Modeling Human Dynamics of Face-to-Face Interaction Networks

Michele Starnini,[1] Andrea Baronchelli,[2] and Romualdo Pastor-Satorras[1]

[1]Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Campus Nord B4, 08034 Barcelona, Spain
[2]Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston Massachusetts 02115, USA

(Received 16 January 2013; published 15 April 2013)

Face-to-face interaction networks describe social interactions in human gatherings, and are the substrate for processes such as epidemic spreading and gossip propagation. The bursty nature of human behavior characterizes many aspects of empirical data, such as the distribution of conversation lengths, of conversations per person, or of interconversation times. Despite several recent attempts, a general theoretical understanding of the global picture emerging from data is still lacking. Here we present a simple model that reproduces quantitatively most of the relevant features of empirical face-to-face interaction networks. The model describes agents that perform a random walk in a two-dimensional space and are characterized by an attractiveness whose effect is to slow down the motion of people around them. The proposed framework sheds light on the dynamics of human interactions and can improve the modeling of dynamical processes taking place on the ensuing dynamical social networks.

DOI: 10.1103/PhysRevLett.110.168701 PACS numbers: 89.75.Hc, 05.40.Fb, 89.65.Ef

Uncovering the patterns of human mobility [1] and social interactions [2] is pivotal to decipher the dynamics and evolution of social networks [3], with wide practical applications ranging from traffic forecasting to epidemic containment. Recent technological advances have made possible the real-time tracking of social interactions in groups of individuals, at several temporal and spatial scales. This effort has produced large amounts of empirical data on human dynamics, concerning letter exchanges [4], email exchanges [5], mobile phone communications [1], or spatial mobility [6], among others.

Especially noteworthy is the data on face-to-face human interactions recorded by the SocioPatterns collaboration [7] in closed gatherings of individuals such as schools, museums, or conferences. SocioPatterns deployments measure the proximity patterns of individuals with a space-time resolution of $\sim$1 m and $\sim$20 sec by using wearable active radio-frequency identification devices. The data generated by the SocioPattern infrastructure show that human activity follows a bursty dynamics, characterized by heavy-tailed distributions for the duration of contacts between individuals or groups of individuals and for the time intervals between successive contacts [8,9].

The bursty dynamics of human interactions has a deep impact on the properties of the temporally evolving networks defined by the patterns of pairwise interactions [10], as well as on the behavior of dynamical processes developing on top of those dynamical networks [9,11–16]. A better understanding of these issues calls for new models, capable to reproduce the bursty character of social interactions and trace back their ultimate origin, beyond considering their temporal evolution [17]. Previous modeling efforts mostly tried to connect the observed burstiness to some kind of cognitive mechanisms ruling human mobility patterns, such as a reinforcement dynamics [18], cyclic closure [19] or preferential return rules [20], or by focusing on the relation between activity propensity and actual interactions [17].

In this Letter, we present a simple model of mobile agents that captures the most distinctive features of the empirical data on face-to-face interactions recorded by the SocioPatterns collaboration. Avoiding any *a priori* hypothesis on human mobility and dynamics, we assume that agents perform a random walk in space [21] and that interactions among agents are determined by spatial proximity [22]. The key ingredients of the model are the following: We consider that individuals have different degrees of social appeal or *attractiveness*, due to their social status or the role they play in social gatherings, as observed in many social [23], economic [24], and natural [25] communities. The effect of this social heterogeneity is that interactions, as well as the random walk motion of the agents, are biased by the attractiveness of the peers they met over time. Additionally, we assume, according to experimental data, that not all the agents are simultaneously present in the system, but can jump in and out of an active state in which they can move and establish interactions. We will see that these simple assumptions allow the model to reproduce many of the properties of face-to-face interaction networks.

