq j\leq K} n_j$ denoting the total number of active users.

PROPOSITION 2.1. ***Let** $\rho_k = \lambda_k/\mu_k$ **be the offered load of class** k. **The Markov occupancy process** $(\mathbf{N}(t))_{t\geq 0}$ **has a stationary regime if and only if***

$$\rho_S = \sum_{k\in S} \rho_k < 1, \tag{1}$$

where** S **is the set of static user classes.

Although condition (1) is intuitive, its derivation is not straightforward; we refer to [12] for the detailed proof which exhibits a proper Lyapounov function ensuring the sufficiency of (1). Now, given (1), the equilibrium equations of process $(\mathbf{N}(t))_{t\geq 0}$ in stationary regime read

$$\sum_{k=1}^{K} \left[\lambda_k + n_k\left(\frac{\mu_k}{L(\mathbf{n})} + \theta_k\right)\right] \Pi(\mathbf{n}) = \sum_{k=1}^{K} \lambda_k\, \Pi(\mathbf{n} - \mathbf{e}_k) + \sum_{k=1}^{K} (n_k + 1)\left(\frac{\mu_k}{L(\mathbf{n}) + 1} + \theta_k\right) \Pi(\mathbf{n} + \mathbf{e}_k) \tag{2}$$

where $\Pi(\mathbf{n}) = \mathbb{P}(\mathbf{N} = \mathbf{n})$ and with $\sum_{n\in\mathbb{N}^K} \Pi(\mathbf{n}) = 1$. For given k, multiplying each equation of (2) by n_k and then summing over all state vectors $\mathbf{n} \in \mathbb{N}^K$ provides the following relation.

COROLLARY 2.1. ***For any class** k**, the average arrival and departure rates verify the conservation law***

$$\lambda_k = \mu_k\, \mathbb{E}\left(\frac{N_k\, \mathbf{1}_{N_k>0}}{L(\mathbf{N})}\right) + \theta_k\, \mathbb{E}(N_k). \tag{3}$$

Process $(\mathbf{N}(t))_{t\geq 0}$ is not reversible unless all classes are static; its stationary distribution $\Pi(\cdot)$ is thus not amenable to a simple closed form. Though we can evaluate this distribution by solving system (2) numerically and derive the performance indicators of interest. We define two performance indicators per user-class, the *average throughput* and the *handover probability*. Considering data (elastic) traffic, the user-perceived QoS can be measured by the average throughput defined as the ratio of the mean volume of transferred data to the mean transfer time [5]. We also define the *handover probability* for class-k users as the proportion of users that exit the cell before the completion of their transmission, i.e., the ratio of the mean handover rate λ_k^{Out} to the mean flow arrival rate λ_k. The latter definitions read

$$\gamma_k \triangleq \frac{\mathbb{E}(X_k)}{\mathbb{E}(\Delta_k)}, \quad H_k \triangleq \frac{\lambda_k^{Out}}{\lambda_k}, \qquad \forall\, k = 1, ..., K, \tag{4}$$

where X_k denotes the part of the total data volume Σ_k which is actually transferred by a class-k user during its transmission time Δ_k (less than or equal to T_k) in the cell.

PROPOSITION 2.2. ***For any class** k**, the throughput** γ_k **and the handover probability** H_k **are given by***

$$\gamma_k = C\left(\frac{\rho_k}{\mathbb{E}(N_k)} - \frac{\theta_k}{\mu_k}\right), \qquad H_k = \frac{\mathbb{E}(N_k)\,\theta_k}{\lambda_k}, \tag{5}$$

which depend on the mean number of class-**k **users only. Moreover, the relation

$$H_k = \frac{\theta_k\,\sigma_k}{\gamma_k + \theta_k\,\sigma_k} \tag{6}$$

holds between** H_k **and** γ_k**.

PROOF. By Little's law, $\mathbb{E}(N_k)/\lambda_k$ equals the mean time $\mathbb{E}(\Delta_k)$ to transfer the average volume $\mathbb{E}(X_k)$, hence

$$\gamma_k = \frac{\lambda_k \mathbb{E}(X_k)}{\mathbb{E}(N_k)}. \tag{7}$$

The average transferred volume $\mathbb{E}(X_k)$ is not directly computable from distribution $\Pi(\cdot)$. We can nevertheless observe that $\lambda_k \mathbb{E}(X_k)$ represents the *carried* traffic intensity for class k. In a way similar to the derivation of Little's formula, this traffic intensity can be shown to equal the mean bandwidth $\mathbb{E}(\phi_k)$ allocated to class-k users, so that (7) eventually reads

$$\gamma_k = \frac{\mathbb{E}(\phi_k)}{\mathbb{E}(N_k)} = \frac{C}{\mathbb{E}(N_k)}\,\mathbb{E}\left(\frac{N_k\,\mathbf{1}_{N_k>0}}{L(\mathbf{N})}\right). \tag{8}$$

Using conservation law (3), expression (5) for γ_k then follows. As to the handover probability, the mean handover rate λ_k^{Out} is by definition $\mathbb{E}(N_k)\,\theta_k$, hence (5) for H_k.

