



**HAL**  
open science

# An Interaction-Based Mobility Model for Dynamic Hot Spot Analysis

Frédéric Morlot, François Baccelli, Salah Eddine Elayoubi

► **To cite this version:**

Frédéric Morlot, François Baccelli, Salah Eddine Elayoubi. An Interaction-Based Mobility Model for Dynamic Hot Spot Analysis. 2012. hal-00689649v1

**HAL Id: hal-00689649**

**<https://hal.science/hal-00689649v1>**

Preprint submitted on 19 Apr 2012 (v1), last revised 8 Jul 2012 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Disclaimer**

This paper is the second version of a paper we published in Infocom, Proceedings IEEE (2010). Equations (5) and (6) (here equations (6) and (7)) have been fixed, some proofs have been simplified and there are some additional results and explanations.

# An Interaction-Based Mobility Model for Dynamic Hot Spot Analysis

Frédéric Morlot  
INRIA - École Normale Supérieure  
Paris, France  
frederic.morlot@polytechnique.org

François Baccelli  
INRIA - École Normale Supérieure  
Paris, France  
francois.baccelli@ens.fr

Salah Eddine Elayoubi  
Orange Labs  
Paris, France  
salaheddine.elayoubi@orange.com

**Abstract**—In this paper, we analyze phenomena related to user clumps and hot spots occurring in mobile networks at the occasion of large urban mass gatherings in large cities. Our analysis is based on observations made on mobility traces of GSM users in several large cities. Classical mobility models, such as the random waypoint, do not allow one to represent the observed dynamics of clumps in a proper manner. This motivates the introduction and the mathematical analysis of a new interaction-based mobility model, which is the main contribution of the present paper. This model is shown to allow one to describe the dynamics of clumps and in particular to predict key phenomena such as the building of hot spots and the scattering between hot spots, which play a key role in the engineering of wireless networks during such events. We show how to obtain the main parameters of this model from simple communication activity measurements and we illustrate this calibration process on real cases.

## I. INTRODUCTION

Traces collected by wireless operators in urban environments during multi event mass gatherings reveal that user mobility is extremely volatile and leads to the formation of hot spots, exemplified on the snapshot of Figure 1 which was gathered during one the 2008 Paris *Fête de la Musique*, where unpredicted gatherings are triggered by a large number of spontaneous street concerts taking place throughout the city. Such observations have also been made during other events, like a summer festival in Poland and Euro2008 soccer matches in Spain and in Romania. A key phenomenon within this context concerns the dynamics of these hot spots, namely their formation and their disaggregation which is random both in space and time.

Accurate and yet tractable user mobility models allowing one to represent these phenomena would significantly improve the robustness and the realism of the simulators and the analytical models used by wireless communication operators. They would hence ameliorate the understanding of the impact of mass mobility on QoS within this setting. Predicting hot spots dynamics could also help in the design and the tuning of adaptive radio resource allocation schemes, where it is fundamental to have an idea of some time characteristics of the system. This allows a better utilization of the corresponding infrastructure and eventually a better QoS and better services to end users. As existing models are not completely adapted to the phenomena we observe on traces, such as the random dynamic hot spot phenomena alluded to above, we propose a

new mobility model taking into account interaction between different users and show by a mathematical analysis based on Markov chain theory that this interaction-mobility model is sufficient to generate random hot spots. For this, we will represent an urban area as a collection of potential hot spot zones. These zones represent potentially attractive points in the city during such mass gatherings.

The basic assumption in our model is that users influence each other. In Section III, we first study the "sheep model", where this mutual influence is particularly strong and where hot spots form at random in time and space, but are stable once formed (i.e. there is no disaggregation of hot spots in this model yet). We start with the 2-zone case, in which we find closed forms for the mean time for the formation of a stable hot spot and for the law of its location. We also give asymptotic estimates for these quantities, and upper bounds on the convergence speed. We then generalize the previous results to an arbitrary number of zones.

Section IV introduces the "sheep and maverick" model, which is more realistic, yet tractable, and which features all key phenomena identified on the traces and alluded to above: the formation of hot spots at random times and places, that we will refer to as *filling*, their disaggregation that we will refer to as *scattering*, and the dynamics between them.

In a last part, we leverage filling and scattering in the "sheep and maverick" model to calibrate our mobility model and show that it can be made consistent with the traces.

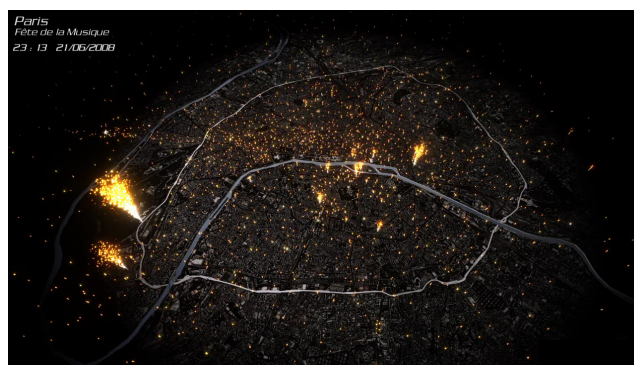


Fig. 1. Communication activity during *Fête de la Musique* in Paris, on June 21<sup>st</sup>, 2008. It is 11:23 pm. One can observe important hot spots in the center (*Châtelet*, *Bastille*, *Saint-Michel*) as well as at the *Parc des Princes* (the city stadium) and *Auteuil* (the city park).

## II. THE NEED FOR A NEW MODEL

First, let us describe more precisely the phenomenon we want to model. A first remark is the existence of hot spots, that follow some dynamics. On the snapshots of Fig. 2, the crowd moves in large numbers outside a zone: we shall call that a *hot spot scattering*. A new or a few new hot spots then build up at other locations: we shall call that a *hot spot filling*.

We now survey existing mobility models and discuss how they could be adapted to cope with these phenomena.

The most popular mobility model used in communication networks is probably the Random Waypoint Model (see [9]). This model is easy to simulate and is analytically tractable in many situations depending of the mobility rules of the individuals on the domain (wrap around, edge reflection...). However, it applies to a single mobile user, or to a collection of users moving independently.

Using this model, in order to take clumping into account, we would have to choose a trip selection rule which gives more mass to potentially attractive zones. As a consequence, the law of the location of a single user would clearly favour these zones. However, such a situation prevents any macroscopic time dynamics. Indeed, with a large number of independent users, all following the same law, each snapshot would be more or less the same, very close to the theoretical location law<sup>1</sup>.

Thus, all clumps would happen at predefined locations and the sudden hot spot scattering as well as the sudden new hot spot aggregation observed in our traces would happen with extremely small probability.

More generally, any multi user model built from the superposition of many independent single user motions will exhibit the same "lack of hot spot dynamics", while a key point observed within traces is a strong dependence between users, since there is clearly a joint motion from a hotspot to another leading to *reinforced clumping* in the corresponding zone.

Nomadic Community Mobility Model or Reference Point Group Mobility Model (see [1]) consist in making small groups of users move in a mean given direction, while the users themselves are allowed to walk here and there around the imposed direction. These models are well adapted for a group of visitors in a museum for instance, where the trajectory is more or less the same for everyone. However, such a system does not allow us to make groups scatter.

Let us now review some ideas about modelling users' interactions. Since we need a strong dependence among users,

<sup>1</sup>This follows directly from Sanov's theorem (see [3], Theorem 6.2.10). Let us denote by  $\mu$  the law of the location of a single user. Let  $X_1, X_2, \dots, X_n$  be the locations of  $n$  users randomly and independently located according to  $\mu$ . Let  $\hat{\mu}_n = \frac{1}{n} \sum_{k=1}^n \delta_{X_k}$  be the empirical measure associated with a snapshot: Then for all given events  $B$ , Sanov's theorem states that, for large values of  $n$ ,  $\mathbb{P}(\hat{\mu}_n \in B) \sim \exp[-n \inf_{\nu \in B} H(\nu|\mu)]$ , where  $H(\nu|\mu)$  denotes the Kullback-Leibler divergence between the two probability measures  $\mu$  and  $\nu$ . In particular, if  $B$  represents the fact that users do not clump where they should according to  $\mu$ ,  $H(\nu|\mu)$  will be large for  $\nu \in B$ , and the probability to make such an observation will be extremely small when  $n$  is large.