The model is defined as follows (see Fig. 1): N agents are placed in a square box of linear size L with periodic boundary conditions, corresponding to a density $\rho = N/L^2$. Each individual i is characterized by her attractiveness or social appeal, a_i which represents her power to raise interest in the others. The attractiveness a_i of the agents is a (quenched) variable randomly chosen from a prefixed distribution $\eta(a)$, and bounded in the interval $a_i \in [0, 1)$. Agents perform a random walk biased by the attractiveness of neighboring individuals. Whenever an

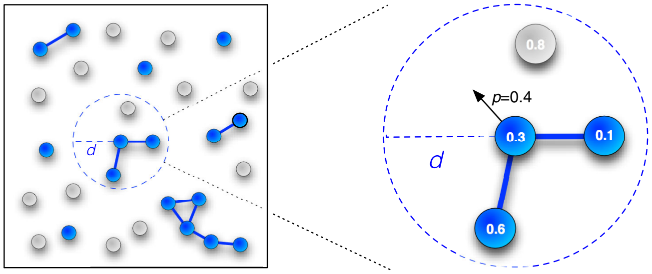

FIG. 1 (color online). Left: Blue (dark) colored agents are active, gray (light) agents do not move nor interact. Interacting agents, within a distance d, are connected by a link. Right: Each individual is characterized by a number representing her attractiveness. The probability for the central individual to move is $p = 1.0 - 0.6 = 0.4$, since the attractiveness of the inactive agent is not taken into account.

agent intercepts, within a distance smaller than or equal to d, another individual, they start to interact. We consider that both agents are interacting as long as their relative distance in smaller than d. Crucially, the more attractive an agent j is (the largest her attractiveness a_j), the more interest she will raise in the other agent i, who will slow down her random walk exploration accordingly. This fact is taken into account by a walking probability $p_i(t)$ which takes the form

$$p_i(t) = 1 - \max_{j \in \mathcal{N}_i(t)} \{a_j\}, \qquad (1)$$

where $\mathcal{N}_i(t)$ is the set of neighbors of agent i at time t, i.e., the set of agents that, at time t, are at a distance smaller than or equal to d from agent i. Hence, the biased random walk performed by the agents is defined as follows: At each time step t, each agent i performs, with probability $p_i(t)$, a step of length v along a direction given by a randomly chosen angle $\xi \in [0, 2\pi)$. With the complementary probability $1 - p_i(t)$, the agent does not move. Thus, according to Eq. (1), if an agent i is interacting with other agents, she will keep her position in the following time step with a probability proportional to the appeal of his most interesting neighbor.

Furthermore, the empirical observations of SocioPatterns data show that not all the agents involved in a social event are actually present for its entire duration: Some agents leave the event before the end, some join it later after the beginning, and some others leave and come back several times. Therefore we assume that agents can be in an active or an inactive state. If an individual is active, she moves in space and interacts with the other agents; otherwise she simply rests without neither moving nor interacting. At each time step, one inactive agent i can become active with a probability r_i, while one active and isolated agent j (not interacting with other agents) can become inactive with probability $1 - r_j$. The activation probability r_i of the individual i thus represents her activeness in the social event, the largest the activity r_i, the more likely agent i will be involved in the event. We choose the activation probability r_i of the agents randomly from a uniform distribution $\zeta(r)$, bounded in $r_i \in [0, 1]$, but we have verified that the model behavior is independent of the activity distribution functional form (even if we consider a constant activity rate, $r_i = r$ for all agents, we obtain very similar results, see Supplemental Material [26], Fig. 1).

Within this framework, each individual performs a discrete random walk in a 2D space, interrupted by interactions of various duration with peers. The movement of individuals is performed in parallel in order to implement the time resolution (20 sec) at which empirical measurements are made [8]. The model is Markovian, since agents do not have memory of the previous time steps. The full dynamics of the system is encoded in the collision probability $p_c = \rho \pi d^2$, the activation probability distribution $\zeta(r)$, and the attractiveness distribution $\eta(a)$. The latter can hardly be accessed empirically, and is likely to be the combination of different elements, such as prestige, status, role, etc. Moreover, in general, attractiveness is a relational variable, the same individual exerting different interest on different agents. Avoiding any speculations on this point, we assume the simplest case of a uniform distribution for the attractiveness [27]. Remarkably, this simple assumption leads to a rich phenomenology, in agreement with empirical data.