Relation (6) directly follows from eliminating $\mathbb{E}(N_k)$ between both expressions (5). ☐

2.2 Quasi-Stationary Approximation

Consider now the important case where users are gathered into two classes, namely static users (with class index "S") and mobile users (with index "M"). We thus set $\theta_S = 0$, $\theta_M > 0$, and define loads $\rho_S = \lambda_S/\mu_S$, $\rho_M = \lambda_M/\mu_M$.

Although the Markovian model can be solved numerically, it does not provide explicit expressions for performance indicators. To also circumvent the explosive computation time of either the simulation or the numerical resolution when the load increases, we here develop an approximation framework suitable for this two-class scenario. The results of this approach can be summarized as follows.

PROPOSITION 2.3. ***Let** $A > 0$ **be the solution to***

$$e^{-A}\,(1-\rho_S) = \frac{\theta_M}{\mu_M}A + 1 - \rho_S - \rho_M. \tag{9}$$

The throughputs** γ_S **and** γ_M **can be approximated by

$$\gamma_S \approx \frac{C(1-\rho_S)}{1+A}, \quad \gamma_M \approx \sigma_M\left(\frac{\lambda_M}{\mathbb{E}(N_M)} - \theta_M\right) \tag{10}$$

with** $\mathbb{E}(N_M) = \sum_{(\ell,m)\in\mathbb{N}^2} m\Psi(m|\ell)q(\ell)$**, where

$$\left[\begin{array}{l} q(\ell) = e^{-A\rho_S}\,\rho_S^{\ell}(1-\rho_S)\displaystyle\sum_{k=0}^{\ell}\binom{\ell}{k}\frac{[A(1-\rho_S)]^k}{k!}, \\ \Psi(m|\ell) = \Psi(0|\ell)\left(\dfrac{\rho_M}{\rho_\theta}\right)^m \dfrac{1}{m!}\displaystyle\prod_{k=1}^{m}\frac{\ell+k}{\ell+k+1/\rho_\theta} \end{array}\right. \tag{11}$$

for all** $(\ell,m)\in\mathbb{N}^2$**, with constant** $\rho_\theta = \theta_M/\mu_M$ **and where the factor** $\Psi(0|\ell)$ **is determined by the normalization condition** $\sum_{m\geq 0}\Psi(m|\ell) = 1$ **for all** ℓ**.

The handover** H **is calculated by applying (6) to the value** γ_M **given in (10).

PROOF. As a preamble, applying conservation law (3) to each class, we obtain

$$\rho_S = \mathbb{E}\left(\frac{N_S\mathbf{1}_{N_S>0}}{L(\mathbf{N})}\right), \rho_M = \mathbb{E}\left(\frac{N_M\mathbf{1}_{N_M>0}}{L(\mathbf{N})}\right) + \frac{\theta_M}{\mu_M}\mathbb{E}(N_M).$$

Then, noting that (with $L(\mathbf{N}) = N_S + N_M$)

$$\mathbb{E}\left(\frac{N_S\,\mathbf{1}_{N_S>0}}{L(\mathbf{N})}\right) + \mathbb{E}\left(\frac{N_M\,\mathbf{1}_{N_M>0}}{L(\mathbf{N})}\right) = \mathbb{E}(\mathbf{1}_{\mathbf{N}\neq(0,0)}),$$

we obtain the relation

$$\mathbb{E}(N_M) = \frac{\mu_M}{\theta_M}\left(\rho_S + \rho_M + \Pi(0,0) - 1\right). \tag{12}$$

Now, the underlying idea for the proposed approximation is a Quasi-Stationary (QS) assumption which can be applied in two steps.

(I) First, note that the largest the load, the highest tendency of mobile users to leave the cell before the end of their transmission. The stationary distribution of their number N_M may thus be approximated by that of a queuing system without any constraint, i.e., a Poisson distribution describing the occupancy of an M/G/∞ queue. The mean value $A = \mathbb{E}(N_M)$ of this Poisson distribution remains to be evaluated. Given a number $N_M = m$ of mobile users in the cell, we assume it to remain constant in time (*Quasi-Stationary*) so that the number N_S of static users evolves as the occupancy process of a Processor Sharing queue with a set of m permanent customers. This enables us to obtain the conditional distribution $\Phi(\cdot|m)$ of N_S, which is readily derived from the local balance equations of the one-dimensional Markov process $(N_S(t))$, that is,

$$\Phi(\ell|m) = \Phi(0|m)\,\rho_S^{\ell}\,\frac{(\ell+m)!}{\ell!\,m!}, \quad \ell \geq 0, \tag{13}$$

where $\Phi(0|m)$ is determined by the normalization condition $\sum_{\ell\geq 0}\Phi(\ell|m) = 1$, hence $\Phi(0|m) = (1-\rho_S)^{m+1}$. Deconditioning on N_M, the joint distribution of (N_S, N_M) is then

$$\Pi^{(1)}(\ell,m) = e^{-A}\,\frac{A^m}{m!}\binom{\ell+m}{m}\rho_S^{\ell}(1-\rho_S)^{m+1} \tag{14}$$

for $(\ell,m)\in\mathbb{N}^2$ and, in particular, $\Pi(0,0) = e^{-A}\,(1-\rho_S)$. Using relation (12) and the latter value of $\Pi(0,0)$, we deduce that the unknown mean A is determined as the unique positive solution to the implicit equation (9).

