Characteristics of the Dynamic of Mobile Networks

Pierre Borgnat, Eric Fleury, Jean-Loup Guillaume, Céline Robardet

Characteristics of the Dynamic of Mobile Networks⋆

Pierre Borgnat[1], Éric Fleury[2], Jean-Loup Guillaume[3], and Céline Robardet[4]

[1] Université de Lyon, ENS Lyon,
Laboratoire de Physique, UMR 5672 CNRS
69364 Lyon cedex, France
[2] Université de Lyon, ENS Lyon
INRIA/D-NET, LIP, UMR 5668 CNRS
69364 Lyon cedex, France
[3] Université Pierre & Marie Curie
LIP6 UMR 7606 CNRS
75252, Paris Cedex 05, France
[4] Université de Lyon, INSA-Lyon
LIRIS UMR 5205 CNRS,
69622 Villeurbanne cedex, France

Abstract. We propose in this paper a novel framework for the study of dynamic mobility networks. We address the characterization of dynamics by proposing an in-depth description and analysis of two real-world data sets. We show in particular that links creation and deletion processes are independent of other graph properties and that such networks exhibit a large number of possible configurations, from sparse to dense. From those observations, we propose simple yet very accurate models that allow to generate random mobility graphs with similar temporal behavior as the one observed in experimental data.

Key words: Dynamic Networks, Network Models, Complex Systems, Random Graphs, Statistical Analysis, Stochastic Process, Data Mining.

1 Introduction

During the last decade, the study of large scale complex networks has attracted a substantial amount of attention and works from several domains: sociology [15], biology [7], computer science [1], epidemiology [12]. This emerging domain has proposed a large set of tools that can be used on any complex network in order to get a deep insight on its properties and to compare it to other networks. Such *fundamental* properties [1, 10, 11] are used as characterization parameters in the

---

study of various problems such as virus spreading [5, 9, 12] in the epidemiology context, or information / innovation diffusion [2, 6] for instance. Whereas most of such complex networks are inherently dynamic, this aspect has less been studied. Most approaches consider *growing models*, such as the preferential attachment model [2, 8] or analyze the aggregation of all interactions. Both approaches may miss the real dynamic behavior while there is a strong need for dynamic network models in order to sustain protocol performance evaluations and fundamental analyzes.

In this talk, we address the description and the simulation of sensor mobility networks. The proposed methods come from various research domains (signal processing, graph theory and data mining). This emphasizes the necessity of interdisciplinary research since dynamic networks are becoming a central point of interest, not only for engineers and computer scientists but also for people in many other fields.

We apply those methods on mobility networks. Mobile devices with wireless capabilities are a typical example of evolving networks where users are spread in the environment and communications can only take place if they are near each others. We study an empirical mobility network, called IMOTE [3], based on 41 Bluetooth sensors whose interactions have been recorded during 3 days. This who-is-near-whom network evolves every time users move.

We introduce some simple methods to describe the network dynamics and propose models of dynamic networks. The complete methodology of analysis was reported in a full version of this communication [14].

1. We study graph properties as function of time to provide an empirical statistical characterization of the dynamics.
2. We also compute global indicators from the dynamics of the network (connected components, triangles, and communities).
3. We propose models to perform random dynamic networks simulations.

The descriptive analysis show that link (or edge) creation and deletion processes is mostly independent of other graph properties and that such networks exhibit a large number of possible configurations, from sparse to dense. From those observations, we propose simple yet accurate models that allow to generate random mobility graphs with similar temporal behavior as the one observed in experimental data.

Even though such networks have obvious specificities, the in-depth study of their dynamic is an original work, and can have a broader impact on the complex system community. It is noteworthy that our approach does not make any assumption on the specificities unlike agent-based models or geographical approaches.

2   Statistical analysis of snapshots of graphs

We first propose and study a set properties usable as a practical basis for the analysis of dynamic mobility networks that can be easily extended to large complex networks. The studied graph properties are the distributions of contact

and inter-contact durations, the correlation between various graph properties as function of time and the links correlation, so as to give an empirical statistical characterization of the dynamics.

2.1   Contact and inter-contact durations

The contact and inter-contact duration distributions are dynamic characteristics that are interesting for mobility networks. The contact duration is the time during which two vertices remain directly and continuously adjacent. The inter-contact duration is the duration between two periods of contact for two vertices.
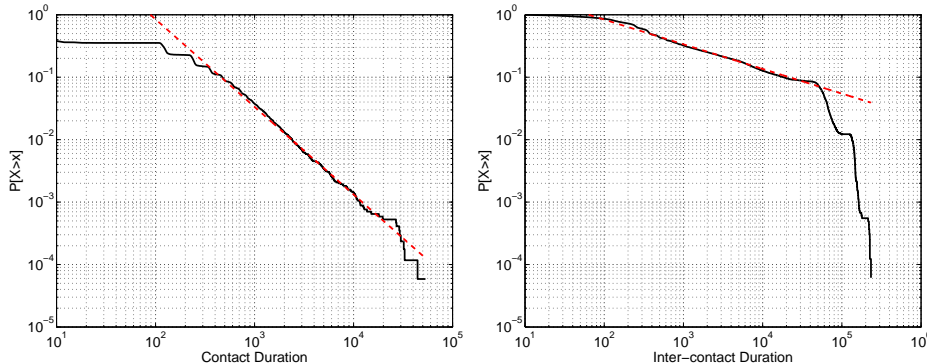

Fig. 1. Contact (left) and Inter-contact (right) duration distributions (CCDF).