In the following we will contrast results obtained by the numerical simulation of the model against empirical results from SocioPatterns deployments in several different social contexts: a Lyon hospital ("hosp"), the Hypertext 2009 conference ("ht"), the Societè Francaise d'Hygiène Hospitalière congress ("sfhh"), and a high school ("school"). A summary of the basic properties of the data sets is provided in Table I (see Refs. [8,9,28] for further description and details). The model has been simulated adopting the parameters $v = d = 1$, $L = 100$, and $N = 200$. Different values of the agent density ρ are obtained by changing the box size L. In the initial conditions, agents are placed at randomly chosen positions,

TABLE I. Some properties of the SocioPatterns data sets under consideration: N, number of different individuals engaged in interactions; T, total duration of the contact sequence, in units of the elementary time interval $t_0 = 20$ sec; $\bar{p}$, average number of individuals interacting at each time step; $\langle \Delta t \rangle$, average duration of a contact; $\langle k \rangle$ and $\langle s \rangle$: average degree and average strength of the projected network, aggregated over the whole sequence (see main text).

Data set	N	T	$\bar{p}$	$\langle \Delta t \rangle$	$\langle k \rangle$	$\langle s \rangle$
hosp	84	20 338	0.049	2.67	30	1145
ht	113	5093	0.060	2.13	39	366
school	126	5609	0.069	2.61	27	453
sfhh	416	3834	0.075	2.96	54	502

and are active with probability $1/2$. Numerical results are averaged over 10^2 independent runs, each one of duration T up to $T_{\max} = 2 \times 10^4$ time steps.

The temporal pattern of the agents' contacts is probably the most distinctive feature of face-to-face interaction networks [8,9]. We therefore start by considering the distribution of the duration Δt of the contacts between pairs of agents, $P(\Delta t)$, and the distribution of gap times τ between two consecutive conversations involving a common individual, $P(\tau)$. The bursty dynamics of human interactions is revealed by the long-tailed form of these two distributions, which can be described in terms of a power-law function [8]. Figure 2 show the distribution of the contacts' duration $P(\Delta t)$ and gap times $P(\tau)$ for the various sets of empirical data along with the same distributions obtained by simulating the model described above with density $\rho = 0.02$. In the case of the contact duration distribution, numerical and experimental data match almost perfectly, see Fig. 2 (top). Moreover, numerical results are robust with respect to variations of the collision probability $p_c = \pi d^2 \rho$, as shown in the inset. It also worth highlighting the crucial role played by the heterogeneity of attractiveness a_i. In fact, assuming it is constant, $a_i = a$ (and neglecting excluded volume effects between agents) our model can be mapped into a simple first passage time problem [29], leading to a distribution $P(\Delta t) \sim (\Delta t)^{-3/2}$ with an exponential cutoff proportional to $d^2/(1-a)$. The (nonlocal) convolution of the exponential tails induced by the heterogeneous distribution of attractiveness leads in our model to a power law form, with no apparent cutoff, and with an exponent larger than $3/2$, in agreement with the result observed in the SocioPatterns data. Regarding the distribution of gap times, $P(\tau)$, the model also generates a long-tailed form, which is compatible, although in this case not exactly equal, to the empirical data, see Fig. 2 (bottom). The behavior of the distribution $P(\tau)$ yielded by the model is substantially independent of the agent density ρ also in this case, as shown in the inset.

Sociopatterns data can be naturally analyzed also in terms of temporally evolving graphs [10], whose nodes are defined by the agents, and whose links represent interactions between pairs of agents. Instantaneous networks are thus formed by isolated nodes and small groups of interacting individuals, not necessarily forming a clique. Integrating the information of these instantaneous graphs over a time window T, which we choose here equal to the total duration of the contact sequences defining each data set [30], produces an aggregated weighted network [3], where the weight w_{ij} between nodes i and j represents the total temporal duration of the contacts between agents i and j. The weight distribution $P(w)$ of the various data sets are broad [8,9], see Fig. 3 (main), showing that the heterogeneity in the duration of individual contacts persists even when contact durations are accumulated over longer time intervals. Figure 3 shows that the outcome of the model is again in excellent agreement with all empirical data, with the exception of the "hosp" database. The reason for the departure of this data set with respect to both other data sets and the model could be attributed to the duration T of the