Applying relation (5) with $\mathbb{E}(N_M) = A$ then provides the expression (10) for γ_M. As regards static users, we use the distribution (14) to derive $\mathbb{E}(N_S) = \rho_S\mathbb{E}(N_S) + \rho_S\mathbb{E}(N_M+1)$, hence $\mathbb{E}(N_S) = (1+A)\rho_S/(1-\rho_S)$ and the application of (5) with $\theta_S = 0$ eventually leads to formula (10) for γ_S.

(II) This first step, however, does not provide accurate enough values for throughputs (in particular, γ_M) when comparing them to that of the exact Markovian Model. We thus iterate the QS approximation scheme by considering now that mobile users "see" a succession of stationary regimes conditioned on the number of static users. We first compute the marginal distribution $q(\cdot)$ of N_S from the joint distribution (14) as

$$q(\ell) = \sum_{m\geq 0} e^{-A}\,\rho_S^{\ell}(1-\rho_S)\binom{\ell+m}{m}\frac{[A(1-\rho_S)]^m}{m!}$$

for all $\ell \geq 0$. The latter formula reduces after simple manipulations to the expression for $q(\ell)$ given in (11). For any given $N_S = \ell$, the number N_M is then supposed to evolve as the occupancy process of a Processor Sharing queue with impatient (mobile) customers and a set of ℓ permanent customers. The conditional distribution $\Psi(\cdot|\ell)$ for the number of mobile users thus verifies the local balance equations

$$\lambda_M \Psi(m-1|\ell) = m\left(\frac{\mu_M}{\ell+m} + \theta_M\right)\Psi(m|\ell) \qquad (15)$$

for all $m > 0$. By recursion, we obtain expression (11) for $\Psi(m|\ell)$, $m \geq 0$, with $\Psi(0|\ell)$ given by the normalization condition. By deconditioning on N_S, the joint distribution of (N_S, N_M) is now given by

$$\Pi^{(2)}(\ell, m) = \Psi(m|\ell)\, q(\ell), \quad (\ell, m) \in \mathbb{N}^2. \qquad (16)$$

The throughput γ_S in this second step QS approximation will leave expression (10) unchanged, since $\gamma_S = C\rho_S/\mathbb{E}(N_S)$ and the marginal distribution $q(\cdot)$ of N_S is now known *a priori*. Besides, relations (3) and (5) again enable us to write γ_M as in (10) where the mean $\mathbb{E}(N_M)$ has to be re-evaluated from the joint distribution (16). $\square$

The latter evaluations associated with the second step distribution $\Pi^{(2)}$ indeed provide more accurate results (see Sections 4.1.6 and 4.2) than that obtained at the first step. Besides, it has been observed that further iterations of this QS approximation scheme do not bring better accuracy.

3. NETWORK WITH MOBILITY

3.1 A closed network of queues

Consider now a cellular network of I cells with possibly distinct capacities. Users from K traffic classes may appear and move during their communications. When leaving a cell during transmission, users join one of the neighboring cells according to some routing probabilities. They consequently generate supplementary flows of new arrivals, hereafter called *handover arrivals*, which are to be added to *fresh arrivals* in each cell.

We assume that class-k users generate requests for transmission in cell i according to a Poisson process with rate $\lambda^0_{i,k}$, $i = 1, ..., I$, $k = 1, ..., K$; this corresponds to the fresh traffic offered to cell i. To account for class-k users that became active outside cell i and experienced one or more handovers, the total flow arrival to cell i is written as

$$\lambda_{i,k} = \lambda^0_{i,k} + \lambda^{In}_{i,k} \qquad (17)$$

where $\lambda^{In}_{i,k}$ denotes the handover arrival rate from neighboring cells. For all i and k, we will assume that the handover arrival process to cell i from class-k users can be approximated by a Poisson process so that it can be superposed to the fresh arrivals to build up a total Poisson arrival process with rate $\lambda_{i,k}$ given in (17). All Poisson processes introduced above are assumed to be mutually independent.

This Poisson assumption is fundamental in our modeling approach, which allows a local description of the traffic flows at each cell as explained below. This notably simplifies the global description of the system by reducing it to a network of queues which is *closed* regarding the handover flows. This is, in particular, in contrast to the overall multi-class multi-cell process considered in other papers quoted in the Introduction.

The handover arrival rate $\lambda^{In}_{i,k}$ to cell i results from the superposition of handovers departing from neighboring cells according to routing probabilities $p_k(j,i)$ from cell $j \neq i$ to cell i for class-k users. We can write the flow equation

$$\lambda^{In}_{i,k} = \sum_{j \neq i} p_k(j,i) \cdot \lambda^{Out}_{j,k} \qquad (18)$$

where $\lambda^{Out}_{j,k}$ is the handover departure rate from cell j emanating from class k-users. Rate $\lambda^{Out}_{j,k}$ can in turn be considered as an output of a queuing model for cell j (typically, the generic queuing model studied in Section 2) and calculated by means of some *performance function* $\mathcal{F}_{j,k}(.)$, that is,