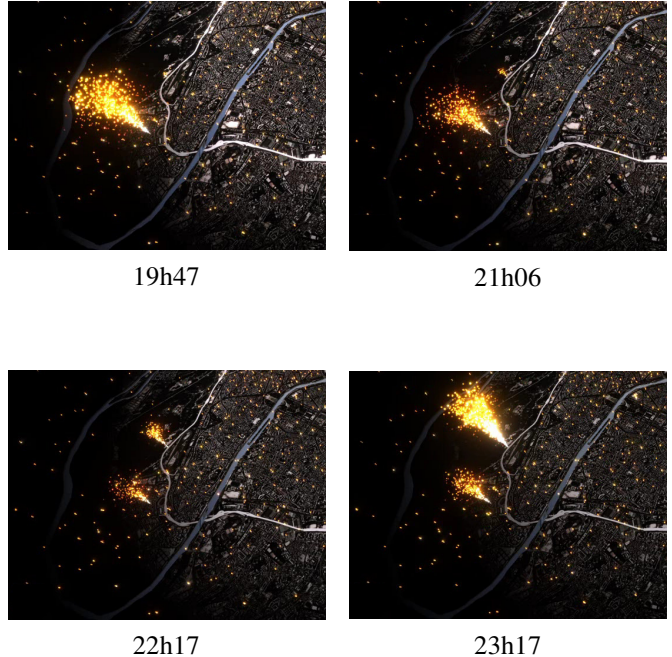


Fig. 2. Communication activity in Paris on June 21st, 2008, between 07:47 pm and 11:17 pm. One can observe scattering from the Parc des Princes (south) to Auteuil (north).

a natural attempt could be to work with permanent processes ([12]) that are known to exhibit frequent clumping. However, permanent processes are nothing but particular Cox processes (see [12], Remark 3.6.3.). This means that we would have to choose an underlying intensity measure which varies over time to cope with the dynamics of clumping zones, which is as difficult as building directly a convenient model.

As for determinantal processes, very little has been written yet about time-dependent versions. One way consists in starting from a gaussian analytic function, for instance  $f(z) = \sum_{n \geq 0} a_n z^n$ , where the  $a_n$ 's are i.i.d. Gaussian. Such a series converges almost surely on the unit disk, and its zeros form a determinantal process. It is then easy to make the  $a_n$ 's time-dependent, by using Ornstein - Uhlenbeck processes for instance (see [14]). The spatial processes we obtain are in fact almost exclusively distributed on the edge of the unit disk<sup>2</sup>. Such a phenomenon is too extreme for our purpose. For instance, it would be better adapted to model a wave of panick.

Another class of mobility models that could be considered is the class of spatial birth and death processes (see [4]) first studied by Preston in 1975 (see [13]). Such a process can simulate characteristic point patterns if one chooses its

<sup>2</sup>In fact, this is a direct consequence of the Peres & Virág theorem (see [12], Theorem 4.1.1.), which states that the  $n$ -th joint intensities of the zeros is given by:

$$\rho_n(z_1, \dots, z_n) = \frac{1}{\pi^n} \det \left[ \frac{1}{(1 - z_i \bar{z}_j)^2} \right].$$

In particular,  $\rho_1(z) = 1/\pi(1 - |z|^2)$ .

stationary distribution in a proper manner. Furthermore, in contrast to the random waypoint model, it provides a real time dynamics. However, this model is known to lead to Gibbs point processes which are usually not analytically tractable.

This state of the art leads us to the conclusion that there is a need for a tractable model allowing one to exhibit the type of hot spot dynamics observed in communication traces.

### III. THE SHEEP MODEL

The aim of this section is to model user clumping in some predefined zones. The basic idea is that users tend to move towards places where there are already more people, creating a self-enhanced clumping.

#### A. General Overview

Let us divide the city into  $K$  zones, that are exchanging mobiles one with another. In each zone, a mobile waits during an exponential time with parameter  $\mu$  before deciding whether to change its zone or not.

In a first step, let us suppose that the mobiles are totally influenced by their environment: they behave as sheep. They decide to move towards one zone or another with a probability proportional to the current repartition of the population (see Fig. 3) of the zone. In other words, if we denote by  $N$  the total number of mobiles, by  $n_1$  the number of mobiles in the first zone, by  $n_2$  the number of mobiles in the second zone, etc., a mobile from the  $k$ -th zone decides to move to the  $k'$ -th zone with probability  $n_{k'}/N$ . Conversely, he decides to stay in the  $k$ -th zone with probability:

$$\frac{n_k}{N} = 1 - \sum_{k' \neq k} \frac{n_{k'}}{N}.$$

Obviously, there are  $K$  absorbing states, which correspond to the situation where all mobiles are in the same zone (these are the  $K$  stable hot spots mentioned in the introduction).

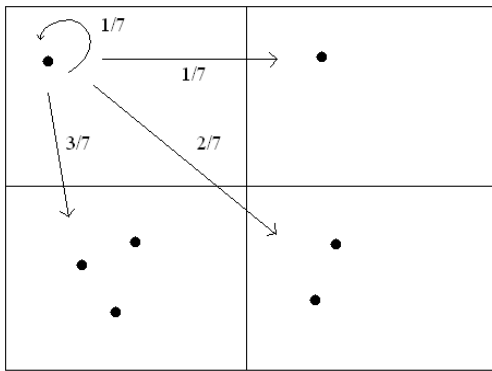


Fig. 3. The sheep model. Here,  $N = 7$  and  $K = 4$ . The mobile in the upper-left corner has 4 choices. The four decision probabilities are proportional to the numbers of mobiles.

If we denote a state by  $\mathbf{n} = (n_1, \dots, n_K)$ , then an absorbing state is given by:

$$N\mathbf{e}_k = \underbrace{(0, \dots, N, \dots, 0)}_{k\text{-th position}}, \quad 1 \leq k \leq K.$$

We first study the 2-zone case for the sake of progressive exposition.

#### B. The 2-zone Problem

In this section we take  $K = 2$ . Let us denote by  $\mathbf{n}(t) = (m(t), N - m(t))$  the state of the system at time  $t$ . Each mobile stays in his zone during a random exponential time with parameter  $\mu$ , before deciding whether to move or not. All waiting times are supposed to be independent. In other words, the system features a superposition of  $N$  exponential clocks ( $m$  clocks in the left zone and  $N - m$  in the right zone).

In the left zone for instance, the superposition of all waiting times should be an exponential clock with parameter  $\mu m$ , but there is not always a transition, since mobiles have the choice between moving and staying. Since the probability of moving is  $(N - m)/N$ , the actual resulting clock is a thinning of the  $\mu m$ -clock with probability  $(N - m)/N$ . As a conclusion, it is a  $\mu m(N - m)/N$ -clock. The same proof could be derived for the right zone.

Thus, the whole process is Markov, with transitions described in Fig. 4.

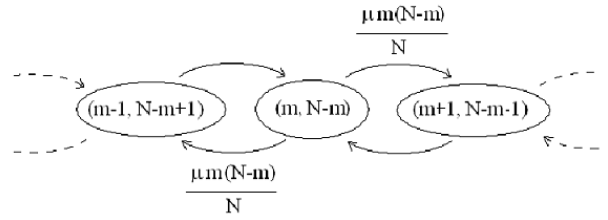


Fig. 4. Diagram of the markovian transitions.

Let us denote by  $Q = (q_{m,m'})$  the infinitesimal generator.