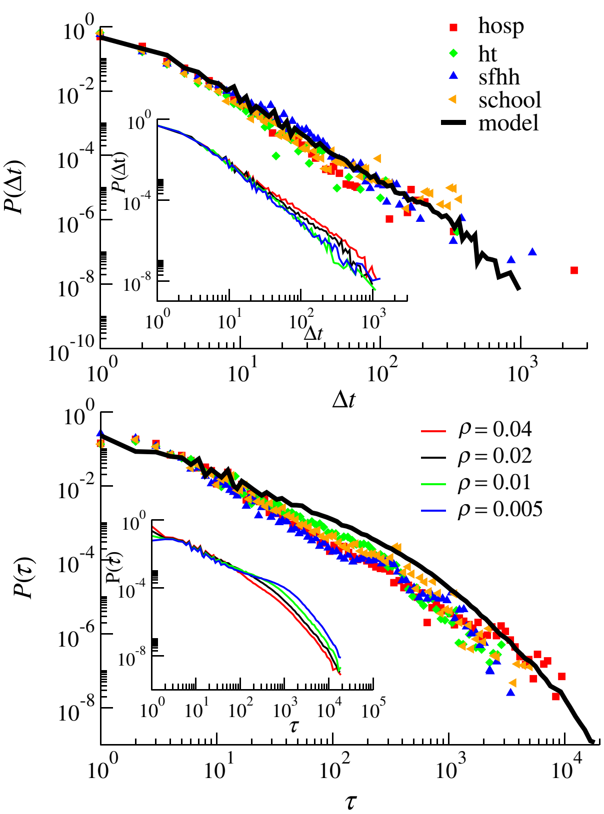

FIG. 2 (color online). Distribution of the contact duration, $P(\Delta t)$, (top) and distribution of the time interval between consecutive contacts, $P(\tau)$, (bottom) for various data sets and for the attractiveness model. Insets: Same distributions for the attractiveness model with different density. Symbols refer to empirical data; lines to results of the model, for different densities ρ.

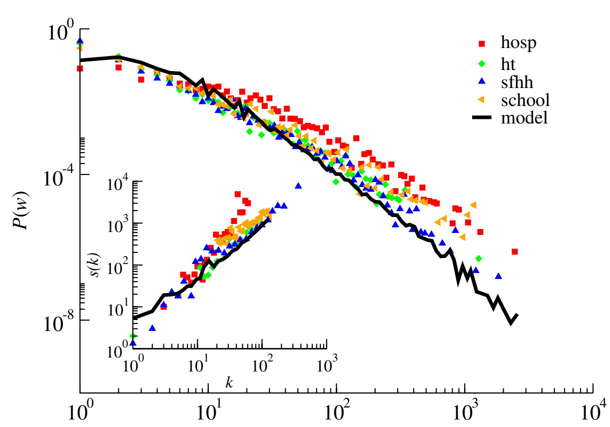

FIG. 3 (color online). Weight distribution $P(w)$ (main) and average strength of nodes of degree k, $s(k)$, as a function of k, (inset) for various empirical data sets and for the aggregate network obtained by simulating the attractiveness model.

corresponding sequence of contacts (see Table I), which is up to four times longer than the other data sets. In the limit of large T, sporadic interactions can lead to a fully connected integrated network, very different from the sparser networks obtained for smaller values of T. These effects extend also to the pattern of weights, which have in the "hosp" database a much larger average value.

Face-to-face networks can be further characterized by looking at the correlation between the number of different contacts and the temporal duration of those contacts. These correlations can be estimated by measuring the strength s_i of a node i, defined as $s_i = \sum_j w_{ij}$ and representing the cumulated time spent in interactions by individual i, as a function of its degree k_i, defined as the total different agents with which agent i has interacted. Figure 3 (inset) shows the growth of the average strength of nodes of degree k, $s(k)$, as a function of k in the empirical data sets and in the aggregated network obtained with the attractiveness model. As one can clearly see, all distributions (again with the exception of the "hosp" data set) are well fitted by a power law function $s(k) \sim k^{\alpha}$ with $\alpha > 1$, with good agreement between real data and the model results. The observed superlinear behavior implies that on average the nodes with high degree are likely to spend more time in each interaction with respect to the low-connected individuals [8].