$$\lambda^{Out}_{j,k} = \mathcal{F}_{j,k}\left(\lambda^{In}_{j,1}, ..., \lambda^{In}_{j,K}\right) \qquad (19)$$

for $j = 1, ..., I$ and $k = 1, ..., K$. In (19), only handover arrival rates are considered as variables, all other intrinsic parameters (such as cell capacities, per-class offered traffic and mobility rates) being kept constant. From (18)-(19), it follows that a stationary network regime can be characterized by a system of $I \times K$ flow equations with unknowns, the handover arrival rates $\lambda^{In}_{i,k}$, namely

$$\lambda^{In}_{i,k} = \sum_{j \neq i} p_k(j,i) \cdot \mathcal{F}_{j,k}\left(\lambda^{In}_{j,1}, ..., \lambda^{In}_{j,K}\right). \qquad (20)$$

The problem of existence and uniqueness of a solution to the non-linear system (20) for a general network is out of the scope of the present paper. As the performance functions $\mathcal{F}_{j,k}$ may not be explicit in terms of input parameters, the practical determination of a solution to (20) involves a numerical iterative procedure, e.g. a fixed-point algorithm.

3.2 The case of a homogeneous network

Now assume that the network is *homogeneous*, with identical cells, in the following sense:

(i) all intrinsic parameters (capacities, arrival rates, ...) are the same for all cells, so that performance functions do not depend on the cell index, that is, $\mathcal{F}_{j,k}(.) = \mathcal{F}_k(.)$;

(ii) handover routing is symmetric, i.e., for each class k, cell i receives handover traffic from a set $\mathcal{J}_k(i)$ of neighboring cells with identical probability $p_k(j,i) = 1/J_k$, where J_k is the common cardinal of sets $\mathcal{J}_k(i)$.

Clearly, any solution to the simpler system

$$\begin{cases} \lambda^{In}_{i,k} = \lambda^{In}_{j,k} = \lambda^{In}_k, & \forall\, i, j, k, \\ \lambda^{In}_k = \lambda^{Out}_k = \mathcal{F}_k\left(\lambda^{In}_1, ..., \lambda^{In}_K\right), & \forall\, k, \end{cases} \qquad (21)$$

will provide a particular solution to (20), hence *the* solution if uniqueness is ensured. For a homogeneous network, the problem thus reduces to the study of a single cell, hereafter called the *representative cell*. System (21) expresses the fact that the outgoing handover traffic is fed back in a balanced way as a supplementary traffic to the ingress of this cell.

In this homogeneous network model with inter-cell mobility, we will define the total and per-class loads by referring to the fresh offered traffic, that is,

$$\rho^0_k = \frac{\lambda^0_k\, \sigma_k}{C} \;\; (k = 1, ..., K), \quad \rho^0 = \sum_{k=1}^{K} \rho^0_k.$$

We conjecture that the stability region is now characterized by $\rho^0 < 1$. This can be understood since mobile users re-enter the system until their transfer is completed.

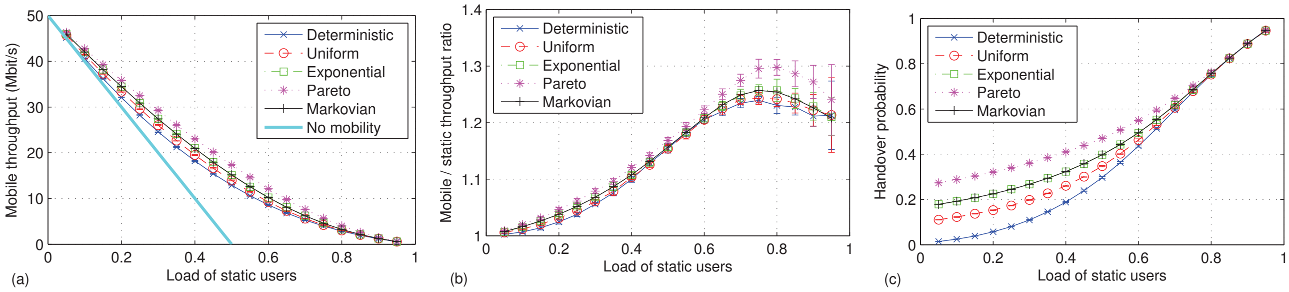

Figure 1: Impact of the mobile sojourn time distribution on (a) mobile user throughput, (b) mobile throughput to static throughput ratio and (c) handover probability, computed from simulation and the Markovian IM (proportion of 50% mobile users and mobility rate $\theta_M = 0.1$ s^{-1}).

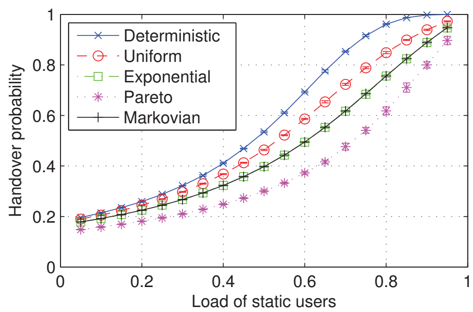

Figure 2: Impact of the flow volume distribution on the handover probability, computed from simulation and the Markovian IM for a proportion of 50% mobile users and a mobility rate $\theta_M = 0.1$ s^{-1}.