1) *Mean Absorption Time*: We are interested in characterizing mobility between different zones. For instance, we would like to give an estimate of  $\tau$ , the first time to reach an absorbing state, starting from the current state of the network  $\mathbf{n}$ :  $\tau = \inf\{t/\mathbf{n}(t) = (0, N) \text{ or } (N, 0)\}$ . The average value of  $\tau$  is given by  $h_{\mathbf{n}} = \mathbb{E}[\tau|\mathbf{n}(0) = \mathbf{n}]$ . We know (see [11], Theorem 3.3.3. about hitting times) that the vector  $\mathbf{h}$  is the minimum solution of the equation:

$$\begin{cases} h_{(0,N)} = h_{(N,0)} = 0 \\ \forall m \notin \{0, N\}, \quad \sum_{0 < i < N} q_{mi} h_{(i, N-i)} = -1. \end{cases} \quad (1)$$

We prove in appendix A that:

$$h_{m, N-m} = \frac{1}{\mu} \left( \sum_{0 \leq i < m} \frac{N-m}{N-i} + \sum_{0 \leq i < N-m} \frac{m}{N-i} \right). \quad (2)$$

2) *Asymptotic Estimate*: Typically,  $N$  corresponds to the number of users in a city district, thus it is quite large. In order to speed up computations, it can be useful to use an asymptotic estimate of  $h_{\mathbf{n}}$ . Let  $x = n_1/N$ , for  $1 \leq n_1 \leq N - 1$ . We want an estimate that can be used for large values of  $N$ , and for all values of  $x$ . This means that the quality of this estimate should only depend on  $N$ . Since  $N$  is large, we consider  $x$  as a

continuous variable taking its values in  $]0, 1[$ . The computation in appendix A gives:

$$h_{\mathbf{n}} \approx -\frac{N}{\mu} \left( (1-x) \log(1-x) + x \log(x) \right) \quad (3)$$

and  $r_{\mathbf{n}} \leq 2/\mu$ , where  $r_{\mathbf{n}}$  is the absolute error of the approximate. Note that (3) is always an equivalent in the mathematical sense, even when  $h_{\mathbf{n}}$  is minimal, that is, when  $n = 1/N$  or  $(n-1)/N$ . Indeed, in that case the approximate is equal to:

$$-\frac{N}{\mu} \left( \frac{1}{N} \log\left(\frac{1}{N}\right) + \frac{N-1}{N} \log\left(\frac{N-1}{N}\right) \right) \sim \frac{\log(N)}{\mu},$$

whereas  $r_{\mathbf{n}}$  remains bounded. Fig. 5 exemplifies the approximation curve.

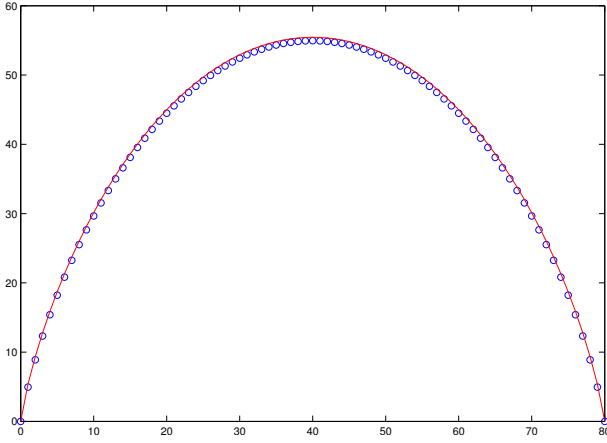


Fig. 5. Asymptotic estimate of  $h_{\mathbf{n}}$  [in min].  $N = 80$  and  $\mu = 1 \text{ min}^{-1}$ . The circles represent  $h_{\mathbf{n}}$ , and the continuous line represents the estimate.

3) *Absorption Probability*: So far we have obtained a closed form for the mean absorption time, as well as an asymptotic estimate. Now we would like to determine the probability to reach state  $(0, N)$  rather than state  $(N, 0)$ , starting from  $\mathbf{n} = (m, N-m)$ . We shall denote this probability by  $p_{\mathbf{n}}$ . Using  $Q$ , we obtain the transitions for the imbedded Markov chain of the process (see Fig. 6).

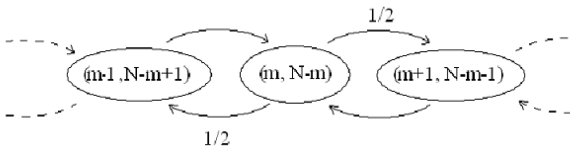


Fig. 6. Imbedded Markov chain of the process.

Conditioning with respect to  $T_1$ , the first transition time, and using Markov property, we get

$$p_{\mathbf{n}} = \frac{1}{2} p_{\mathbf{n}+(-1,1)} + \frac{1}{2} p_{\mathbf{n}+(1,-1)}.$$

Using the fact that  $p_{(0,N)} = 1$  and  $p_{(N,0)} = 0$ , we conclude immediately that:

$$p_{\mathbf{n}} = n_2/N. \quad (4)$$

### C. More Clumping Zones

We now consider the same problem with  $K$  zones.

1) *Description of the system*: Now a state is a  $K$ -uple  $\mathbf{n} = (n_1, n_2, \dots, n_K)$ , with  $\sum n_k = N$ . Hence the space of possible states is a  $(K-1)$ -dimensional variety. A transition from the  $k$ -th zone to the  $k'$ -th zone is given by:

$$\mathbf{n} \mapsto \mathbf{n}' = \mathbf{n} + \mathbf{e}_{k'} - \mathbf{e}_k \quad (5)$$

(we merely introduced a migration operator in Kelly's sense, see [8], 2.3.). In the most general case, there are  $K(K-1)$  possible transitions (corresponding to existing couples  $(k, k')$ ). The transition rate corresponding to the mobility  $k \rightarrow k'$  is given by:

$$q_{\mathbf{n}\mathbf{n}'} = \frac{\mu n_k n_{k'}}{N}.$$

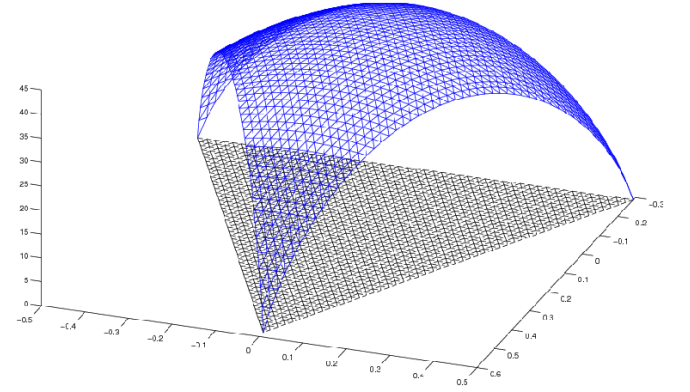


Fig. 7. Mean absorption time [in min] for  $N = 50$ ,  $K = 3$  and  $\mu = 1 \text{ min}^{-1}$ . The z-axis represents the time, whereas the initial state is represented on the horizontal plane, in barycentric coordinates with respect to each of the three absorbing states.

2) *Absorption Time*: Generalizing equation (2) leads to:

$$h_{\mathbf{n}} = \frac{1}{\mu} \sum_{1 \leq k \leq K} \sum_{0 \leq i < n_k} \frac{N - n_k}{N - i} \quad (6)$$

(for a proof, see appendix B).

On fig. 7, we propose a 2-dimensional surface representing the mean absorption time for  $K = 3$ .

Then, defining  $x_k = n_k/N$ , it is immediate to generalize equation (3) to find:

$$h(\mathbf{x}) \approx -\frac{N}{\mu} \sum_k (1 - x_k) \log(1 - x_k) \quad (7)$$

and  $|r(\mathbf{x})| \leq K/\mu$ .

3) *Absorption Probability*: In the same way as in section III-B3, we can compute the probability to reach an absorbing state rather than another one, for example say  $N\mathbf{e}_K$ . We use recursive formulas between a state and its neighbours. For example, in the most general case, we have:

$$p_{\mathbf{n}} = \frac{\sum_{(k,k') \neq} n_k n_{k'} p_{\mathbf{n}+\mathbf{e}_{k'}-\mathbf{e}_k}}{\sum_{(k,k') \neq} n_k n_{k'}}.$$



Using then the fact that  $\forall k \neq K, p_{Ne_k} = 0$ , we obtain all the  $p_n$ 's recursively. In fact, it is quite easy to generalize equation (4) to obtain  $p_n(K) = n_K/N = x_K$ .

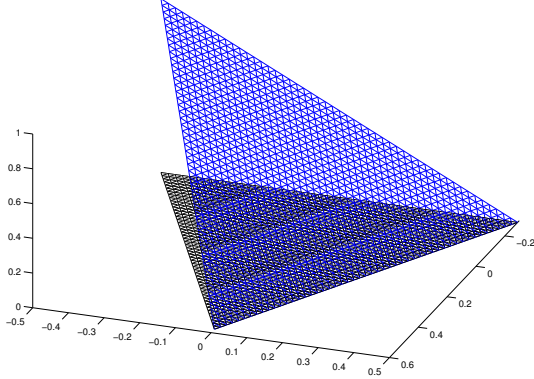


Fig. 8. Reaching probability for  $N = 50$ ,  $K = 3$  and  $\mu = 1 \text{ min}^{-1}$ .