A final important feature of face-to-face interactions, also revealed in a different context involving human mobility [20], is that the tendency of an agent to interact with new peers decreases in time. This fact translates into a sub-linear temporal growth of the number of different contacts of single individuals [i.e., the aggregated degree $k_i(t)$], $k(t) \sim t^{\mu}$, with $\mu < 1$. Figure 4 shows the evolution of $k(t)$ versus time for several agents with a final aggregated degree $k(T)$, both belonging to a single data set (main) and for the different data sets (inset). The sublinear behavior of $k(t)$ is clear, with $\mu = 0.6 \pm 0.15$ depending on the data set. Moreover, the shapes of the $k(t)$ functions can be collapsed in a single curve by appropriately rescaling the data as $k(t)/k(T)$ as a function of t/T, Fig. 4 (inset). Figure 4 shows that, remarkably, the attractiveness model is also capable of reproducing the behavior of $k(t)$, up to the rescaling with total T time, again with the exception of the "hosp" data set.

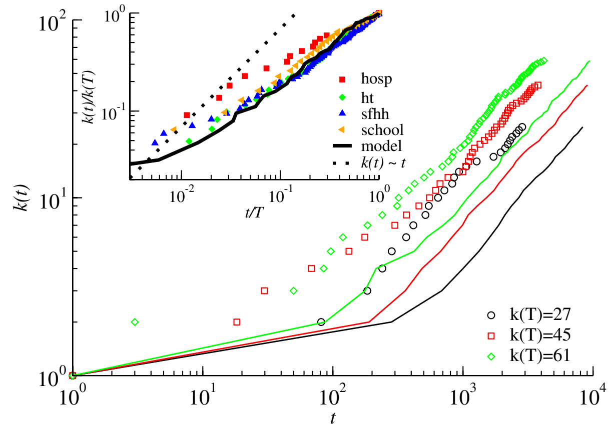

FIG. 4 (color online). Main: Aggregated degree $k(t)$ versus time for various individuals with different final degree $k(T)$, for the "ht" data set (symbols) and for the network obtained by simulating the attractiveness model (lines). Inset: Rescaled aggregated degree $k(t)/k(T)$ as a function of time t/T for various empirical data sets and for the attractiveness model.

In summary, in this Letter we have introduced a simple model of mobile agents that naturally reproduces the social context described by the SocioPatterns deployments, where several individuals move in a closed environment and interact between them when situated within a small distance (the exchange range of radio-frequency identification devices devices). The main ingredients of the model are (i) agents perform a biased random walk in two-dimensional space, (ii) their interactions are ruled by a heterogeneous attractiveness parameter, Eq. (1), and (iii) not all agents are simultaneously active in the system. Without any data-driven mechanism, the model is able to quantitatively capture most of the properties of the pattern of interactions between agents, both at the level of the broad distributions of contact and intercontact times, and at the level of the ensuing temporal network. Importantly, results are robust with respect to variations of the model parameters, i.e., the collision probability p_c and the activity distribution functional form, $\zeta(r)$. We have additionally checked that results do not depend qualitatively on the nature of the motion rule, given by Eq. (1). Indeed, other rules for the walking probability, such as considering the average of the attractives of the neighbors, i.e., $p_i(t) = 1 - \sum_{j\in\mathcal{N}_i(t)} a_j/k_i(t)$, lead substantially to the same behavior produced by Eq. (1) (see Supplemental Material [26], Fig. 2). Overall, the proposed framework represents an important step forward in the understanding of face-to-face dynamical networks. Confronted with other modeling efforts of SocioPatterns data [18], our model is not based on any cognitive assumption (reinforcement dynamics in Ref. [18]) and furthermore it leads to good agreement with experimental data without any fine tuning of internal parameters. It thus opens new interesting directions for future work, including the study of dynamical processes taking place in face-to-face networks and possible extensions of the model to more general settings.