CONJECTURE 3.1. ***In the homogeneous network with inter-cell mobility, the condition $\rho^0 < 1$ ensures the existence of a fixed-point solution to equilibrium equation (21), that is, $\lambda_k^{In} = \lambda_k^{Out}$, $\forall\, k = 1, ..., K$.***

Note that, for simplicity, this condition will still be called a *stability condition*, as it translates the fact that the resource must have enough capacity to handle all offered traffic. Although the proof of its sufficiency is open, its necessity is straightforward: in fact, applying conservation equation (3) to any class k and recalling that $\lambda_k = \lambda_k^0 + \lambda_k^{In}$ with $\lambda_k^{In} = \lambda_k^{Out} = \theta_k\, \mathbb{E}(N_k)$, we obtain

$$\rho_k^0 = \mathbb{E}\left(\frac{N_k \mathbf{1}_{N_k>0}}{L(\mathbf{N})}\right), \quad k = 1, ..., K,$$

hence $\rho^0 = \sum_{k=1}^K \rho_k^0 = \mathbb{E}(1_{L(\mathbf{N})>0}) = 1 - \Pi(\mathbf{0}) < 1$.

4. NUMERICAL RESULTS

In this section, we validate the robustness of the proposed approach regarding some important assumptions. For the sake of clarity, we distinguish

- the *Impatience Model* (IM) consisting in the generic Markovian model analyzed in Section 2, where the arrival rate λ_k for any class k is an exogenous input parameter,

- the *Mobility Model* (MM) consisting in the same Markovian model but where, as discussed in Section 3.2, λ_k is the sum of a given exogenous arrival rate λ_k^0 (the fresh offered traffic) and the handover arrival rate λ_k^{In} which exactly balances the outgoing handover rate λ_k^{Out}.

Moreover, we focus on the two-class model introduced in Section 2.2, with all customers gathered either in a static class or in a mobile class with a single mobility rate.

4.1 Impatience Model

4.1.1 Numerical set-up

Event-driven simulations have been performed at the flow level (i.e., neglecting detailed packet-scale phenomena and assuming all system level parameters are summarized in the cell capacity). The purpose of such experiments is to test the sensitivity to the Markovian assumptions made on the users' traffic and mobility. We fix a cell capacity $C = 50$ Mbit/s, a proportion of 50% mobile users and a mean flow volume $\sigma = 12.5$ MB (100 Mbit) for both classes. The mobile users speed is set to $v = 36$ km/h and the mean distance crossed by a mobile user, after its data transfer has started, is $\mathbb{E}(D) = 100$ m, hence a mean exit rate $\theta_M = v/\mathbb{E}(D) = 0.1$ s^{-1}.

The accuracy of results drawn from simulation has been tightly controlled. Specifically, in every configuration, ten independent simulation runs have been performed, generating around 1 million discrete events each, so as to guarantee a confidence interval with range equal to two standard deviations around the mean. The confidence intervals plotted on Fig. 1 and 2 are very small and can hardly be distinguished in most cases. Thus, for simplicity, such confidence intervals will not be represented anymore in further figures.

4.1.2 Influence of sojourn time distribution

We envisage several theoretical distributions for T_M so as to obtain a wide range of values for its variance (with given mean). Beside the exponential distribution, we thus consider the Deterministic, the Uniform and the Pareto distribution with power index 2.

For such distributions of T_M, Fig. 1 depicts the variations of γ_M, γ_M/γ_S and H with varying static load ρ_S. We observe that the throughput of each class is only marginally impacted by the distribution of T_M, indicating that results derived in the Markovian framework remain valid for more realistic sojourn time distributions. The handover H is, however, noticeably more impacted (particularly at low load) and increases with the variance of T_M.

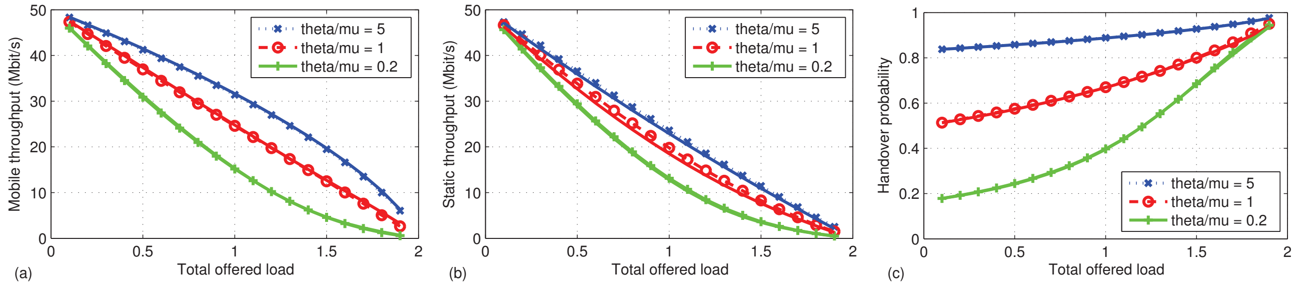

Figure 3: Comparison of the (a) mobile throughput, (b) static throughput and (c) handover probability obtained from the Markovian IM (marks) with those from the QS approximation (lines), for a proportion of 50% mobile users and a normalized mobility rate of 0.2, 1, or 5.