#### IV. MAVERICKS AND HOT SPOTS DYNAMICS

So far, we have modeled the phenomenon of clumping in one given zone, which is only an aspect of what can be observed in dense urban networks during exceptional events. The sheep model has a major drawback, which prevents it to fit reality: as soon as it has reached an absorbing state, it does not evolve any more. This does not account for scattering, where a hot spot suddenly disaggregates before reforming in other zones.

This section aims at modeling scattering. We introduce some users, the mavericks, who are not influenced by the others. In the model, mavericks coexist with sheep. During a transition, mavericks choose uniformly their target zone, including their current zone, independently from other mobiles (which is a simple instance of independent random waypoint motion). In this new model, states of the form  $Ne_k$  are not absorbing anymore.

We introduce maverickness as follows<sup>3</sup>: at each transition, an individual decides to behave as a maverick with probability  $\alpha$  ( $\alpha$  is the maverickness rate), and as a sheep with probability  $1 - \alpha$ . The new rate of the  $(k \rightarrow k')$  transition is hence given by:

$$q_{nn'} = \mu n_k \left( \alpha \cdot \frac{1}{K} + (1 - \alpha) \frac{n_{k'}}{N} \right).$$

##### A. Mean Reaching Time

We call reaching time the time to reach any extremal state, namely a state of the form  $Ne_k$  from some  $k$ . On Fig. 9, we have plotted the evolution of the mean reaching time with respect to  $\alpha$ .

As  $\alpha$  grows, one can observe several phenomena:

<sup>3</sup>There are several possible coexistence models. A first model would consist in choosing a fixed population of sheep and mavericks at the beginning of the evolution; the main objection against this model is that it is quite constrained: suppose there are only 20 individuals; if we want to reach a "maverickness" rate of 0.01% for example, this is impossible.

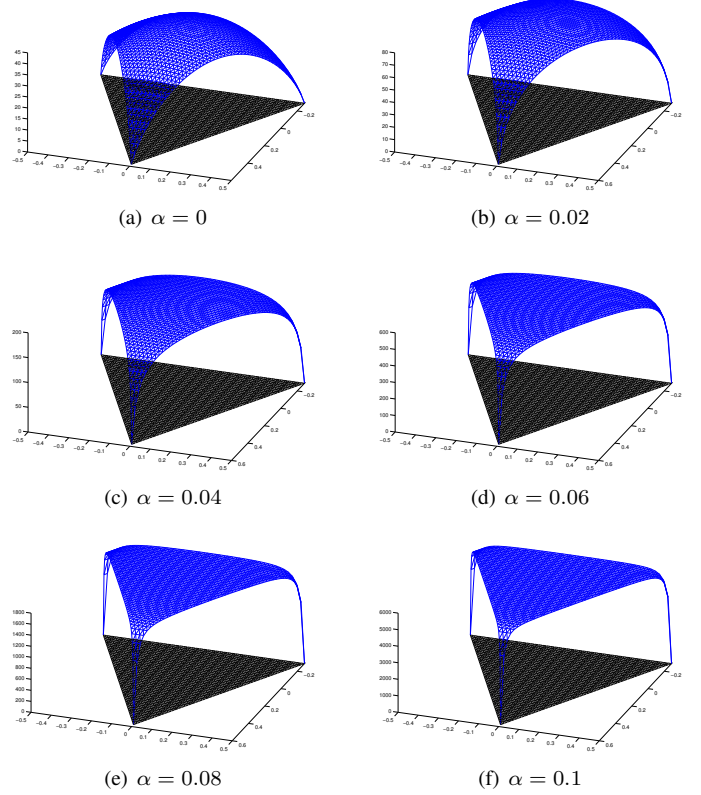


Fig. 9. Mean reaching time [in min] for  $N = 50$ ,  $K = 3$  and  $\mu = 1 \text{ min}^{-1}$ . The maverickness rate  $\alpha$  grows from 0 to 0.1. Note that the scale of the z-axis varies.

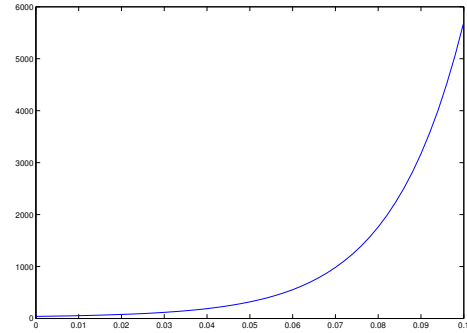


Fig. 10. Maximal mean reaching time [in min] with respect to  $\alpha$ .  $N = 50$ ,  $K = 3$  and  $\mu = 1 \text{ min}^{-1}$ .

- The mean reaching time grows exponentially with  $\alpha$ . On Fig. 10, we have plotted the maximal mean reaching time with respect to  $\alpha$  (this corresponds to the case where all zones initially contain each approximately  $N/K$  individuals).
- On Fig. 9, as  $\alpha$  increases (in other words, as the system becomes saturated by mavericks), the surface becomes very quickly plateau-shaped (this is already the case for  $\alpha = 0.08$ ). Apart from the regions close to extremal states, all reaching times are more or less the same, which means that reaching times become insensitive to initial

state. Even if the system starts from a state close to an extremal point, it takes a large excursion and a long time to reach one of the extreme states. It is interesting to observe that such a small proportion of mavericks can induce such a perturbation (for further details, see section IV-B). This is not yet a proof but just an indication that there is scattering (we propose a systematic study of scattering in section IV-C).

### B. Phase Transition

To characterize the phase transition observed in Fig. 9, we calculate the invariant measure  $\pi$  of the "sheep and maverick" Markov process, and observe whether there is a value of  $\alpha$  for which all configurations are equiprobable.  $\pi$  is given by the equation  $\pi Q = 0$ . Let  $\beta = \frac{\alpha}{K(1-\alpha)}$ , we have :

$$\pi(\mathbf{n}) = A \prod_{k=1}^K \frac{1}{n_k!} \prod_{i=0}^{n_k-1} (N\beta + i), \quad (8)$$

where  $A$  is a normalizing constant. This result is proved in appendix C. On Fig. 11, we plot the invariant measure with respect to  $\alpha$  for  $K = 2$  and  $N = 50$ . As already indicated by Fig. 9, it appears that when  $\alpha$  grows, the influence of mavericks is more and more noticeable. For higher values of  $\alpha$ , the probability is more concentrated on uniformly-spread states.

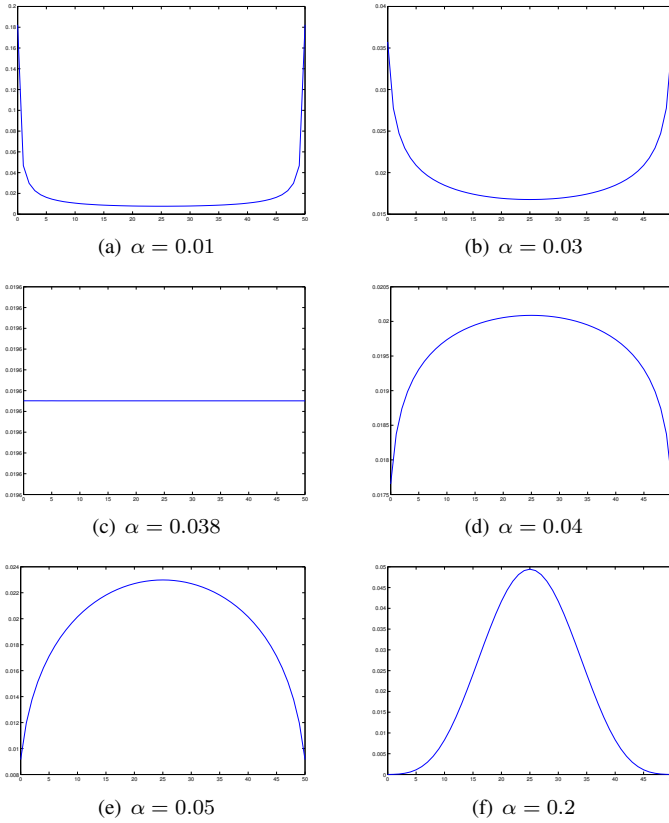


Fig. 11. Invariant measure for  $N = 50$ ,  $K = 2$  and  $\mu = 1 \text{ min}^{-1}$ . When  $\alpha$  exceeds 0.038, the shape of the curve reverses.