We acknowledge financial support from the Spanish MEC, under Project No. FIS2010-21781-C02-01, and the Junta de Andalucía, under Project No. P09-FQM4682. R. P.-S. acknowledges additional support through ICREA Academia, funded by the Generalitat de Catalunya. We thank the SocioPatterns collaboration for providing privileged access to data sets "hosp," "school," and "sfhh." Data set "ht" is publicly available at Ref. [7]. We thank A. Barrat for helpful comments and discussions.

[1] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, Nature (London) **453**, 779 (2008).

[2] M. Jackson, *Social and Economic Networks* (Princeton University Press, Princeton, NJ, 2010).

[3] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).

[4] J. G. Oliveira and A.-L. Barabasi, Nature (London) **437**, 1251 (2005).

[5] A.-L. Barabasi, Nature (London) **435**, 207 (2005).

[6] D. Brockmann, L. Hufnagel, and T. Geisel, Nature (London) **439**, 462 (2006).

[7] http://www.sociopatterns.org.

[8] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, and A. Vespignani, PLoS ONE **5**, e11596 (2010).

[9] M. Starnini, A. Baronchelli, A. Barrat, and R. Pastor-Satorras, Phys. Rev. E **85**, 056115 (2012).

[10] P. Holme and J. Saramäki, Phys. Rep. **519**, 97 (2012).

[11] J. Stehle, N. Voirin, A. Barrat, C. Cattuto, V. Colizza, L. Isella, C. Regis, J.-F. Pinton, N. Khanafer, W. Van den Broeck, and P. Vanhems, BMC Medicine **9**, 87 (2011).

[12] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, and J. Saramäki, Phys. Rev. E **83**, 025102 (2011).

[13] S. Lee, L. E. C. Rocha, F. Liljeros, and P. Holme, PLoS ONE **7**, e36439 (2012).

[14] R. Parshani, M. Dickison, R. Cohen, H. E. Stanley, and S. Havlin, Europhys. Lett. **90**, 38 004 (2010).

[15] N. Fujiwara, J. Kurths, and A. Díaz-Guilera, Phys. Rev. E **83**, 025101 (2011).

[16] A. Vázquez, B. Rácz, A. Lukács, and A. L. Barabási, Phys. Rev. Lett. **98**, 158702 (2007).

[17] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, Sci. Rep. **2**, 469 (2012).

[18] K. Zhao, J. Stehlé, G. Bianconi, and A. Barrat, Phys. Rev. E **83**, 056109 (2011).

[19] H.-H. Jo, R. K. Pan, and K. Kaski, PLoS ONE **6**, e22687 (2011).

[20] C. Song, T. Koren, P. Wang, and A.-L. Barabasi, Nat. Phys. **6**, 818 (2010).

[21] B. D. Hugues, *Random Walks and Random Environments*, Random Walks Vol. I (Clarendon Press, Oxford, 1995).

[22] A. Baronchelli and A. Díaz-Guilera, Phys. Rev. E **85**, 016113 (2012).

[23] S. Valverde and R. V. Solé, Phys. Rev. E **76**, 046118 (2007).

[24] S. Athey, E. Calvano, and S. Jha, "A Theory of Community Formation and Social Hierarchy" (to be published).

[25] R. M. Sapolsky, Science **308**, 648 (2005).

[26] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.110.168701 for further testing of the model robustness against changes in parameter settings.

[27] F. Papadopoulos, M. Kitsak, M. Serrano, M. Boguná, and D. Krioukov, Nature (London) **489**, 537 (2012).

[28] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J. Pinton, and W. Van den Broeck, J. Theor. Biol. **271**, 166 (2011).

[29] S. Redner, *A Guide To First-Passage Processes* (Cambridge University Press, Cambridge, England, 2001).

[30] B. Ribeiro, N. Perra, and A. Baronchelli, arXiv:1211.7052.