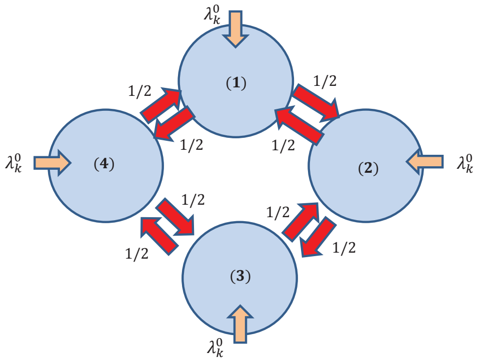

Figure 4: A homogeneous ring network of four identical cells with symmetric routing.

4.1.3 *Influence of flow volume distribution*

Sensitivity tests have also been performed regarding the distribution of the flow volume. Surprisingly, the impact of this distribution (applied to both mobile and static users) on γ_S and γ_M is the same as that of Fig. 1 with an identical parameter setting (discrepancies are indistinguishable and so results are not shown here). On the other hand, the impact on H is reversed, as shown in Fig. 2: the greater the distribution variability, the lower H.

4.1.4 *Throughput gains*

Fig. 1(a) shows a significant throughput gain for mobile users compared to a scenario where all users would be static; besides, Fig. 1(b) highlights a significant gain of mobile users throughput over that of static users. Such throughput gains are expected in this open-loop system: they result from the impatient nature of mobile users who may leave the system without re-entering it, thus decreasing the load (hence the gain w.r.t. the all-static scenario) and tend to do it especially when local congestion occurs (hence the gain mobile/static). As indicated by Fig. 1(c), these throughput gains are obtained at the expense of larger handover.

In relation with these observations, note that the stability region is $\rho_S < 1$, as expected from Proposition 2.1, while it is limited to $\rho_S < 0.5$, in this scenario with 50% static users, when no mobility occurs (the system in that case is a genuine multi-class PS queue).

4.1.5 *Other scenarios*

We have also considered different proportions of mobile users and different values of θ_M. Qualitatively, the corresponding performance curves show a behavior quite similar to that of the above scenario. Based on these observations, we will further pay no more attention to the sensitivity of performance indicators and always refer to the Markovian framework with exponential distributions. We have further observed that the throughput gain of mobile over static users appears to be the greatest when the mobility rate is large (say, more than twice the service rate for mobile users) and when the proportion of mobile users is small (say, 20%).

4.1.6 *Validation of the QS Approximation*

The QS approximation is compared in Fig. 3 to the results of the Markovian IM with the same parameter setting, but all random variables are now supposed to be exponentially distributed and we consider a series of three normalized mobility rates: θ_M equals 0.2, 1 or 5 times the service capacity $\mu_M = 0.5$ s^{-1}. The accuracy of the approximation appears very good for all performance indicators. This has also been checked for other parameter settings and, in particular, for other proportions of the mobile user class.

Interpreting the previous results and figures helps us to assess the impact of the cell size. In fact, assuming a constant speed v, the mean distance the mobile user travels in the cell is $\mathbb{E}(D) = v/\theta_M$; this mean distance is typically of order the radius R of a circular cell. Thus if $v = 90$ km/h for example, the values of θ_M considered above, namely $5\,\mu_M$, μ_M and $0.2\,\mu_M$ (with $\mu_M = 0.5\,\text{s}^{-1}$) respectively correspond to a radius of 10 m, 50 m and 250 m, typical of a Femto, Pico, and Micro cell. As expected, users in Femto cells experience the largest throughput since their mobility rate is the highest. As a counterpart, the handover probability they generate is much higher than that in other types of cells.

4.2 Mobility Model in a Network

To handle mobility in a network, first consider the homogeneous four-cell ring network represented in Fig. 4, where all cells are equivalent with the same capacity and traffic parameters as that used in Section 4.1. All random variables are still assumed to be exponentially distributed, and we consider a series of three normalized mobility rates: θ_M equals 0.1, 1 or 10 times the service capacity $\mu_M = 0.5$ s^{-1}.

Fig. 5 depicts the performance indicators obtained from

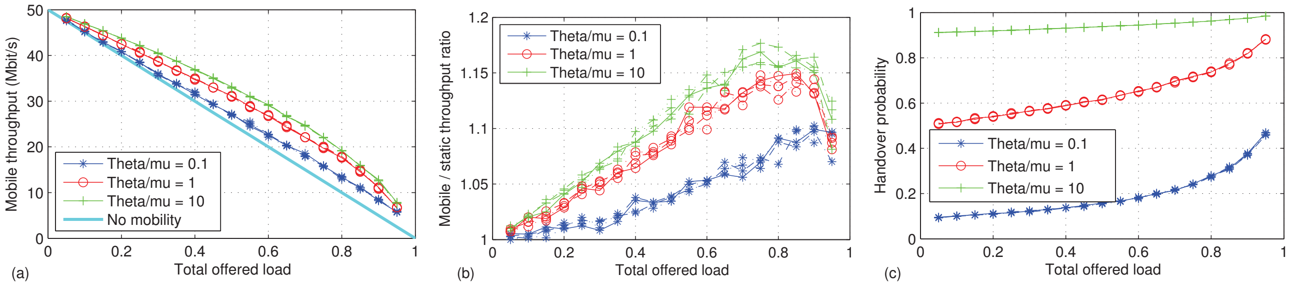

Figure 5: Homogeneous ring network: performance indicators, (a) mobile user throughput, (b) mobile throughput to static throughput ratio and (c) handover probability, obtained from simulation versus the total offered load in each cell (initial proportion of 50% mobile users and $\theta_M/\mu_M = 0.1, 1, 10$).