There is a critical value  $\alpha_c$  where the curve's shape reverses. What does the transition look like? Is it possible to obtain a completely flat curve (or surface in the general case) ?

Expression (8) shows that it is possible. It is sufficient to impose  $N\beta - 1 = 0$  to obtain a uniform distribution. This corresponds to the critical value :

$$\alpha_c = \frac{K}{N + K}.$$

- for  $\alpha < \alpha_c$ , the influence of the sheep remains noticeable. This is the most interesting case, because the system still exhibits clumping.
- for  $\alpha = \alpha_c$ , all states are equivalent<sup>4</sup>.
- for  $\alpha > \alpha_c$ , the sheep loose very fast their influence.

As traces reveal the existence of clumping, we shall assume until the end of the paper that  $\alpha < \alpha_c$ .

### C. Macro-states and Hot Spot Dynamics

The system is fully characterized by four parameters, which are  $N$ ,  $K$ ,  $\mu$  and  $\alpha$ . The three first parameters can be directly measured, but  $\alpha$  cannot. And yet, it is of major importance. In a sense, it represents the propensity of users to clump. How can we make it best fit reality ? We would like to find a typical characteristic of the system, easy to measure and strongly related to  $\alpha$ . This is what this subsection aims at.

Let us study further the case  $\alpha < \alpha_c$ . If we simulate such a system, we observe clumping as expected, but we also observe scattering. We would like to evaluate how often a scattering happens. However, extremal states of the form  $N\mathbf{e}_k$  are extremely rare, so that it would be difficult to base our analysis on them. We will rather consider "macro-states", which are the union of several states close to one of the extremal states.

For example, let us denote by  $A_k$  the macro-state in which the  $k$ -th zone contains a significant proportion of the population:  $A_k = \{\mathbf{n} : n_k > N_{\text{sup}}\}$ , where  $N_{\text{sup}}$  is a lower bound for a zone to be considered as dense. Conversely, we define  $B_k = \{\mathbf{n} : n_k < N_{\text{inf}}\}$ , where  $N_{\text{inf}}$  is a lower bound. Then a scattering is simply a transition from  $A_k$  to  $B_k$ . More precisely, for  $k$  fixed, we define the corresponding scattering time  $S_k$  as the duration between the first instant when we reach  $A_k$  and the first instant when we reach  $B_k$  after that. Another option would have consisted in using the last time we leave  $A_k$  before reaching  $B_k$  in place. But this is not a stopping time, which would make such a choice more difficult to handle using Markov chain theory. Hence the former definition.

The expectation of  $S_k$  only depends on initial state, regarding that we start in  $A_k$  or not. If we do not start in  $A_k$ , then almost surely  $n_k$  will be equal to  $N_{\text{sup}} + 1$  when we reach  $A$ .

<sup>4</sup>As  $N \rightarrow \infty$ , this assumption can be interpreted as follows: since  $\alpha_c \sim K/N$ , it means that, asymptotically, it is enough to have only one maverick per zone to saturate the system, which is quite surprising.



It is shown in appendix D that  $\mathbb{E}[S|\mathbf{n}^{(0)} = \mathbf{n}]$  is given by the following analytic expression :

$$\frac{N}{\mu(1-\alpha)} \sum_{l=N_{\text{inf}}}^{n_k \vee (N_{\text{sup}}+1)} \sum_{j=l}^N \frac{\prod_{l \leq i < j} \delta(i)}{j(N(K-1)\beta + N - j)}, \quad (9)$$

with

$$\delta(i) = \frac{(N-i)(N\beta+i)}{i(N(K-1)\beta+N-i)}.$$

The expression in (9) only depends on the initial state, and in particular on whether it belongs to  $A_k$  or not. If we are not in  $A_k$ , then (9) does not depend on  $\mathbf{n}$ . By symmetry w.r.t.  $k$ , we can denote it by  $\mathbb{E}[S]$ . Fig. 12 depicts the evolution of  $\mathbb{E}[S]$  in function of  $\alpha$  for several values of  $N$ .

Let  $\alpha_{\text{min}}$  be the value of  $\alpha$  such that  $\mathbb{E}[S]$  is minimum. For certain values of  $N$ ,  $\mathbb{E}[S]$  is strictly decreasing over  $[0, 1]$ , so that  $\alpha_{\text{min}} = 1$ , but for some other values of  $N$  (like  $N = 1000$ ),  $\mathbb{E}[S]$  starts increasing after  $\alpha_{\text{min}}$ .

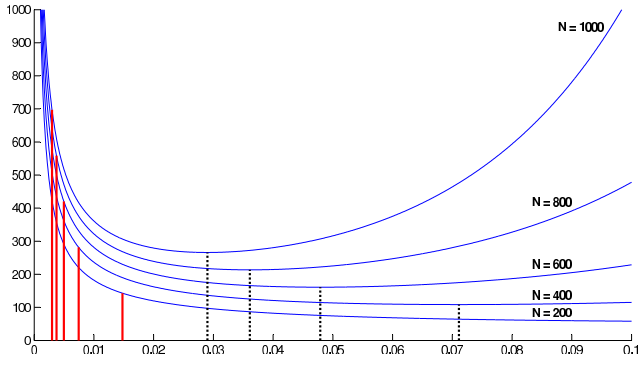


Fig. 12. Mean scattering time [in min] with respect to  $\alpha$ .  $N_{\text{sup}}/N = 80\%$ ,  $N_{\text{inf}}/N = 20\%$ ,  $K = 3$  and  $\mu = 1 \text{ min}^{-1}$ . For each curve, the vertical solid line corresponds to the critical value  $\alpha_c$ , while the dotted line corresponds to  $\alpha_{\text{min}}$  (whenever it is less than the right bound of the axis).

In the same way, we can compute the mean "filling time"  $F$  of a zone, which is the dual of the scattering time, and which we propose to define as the time for the population in a given zone to grow from  $N_{\text{inf}}$  to  $N_{\text{sup}}$ . We have :

$$\mathbb{E}[F] = \frac{N}{\mu(1-\alpha)} \sum_{l=N-N_{\text{sup}}}^{N-N_{\text{inf}}+1} \sum_{j=l}^N \frac{\prod_{l \leq i < j} \epsilon(i)}{j(N\beta + N - j)}, \quad (10)$$

with

$$\epsilon(i) = \frac{1}{\delta(N-i)} = \frac{(N-i)(N(K-1)\beta+i)}{i(N\beta+N-i)}.$$

Note that it is easy to find some recursive formulas for the computation of  $\mathbb{E}[S]$  and  $\mathbb{E}[F]$ , which allows to decrease the complexity of their evaluation down to a Big-O of  $N$ .

## V. CALIBRATION WITH TRACES

The aim of this section is to see whether the crowd movements observed in the traces can be described by our "sheep and maverick" mobility model.

We propose to use  $\mathbb{E}[S]$  or  $\mathbb{E}[F]$  for our calibration. These quantities can be measured. Let us now see that if  $\mathbb{E}[S]$  and

$N$  are known for instance, we can then deduce  $\alpha$  by solving an inverse problem. If the function  $\mathbb{E}[S] = f(\alpha)$  is injective, at least for  $\alpha < \alpha_c$ , this inverse problem can be solved without ambiguity. It appears that  $f(\alpha)$  is always convex, thus decreasing on  $]0, \alpha_{\text{min}}[$ , so that a necessary and sufficient condition would be  $\alpha_c \leq \alpha_{\text{min}}$ , provided we prove convexity. On Fig. 12 this is always the case, but this is not true in general; a counter-example is  $K = 2$ ,  $N = 100$ ,  $N_{\text{sup}}/N = 0.99\%$ . In practice,  $\alpha_{\text{min}}$  seems to be hard to compute analytically, so that we could not find a simple sufficient condition.

Our experiences show that such counter-examples appear to be extremely rare, especially when  $K > 2$ . It seems that most choices of  $N_{\text{inf}}$  and  $N_{\text{sup}}$  satisfy the injectivity of  $f(\alpha)$  on  $]0, \alpha_c[$ . However, if it turned out that it is impossible to choose  $N_{\text{inf}}$  and  $N_{\text{sup}}$  properly, we advise to complement the measure of  $\mathbb{E}[S]$  with a measure of  $\mathbb{E}[F]$  for instance.

In contrast, when measuring just  $\mathbb{E}[F]$ , one has to be more cautious, since the injectivity of  $g(\alpha) = \mathbb{E}[F]$  on  $]0, \alpha_c[$  seems less common.

Let us consider for instance the transfer that happens between *Parc des Princes* and *Auteuil* (see Fig. 2). Observing carefully our data, we conclude that there are three attracting zones :

- *Auteuil* itself
- the *Parc des Princes*, northern part
- the *Parc des Princes*, southern part

so that we take  $K = 3$ . On Fig. 13, one can see the numbers of users in each zone.

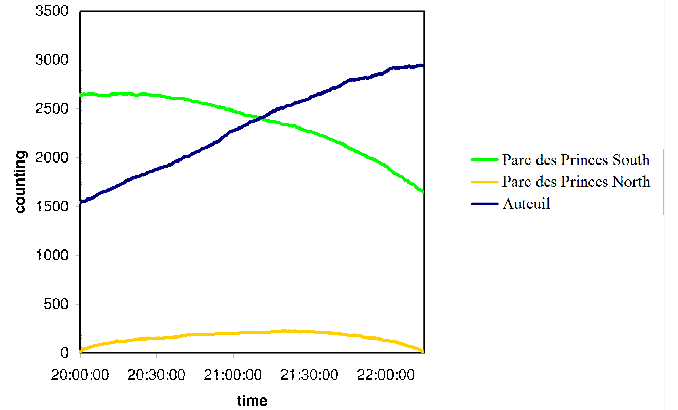


Fig. 13. Traces collected during the *Fête de la Musique* between the three attracting zones.

The total number of traces remains more or less constant, so that we choose an average value of  $N = 4830$ . This gives  $\alpha_c = 6.2 \times 10^{-4}$ . In our data, a SMS signalization lasts about one second: each time a SMS is sent or received, a new user is counted, but he is deleted after one second. Therefore, even if this is not true, but in order to remain consistent, we make as if users were staying one second in the zone before disappearing. Thus we put  $\mu = 1 \text{ sec}^{-1}$ . Choosing to calibrate first on *Auteuil* filling, we take  $N_{\text{sup}}/N = 55\%$  and  $N_{\text{inf}}/N = 45\%$ . After verifying that  $g(\alpha)$  is injective on

$]0, \alpha_c[$ , we find  $\alpha = 0.6 \times 10^{-4}$ , that is,  $\alpha \ll \alpha_c$ . This was to expect, since scattering is very clear in our case.

In order to verify this value, we also measure the *Parc des Princes South* scattering, which yields  $\alpha = 0.5 \times 10^{-4}$ .

Now, if our model is consistent, it has to predict other clumps or scattering over Paris. Let us consider the triangle *Châtelet-Saint Michel-Bastille* for instance, in the heart of Paris (see Fig. 14). These zones are attractive. Besides, we consider a fourth zone which is the complementary region inside the triangle. That one experiences scattering, whereas the three attractive zones experience filling at the same time.

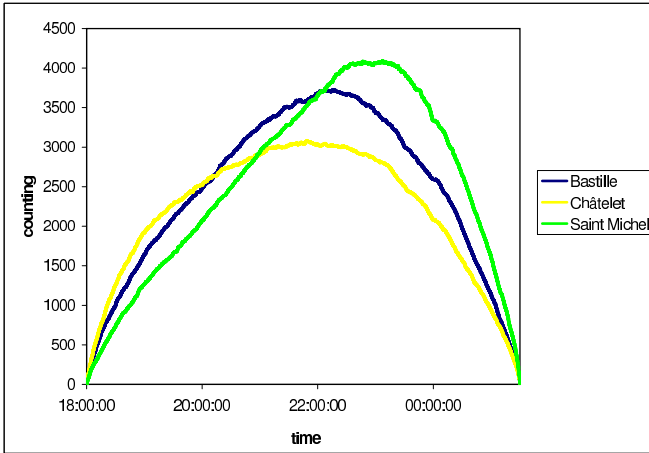


Fig. 14. Traces collected during the *Fête de la Musique* in the triangle *Châtelet-Saint Michel-Bastille*.

Theoretically, a filling of a zone from 5% to 28% for instance should last about 10800 seconds, according to the value of  $\alpha$  we found previously. Now, the filling of *Bastille* lasts about 11300 seconds, so does the filling of *Châtelet*, and the filling of *Saint Michel* lasts about 11000 seconds, which is very close to our prediction.

## VI. CONCLUSION AND PERSPECTIVES

In this paper, we proposed a new mobility model able to represent and quantify the hot spots found on traces, along with their random time dynamics. Our aim was not to build a generic mobility model, adapted to every kind of situation, but rather to focus on multi-event mass gatherings. We have proved our model to be analytically tractable, by derivating closed forms and asymptotic estimates. Using measurements from a GSM network, we also showed how to calibrate the model to fit reality. In addition, our model was proved to be able to predict the order of magnitude of hot spots in different time-space situations.

Apart from the main application, which consists in building an efficient mobility model for network simulators, we describe here some applications, that rely on the ability to predicting clumping or scattering with a good likelihood.

### A. Dynamic Spectrum Allocation and Software Defined Radio

Dynamic Spectrum Allocation (see [7]), consists in allocating variable bandwidth resources, according to different zones

and different densities of population. As for Software Defined Radio (see [10]), it consists in deploying base stations that are able to switch from a system (like GSM) to another one (like UMTS), depending on the demand.

Even if they have not been implemented in real networks yet, these two domains are today quite an active field of research. It is crucial that implemented algorithms can have an idea of the typical evolution durations of the system.

For instance, if we can predict if a zone that is full will remain so for a long time, or if an empty zone will soon fill, one can decide with much more accuracy whether allocating resources or not.

More generally, some optimization algorithms are known to depend massively on the tuning of a key parameter. For instance, [6] proposes an algorithm to optimize radio resources in a 802.11 network, based on Gibbs sampler. The "temperature parameter" of the sampler has to be tuned very carefully, so that the optimizer can react quicker than the typical time of evolution of the system itself. More concretely, in our case, supposing that we want to take users mobility into account in a 802.11 network, if we can anticipate a scattering, we are able to adapt the optimizer in a suitable manner.

Also in the case of a genetic algorithm, it can be interesting to adapt the cross-over and the mutation rates of the optimizer to react to a sudden evolution of the system. Indeed, the higher these rates, the stronger the diversity of the genetic population. In our case, a scattering is likely to modify drastically the system to optimize. Thus it is useful to broaden preventatively the variety of solutions that are explored.

### B. Data Caching in Dense Ad-hoc Networks

In a dense data ad-hoc network, it can be useful to add some fixed servers which keep the most frequently asked data in cache memories. Various algorithms already exist on this subject (see for instance [5][15]). The question is where and when should we cache memory ?

The idea is to select some strategic places where people usually clump together. These are our  $K$  zones. Suppose that, using section V, we have previously evaluated  $\alpha$  over the study area, possibly for a different number of users and zones. Then, measuring in real-time the actual number  $N$  of users and computing the corresponding scattering times, as soon as a zone exhibits clumping, one knows how long on average this will last. In other words, one knows if it is worth caching data at that place.

Conversely, if a zone is almost empty, one knows, using "filling times", if there is a risk that we will soon encounter a peak of population and have to rush to cache memory there.