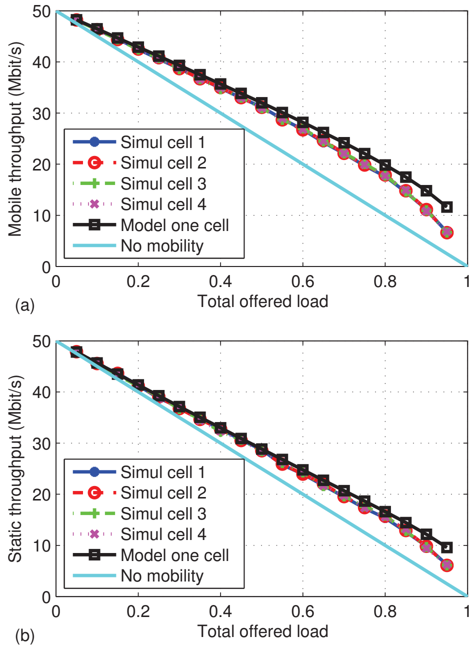

Figure 6: Homogeneous ring network: (a) mobile and (b) static throughput obtained from simulation and the Markovian MM ($\theta_M/\mu_M = 1$).

simulation. For each value of θ_M, the four curves (each corresponding to a given cell) are plotted with the same color code; note that they are almost indistinguishable from each other, except for the ratio γ_M/γ_S.

We observe that the stability region is characterized by $\rho^0 < 1$, as predicted by Conjecture 3.1. From Fig. 5(a)(b), we note that the throughput gains due to mobility increase with the mobility rate. Other complementary results have shown that the mobile/static throughput gain is all the more important that the proportion of mobile users is weak.

The latter simulation results are compared in Fig. 6 to the Markovian MM (applied to the *representative cell*) in the case when $\theta_M = \mu_M$. We observe that the representative cell model provides slightly optimistic throughputs compared to

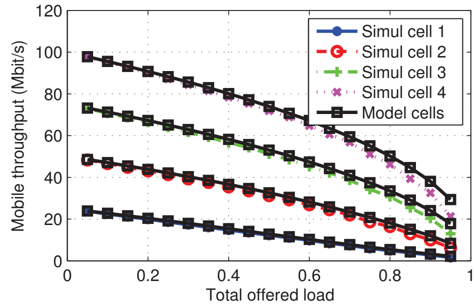

Figure 7: Heterogeneous ring network: mobile throughput obtained from simulation and the generic model ($\theta_M/\mu_M = 10$).

those obtained from simulation. From plots obtained for other values of θ_M/μ_M, we have noted that these discrepancies generally increase with the mobility rate.

The good match between model and simulation results validates our approach for reducing a homogeneous network of small cells to a single representative cell: the critical assumption that the handover traffic flow re-enters the representative cell as a supplementary Poisson flow appears reasonable. We now check this assumption in the case of a heterogeneous network by considering the ring network of Fig. 4 but with different cell capacities: 25, 50, 75 and 100 Mbit/s for cell 1 to 4, respectively. The static/mobile user classes generate fresh traffic in each cell proportionally to its capacity, and mobile users have a mobility rate which is also proportional to the service rate in each cell. The performance indicators are computed for each cell, either from simulation or from the generic model used to solve iteratively the handover flow equations (20). For concision, results are given in Fig. 7 for the mobile throughput only, in the case of 50% mobile users and a ratio $\theta_M/\mu_M = 10$; they again show that our proposed approach (to decorrelate mobility from performance computations in each cell) works quite well.

Note finally that the values of the mobility rate θ_M can be interpreted in terms of various mobile speeds for a given cell size. Specifically, Fig. 8 depicts the static and mobile users throughputs and the handover in terms of the speed v for different values of the total offered load (0.2, 0.5, and 0.8) in

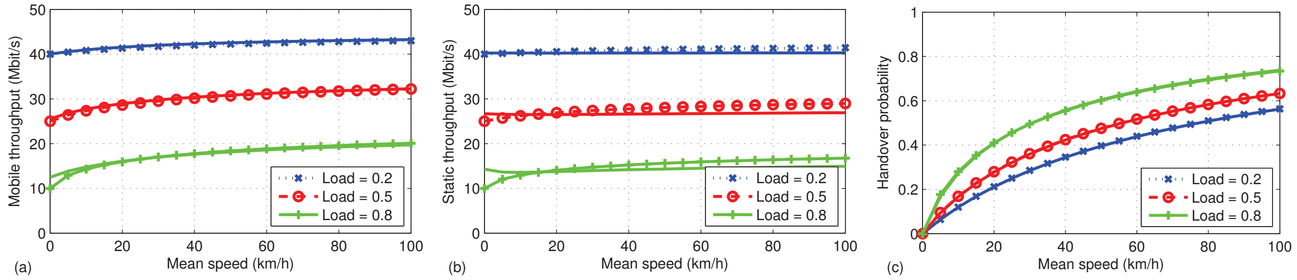

Figure 8: Impact of the users speed on (a) mobile throughput, (b) static throughput and (c) handover probability, obtained from the Markovian MM (marked points) and the QS approximation (plain lines), for an initial proportion of 50% mobile users, cell radius 50 m, and total offered load 0.2, 0.5 or 0.8.

the MM framework. The initial proportion of mobile users is still 50% and the cell radius is 50 m (a Pico cell). Results are derived from the exact Markovian model as well as from the QS approximation, which is shown once again to work nicely enough, and notably for predicting the performance of mobile users. As expected, all performance indicators are increasing functions of the speed. In particular, $H \to 1$ when $v \to \infty$ as predicted by (6); for $v = 0$, both static and mobile users throughputs equal $C(1 - \rho^0)$, the throughput achieved when all users are static.

5. CONCLUSION

We have investigated the impact of inter-cell mobility on the performance of dense networks with small cells. Our approach relies on two main ideas: (1) a simple performance model can be developed to capture mobility on the basis of the multi-class Processor-Sharing queue with impatience; (2) the performance of a network of small cells can be handled by applying the generic model to each individual cell. The main outcomes are: **(i)** the two-class (static / mobile) PS queue can be treated by means of the QS approximation with quite a good level of accuracy; **(ii)** as a step beyond available studies, the handover probability has been evaluated so as to assess the trade-off between throughput gain and signaling overhead due to mobility; **(iii)** both classes of users are shown to benefit from a throughput gain induced by inter-cell mobility; this gain is created by the opportunistic displacement of mobile users within the network according to local load variations in individual cells.

Further work on this topic can be envisaged in two directions: generalize the QS approximate model to an arbitrary number K of traffic classes; extend the performance analysis to cellular networks with spatially varying cell capacities.

6. ACKNOWLEDGMENTS

The authors thank colleagues from the IDEFIX project, and notably Thomas Bonald from Telecom ParisTech, for very fruitful discussions on the considered model.

7. REFERENCES

[1] N. Abbas, T. Bonald, and B. Sayrac. Opportunistic Gains of Mobility in Cellular Data Networks. In *Proc. of WiOpt 13th Intl. Symposium "Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks"*, 2015.

[2] B. Baynat, R.-M. Indre, N. Nya, P. Olivier, and A. Simonian. Impact of Mobility in Dense LTE-A Networks with Small Cells. In *Proc. of VTC2015-Spring Workshop*, Glasgow, Scotland, May 2015.

[3] T. Bonald, S. Borst, and N. Hegde. How Mobility Impacts the Flow-level Performance of Wireless Data Systems. In *Proc. of IEEE Infocom*, 2004.

[4] T. Bonald, S. Borst, N. Hegde, M. Jonckheere, and A. Proutière. Flow-level Performance and Capacity of Wireless Networks with User Mobility. *Queueing Systems*, 63:131–164, 2009.

[5] T. Bonald and A. Proutière. Wireless Downlink Data Channels: User Performance and Cell Dimensioning. In *Proc. of MobiCom'03*, San Diego, USA, Sept. 2003.

[6] S. Borst, A. Proutière, and N. Hegde. Capacity of Wireless Data Networks with Intra- and Inter-cell Mobility. In *Proc. of IEEE Infocom*, 2006.

[7] E. G. Coffman, A. A. Puhalskii, M. I. Reiman, and P. E. Wright. Processor-Shared Buffers with Reneging. *Performance Evaluation*, 19:25–46, 1994.

[8] E. Dahlman, S. Parkvall, and J. Skold. *4G: LTE/LTE-Advanced for Mobile Broadband (2nd ed.)*. Academic Press, 2013.

[9] F. Guillemin, S. ElAyoubi, C. Fricker, P. Robert, and B. Sericola. Controlling Impatience in Cellular Networks Using QoE-aware Radio Resource Allocation. In *Proc. of 27th International Teletraffic Congress*, 2015.

[10] M. K. Karray. User's Mobility Effect on the Performance of Wireless Cellular Networks Serving Elastic Traffic. *Wireless Networks*, 17:247–262, 2011.

[11] X. Lin, R. K. Ganti, P. J. Fleming, and J. G. Andrews. Towards Understanding the Fundamentals of Mobility in Cellular Networks. *IEEE Transactions on Wireless Communications*, 12(4), April 2013.

[12] P. Olivier and A. Simonian. User Performance in Small Cells Networks with Inter-Cell Mobility. arXiv document http://arxiv.org/abs/1605.08353, May 2016.

[13] A. Osseiran et al. Scenarios for 5G Mobile and Wireless Communications: the Vision of the METIS Project. *IEEE Communications Magazine*, 52(5):26–35, 2014.

[14] C. Schindelhauer. Mobility in Wireless Networks. In *Proc. of SOFSEM 2006: Theory and Practice of Computer Science*, Merin, Czech Republic, Jan. 2006.