## APPENDIX

## A. Resolution of Equation (1)

$\sum_{0 < i < N} q_{mi} h_{(i, N-i)} = -1$  can be rewritten as :

$$\begin{cases} 2h_{(1, N-1)} - h_{(2, N-2)} = \frac{N}{\mu(N-1)} \\ 2h_{(N-1, 1)} - h_{(N-2, 2)} = \frac{N}{\mu(N-1)} \\ 2h_{\mathbf{n}} - h_{\mathbf{n}+(-1, 1)} - h_{\mathbf{n}+(1, -1)} = \frac{N}{\mu(N-i)} \quad \text{otherwise.} \end{cases}$$

Thus we have to invert the  $(N-1) \times (N-1)$  matrix  $U = (u_{i,j})$  with  $u_{i,i} = 2$ ,  $u_{i,i-1} = u_{i-1,i} = -1$ ,  $u_{i,j} = 0$  otherwise. One can easily check that  $U^{-1}$  is given by  $V/N$ , where the elements of  $V$  are :

$$v_{ij} = \begin{cases} j(N-i) & \text{if } i \geq j \\ i(N-j) & \text{if } i \leq j. \end{cases}$$

Equation (2) follows immediately.

Let us now give an asymptotic estimate of this expression. We introduce  $x = n_1/N$ ,  $t = i/N$  and  $dt = 1/N$ . We have :

$$\mu h_{\mathbf{n}} = Nx \sum_{i=n_1+1}^N \frac{1}{N} \cdot \frac{1}{i} + N(1-x) \sum_{i=n_2+1}^N \frac{1}{N} \cdot \frac{1}{i}.$$

Both sums of the right handside can be interpreted as Riemann integrals on  $[x, 1]$  and  $[1-x, 1]$ . Let us work out the first one for example :

$$\int_{x+1/N}^{1+1/N} \frac{1}{t} dt \leq \sum_{i=n_1+1}^N \frac{1}{N} \cdot \frac{1}{i} \leq \int_x^1 \frac{1}{t} dt$$

A fortiori :

$$-\log(x) - \frac{1}{Nx} \leq \sum_{i=n_1+1}^N \frac{1}{N} \cdot \frac{1}{i} \leq -\log(x).$$

We would obtain the same inequality replacing  $x$  by  $1-x$ . Defining the remainder :

$$r_{\mathbf{n}} = \left| h_{\mathbf{n}} + \frac{N}{\mu} (x \log(x) + (1-x) \log(1-x)) \right|,$$

we have  $\mu r_{\mathbf{n}} \leq 2$ . Thus, our estimate is

$$-\frac{N}{\mu} (x \log(x) + (1-x) \log(1-x)),$$

and the error is uniformly bounded by  $2/\mu$ .

## B. Proof of equation (6)

We are now looking for a solution of equation (1) in  $K$  dimensions. First, let us start from equation (6), in order to prove it satisfies the equation.

First remark : is one of the  $n_k$ s is zero, then we can remove zone  $k$  without changing the sum in (6), since the subsum indexed over  $0 \leq i < n_k$  is empty. Furthermore, this will not change (1) either, since transitions towards empty zones are forbidden. Thus, without loss of generality we assume that all  $n_k$ s are nonzero.

If  $K = 1$ , we are in an absorbing state and equation (6) holds. Otherwise, following equation (5), let us introduce  $\epsilon_{\mathbf{nn}'} = h_{\mathbf{n}'} - h_{\mathbf{n}}$ . We have :

$$\begin{aligned} \forall \mathbf{n}, \quad \sum_{\mathbf{n}'} q_{\mathbf{nn}'} h_{\mathbf{n}'} &= h_{\mathbf{n}} \underbrace{\sum_{\mathbf{n}'} q_{\mathbf{nn}'} + \sum_{\mathbf{n}'} q_{\mathbf{nn}'} \epsilon_{\mathbf{nn}'}}_{=0} \\ &= \sum_{\mathbf{n}' \neq \mathbf{n}} q_{\mathbf{nn}'} \epsilon_{\mathbf{nn}'}, \end{aligned}$$

and :

$$\begin{aligned} \mu h_{\mathbf{n}'} &= \sum_{l \neq k, k'} \sum_{0 \leq i < n_l} \frac{N - n_l}{N - i} \\ &+ \sum_{0 \leq i < n_k - 1} \frac{N - n_k + 1}{N - i} + \sum_{0 \leq i < n_{k'} + 1} \frac{N - n_{k'} - 1}{N - i} \\ &= \sum_l \sum_{0 \leq i < n_l} \frac{N - n_l}{N - i} - \frac{N - n_k}{N - n_k + 1} + \frac{N - n_{k'}}{N - n_{k'}} \\ &\quad + \sum_{0 \leq i < n_k - 1} \frac{1}{N - i} - \sum_{0 \leq i < n_{k'} + 1} \frac{1}{N - i} \\ &= \mu h_{\mathbf{n}} + \frac{1}{N - n_k + 1} \\ &\quad + \sum_{0 \leq i < n_k - 1} \frac{1}{N - i} - \sum_{0 \leq i < n_{k'} + 1} \frac{1}{N - i}, \end{aligned}$$

whence :

$$\epsilon_{\mathbf{nn}'} = \frac{1}{\mu} \left( \sum_{0 \leq i < n_k} \frac{1}{N - i} - \sum_{0 \leq i \leq n_{k'}} \frac{1}{N - i} \right).$$

Note that the second sum is well-defined, because  $n_{k'} \neq 0$  and  $K > 1$ , so that  $n_{k'} < N$  and  $N - i$  cannot be zero. Now, we can prove that (6) verifies equation (1) :

$$\begin{aligned} \sum_{\mathbf{n}' \neq \mathbf{n}} q_{\mathbf{nn}'} \epsilon_{\mathbf{nn}'} &= \sum_{(k, k')} \frac{n_k n_{k'}}{N} \left( \sum_{0 \leq i < n_k} \frac{1}{N - i} - \sum_{0 \leq i \leq n_{k'}} \frac{1}{N - i} \right) \\ &= \underbrace{\sum_{(k, k')} \frac{n_k n_{k'}}{N} \left( \sum_{0 \leq i < n_k} \frac{1}{N - i} - \sum_{0 \leq i < n_{k'}} \frac{1}{N - i} \right)}_{0 \text{ by symmetry}} \\ &\quad - \sum_{(k, k')} \frac{n_k n_{k'}}{N} \cdot \frac{1}{N - n_{k'}} \\ &= - \sum_{k'} \frac{n_{k'}}{N(N - n_{k'})} \cdot (N - n_{k'}) = -1. \end{aligned}$$

We still have to prove that in our case, there is no other solution of (1). In other words, we have to prove that the only solution of :

$$\begin{cases} \forall \mathbf{n} \text{ absorbing, } h_{\mathbf{n}} = 0 \\ \forall \mathbf{n} \text{ non-absorbing, } \sum_{\mathbf{n}' \text{ non-absorbing}} q_{\mathbf{nn}'} h_{\mathbf{n}'} = 0 \end{cases} \quad (11)$$

is the zero-vector. As usual, we introduce  $q_{\mathbf{nn}} = -\lambda_{\mathbf{n}}$ , and  $q_{\mathbf{nn}'} = \lambda_{\mathbf{n}} p_{\mathbf{nn}'}$  if  $\mathbf{n} \neq \mathbf{n}'$  (where the  $p_{\mathbf{nn}'}$ s are the

transition probabilities). If there exists a transition  $\mathbf{n} \rightarrow \mathbf{n}'$  with probability  $> 0$ , then  $p_{\mathbf{nn}'} > 0$ , and  $p_{\mathbf{nn}'} = 0$  otherwise.

Since  $h_{\mathbf{n}'} = 0$  if  $\mathbf{n}'$  is absorbing, the second line of (11) can be rewritten as:

$$\sum_{\mathbf{n}'} q_{\mathbf{nn}'} h_{\mathbf{n}'} = 0 \Leftrightarrow \sum_{\substack{\mathbf{n}' \neq \mathbf{n} \\ \mathbf{n} \rightarrow \mathbf{n}' \text{ existe}}} p_{\mathbf{nn}'} h_{\mathbf{n}'} = h_{\mathbf{n}}.$$

Consider  $\mathbf{n}$  such that  $h_{\mathbf{n}}$  is maximal. Then by convexity, since all coefficients  $p_{\mathbf{nn}'}$  are nonnegative, we see that necessarily all  $h_{\mathbf{n}'}$ s such that  $\mathbf{n} \rightarrow \mathbf{n}'$  exists with probability  $> 0$  are equal to  $h_{\mathbf{n}}$ . Step by step, since we can reach at least one absorbing state starting from  $\mathbf{n}$ , this proves  $h_{\mathbf{n}} = 0$ .

Finally, changing  $\mathbf{h}$  in  $-\mathbf{h}$  and applying the same result, we see that the minimum of the  $h_{\mathbf{n}}$ s is zero too, which achieves the proof.

### C. Steady-state Computation

In this section, we will explain in several steps how to prove (8). First let us take  $K = 2$  and keep notations of section III-B.

We are facing with a birth and death process, which is time-reversible, so that we can directly try to satisfy local balance equations:

$$\forall m \in \llbracket 0, N-1 \rrbracket, \pi(m) q_{m, m+1} = \pi(m+1) q_{m+1, m}.$$

We deduce at once:

$$\pi(m) = \pi(0) \prod_{i=1}^m \frac{(N-i+1)(N\beta+i-1)}{i(N\beta+N-i)}.$$

In order to generalize this formula, we would like to rewrite it in a symmetric manner w.r.t.  $n_1$  and  $n_2$ . In fact, we have:

$$\pi(m) = \pi(0) \frac{\prod_{i=1}^{n_1} \frac{N\beta+i-1}{i} \prod_{i=1}^{n_2} \frac{N\beta+i-1}{i}}{\prod_{i=1}^N \frac{N\beta+i-1}{i}}.$$

This is equation (8) in case  $K = 2$ . Now we can reasonably suppose that (8) still holds in the general case. In fact, this is quite immediate. Indeed, thanks to the product form of (8), we know by the same calculation that in the general case, the partial balance equations are still satisfied. This completes the proof.

### D. Proof of equation (9)

As previously, let us begin with the case  $K = 2$ . We are in state  $(m, N-m)$  and we are looking for the hitting time  $h_m$  of  $\{\mathbf{n}/n_1 < N_{\text{inf}}\}$ .

We want to obtain an analytical formula of  $h_m$  (let us say for  $N_{\text{inf}} \leq m \leq N$ , since otherwise  $h_m = 0$ ). If we introduce  $u_m = h_m - h_{m-1}$ , we have the following recursion ( $N_{\text{inf}} + 1 \leq m \leq N-1$ ):

$$u_{m+1} = \frac{m(N\beta+N-m)}{(N-m)(N\beta+m)} u_m - \frac{N}{\mu(1-\alpha)(N-m)(N\beta+m)}.$$

Then the idea is to introduce:

$$\delta(i) = \frac{(N-i)(N\beta+i)}{i(N\beta+N-i)} \text{ and } v_m = \frac{u_m}{\prod_{i=m}^{N-1} \delta(i)}.$$

We obtain the following value of  $v_m - v_{m+1}$ :

$$\frac{N}{\mu(1-\alpha)m(N\beta+N-m) \prod_{i=m}^{N-1} \delta(i)},$$

so that, for  $N_{\text{inf}} + 1 \leq m \leq N$ ,  $v_m$  is equal to:

$$v_m + \frac{N}{\mu(1-\alpha)} \sum_{j=m}^{N-1} \frac{1}{j(N\beta+N-j) \prod_{i=j}^{N-1} \delta(i)}.$$

Since  $q_{N,N} h_N + q_{N,N-1} h_{N-1} = -1$ , we also have:  $u_N = -1/q_{N,N} = 2/\mu\alpha N = v_N$ . Finally:

$$v_m = \frac{N}{\mu(1-\alpha)} \sum_{j=m}^N \frac{1}{j(N\beta+N-j) \prod_{i=j}^{N-1} \delta(i)},$$

and:

$$u_m = \frac{N}{\mu(1-\alpha)} \sum_{j=m}^N \frac{\prod_{m \leq i < j} \delta(i)}{j(N\beta+N-j)}.$$

To determine the  $h_m$ s, we still miss an equation. This is the last one before reaching  $\{\mathbf{n}/n_1 < N_{\text{inf}}\}$ :

$$q_{N_{\text{inf}}, N_{\text{inf}}} h_{N_{\text{inf}}} + q_{N_{\text{inf}}, N_{\text{inf}}+1} h_{N_{\text{inf}}+1} = -1.$$

Together with  $h_{N_{\text{inf}}+1} - h_{N_{\text{inf}}} = u_{N_{\text{inf}}+1}$ , this leads successively to:

$$h_{N_{\text{inf}}} = \frac{N}{\mu(1-\alpha)} \sum_{j=N_{\text{inf}}}^N \frac{\prod_{N_{\text{inf}} \leq i < j} \delta(i)}{j(N\beta+N-j)},$$

and, more generally, to:

$$h_m = \frac{N}{\mu(1-\alpha)} \sum_{l=N_{\text{inf}}}^m \sum_{j=l}^N \frac{\prod_{l \leq i < j} \delta(i)}{j(N\beta+N-j)}.$$

In particular, for  $m = N_{\text{sup}} + 1$ , we obtain the claimed formula. Now suppose we are in the general case. It is perfectly correct to consider the union of the  $K-1$  last zones as one big zone. A sheep will stay in this "macro-zone" with probability  $\frac{n_2 + \dots + n_k}{N}$ , and a maverick will stay with probability  $\frac{K-1}{K}$ . This last probability is the only adaptation we have to make in the previous calculation. Introducing  $\beta' = (K-1)\beta$ , we obtain the following new recursion:

$$u_{m+1} = \frac{m(N\beta' + N - m)}{(N - m)(N\beta + m)} u_m - \frac{N}{\mu(1-\alpha)(N - m)(N\beta + m)},$$

The rest of the calculation is exactly the same. In particular, for a fixed  $n$ , the result is the same for each  $\mathbf{n}$  such that  $n_1 = n$ .

## REFERENCES

- [1] Camp T., Boleng J., and Davies V., *A Survey of Mobility Models for Ad Hoc Network Research*, WCMC, Special Issue on Mobile Ad Hoc Networking - Research, Trends and Applications, vol. 2, no. 5, pp. 483502, 2002.
- [2] Cover T.M., Thomas J.A., *Elements of Information Theory, Second Edition*, Wiley, 2006.
- [3] Dembo A. and Zeitouni O., *Large Deviations Techniques and Applications*, Jones and Bartlett, 1993.
- [4] Garcia N.L. and Kurtz T.G., *Spatial birth and death processes as solutions of stochastic equations*, *Alea* 1, 281-303 (2006).
- [5] Hara T., *Replica Allocation Methods in Ad Hoc Networks with Data Update*, Department of Multimedia Engineering, Graduate School of Information Science and Technology, Osaka University, 2003.
- [6] Kauffmann B., Baccelli F., Chaintreau A., Mhatre V., Papagiannaki K. and Diot C., *Measurement-Based Self Organization of Interfering 802.11 Wireless Access Networks*, Proceedings IEEE Infocom'07, Anchorage, Pages :1451 - 1459, May 2007.
- [7] Leaves P. et al., *Dynamic Spectrum Allocation in Composite Reconfigurable Wireless Networks*, IEEE Communications Magazine, May 2004.
- [8] Kelly F. P., *Reversibility and stochastic networks*, J. Wiley (1979).
- [9] Le Boudec J.Y. and Vojnović M., *Perfect Simulation and Stationarity of a Class of Mobility Models*, Proceedings IEEE Infocom'05.
- [10] Mitola J., *Cognitive Radio: an Integrated Agent Architecture for Software Defined Radio*, PhD Thesis, Royal Institute of Technology (KTH), 2000.
- [11] Norris J.R., *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics, 1997.
- [12] Ben Hough J., Krishnapur M., Peres Y., Virg B., *Zeros of Gaussian analytic functions and determinantal point processes*, University Lecture Series, vol. 51. American Mathematical Society, 2009.
- [13] Preston C.J., *Spatial birth-and-death processes*, *Bull. Int. Statist. Inst.* 46 (2) 371391, 1975.
- [14] Uhlenbeck G. E., Ornstein L. S., *On the theory of the Brownian Motion*, *Physical Review*, 36 :823-841, 1930.
- [15] Yin L. and Cao G., *Supporting Cooperative Caching in Ad Hoc Networks*, IEEE Transactions on Mobile Computing, vol. 5, no. 1, pp. 77-89, Jan. 2006.