Fig. 1 shows that the contact and inter-contact durations have a power-law behavior [4]. Inter-contact duration distribution has a very strong variability due to long periods of lack of contact for some nodes, whereas the distribution of contact durations is less heavy-tailed. The heavy-tailed nature of these distributions seems to be an ubiquitous property of dynamic mobility networks.

2.2   Correlation of graph properties

We compute the correlation coefficient between several graph properties seen as functions time t: $E(t)$ is the number of active links, $V(t)$ is the number of connected vertices, $N_c(t)$ is the number of connected components, $D(t)$ is the average degree, $T(t)$ is the number of triangles, $E_\oplus(t)$ is the number of links added at time t and $E_\ominus(t)$ is the number of links removed at that time.

Most of the correlation coefficients are rather high. This is mostly due to the fact that there are constraints on the properties of graphs. For instance the number of links $E(t)$ has a strong influence on the number of connected vertices $V(t)$. Furthermore, the time series are not stationary and there are clear

	$E(t)$	$V(t)$	$N_c(t)$	$D(t)$	$T(t)$	$E_\oplus(t)$	$E_\ominus(t)$
$E(t)$	1	0.85	-0.56	0.95	0.90	**0.19**	**0.15**
$V(t)$	0.85	1	-0.20	0.70	0.66	**0.15**	**0.11**
$N_c(t)$	-0.56	-0.20	1	-0.70	-0.41	**-0.16**	**-0.15**
$D(t)$	0.95	0.69	-0.69	1	0.86	**0.19**	**0.15**
$T(t)$	0.90	0.66	-0.41	0.86	1	**0.15**	**0.11**
$E_\oplus(t)$	**0.19**	**0.15**	**-0.16**	**0.20**	**0.15**	1	**0.03**
$E_\ominus(t)$	**0.15**	**0.11**	**-0.15**	**0.16**	**0.10**	**0.03**	1

Table 1. Correlation coefficients between various graph properties studied as functions of time.

periods of one day and variations between days and nights. Only link creation and deletion processes ($E_\oplus(t)$ and $E_\ominus(t)$) remain mostly uncorrelated with all other properties. Their evolution can be considered mostly independent from the one of other graph properties.

2.3   Links correlations

Let us now turn to individual links. The correlation coefficient of the state evolution of links characterizes the dependency between links. The state evolution $S_e(t)$ of each link e is equal to 1 if link e is in the mobility graph at time t and 0 otherwise. The correlation matrix $Co(e, e')$ for links is computed as:

$$Co(e, e') = CORR(S_e, S_{e'})$$
$$= < S_e(t)S_{e'}(t) >_t - < S_e(t) >_t < S_{e'}(t) >_t .$$

For each link, we also compute its average correlation coefficient with respect to the other links as the average of absolute values. This helps to keep track of the strength of the correlation rather than its direction.

Fig. 2 shows the histogram of the values. Most pairs of links have a very low correlation coefficient. Rare couples of links exhibit a strong correlation.

2.4   Joint distribution

The empirical joint distribution of the number of connected nodes and the number of links gives a finer description of the dependencies between those two properties.

As expected, the correlation between vertices and links shown on Fig. 3 is positive: the more vertices are connected, the more links are present. However it is worth noting that the variation of the number of links is not constant over the number of vertices. For a given number of vertices, the network can have a large number of possible configurations, some of which are very sparse and others more dense, as shown by the gray scale in the plots.
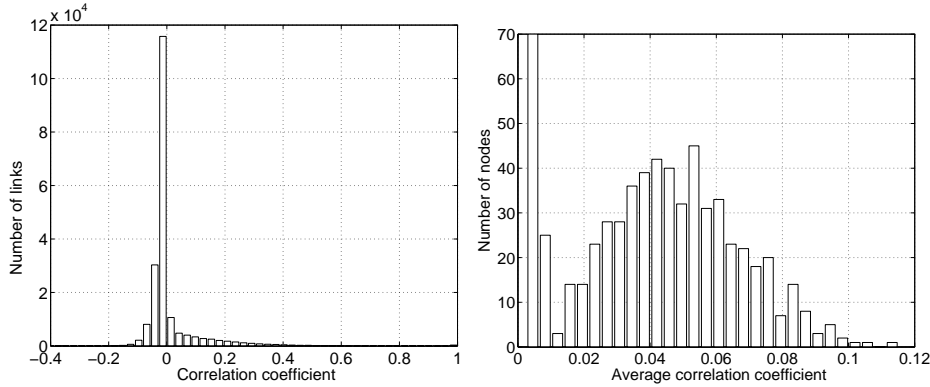
Fig. 2. Correlation between links (left) and average correlation for each node (right)

3 Global view of the dynamics

We also propose to study global indicators from the dynamics of the network (stability of connected components, triangles creations, existence of communities).

3.1 Triangles

The existence and persistence of connected components is generally associated with a rather large number of triangles in the graph. Therefore, an important characteristic of the dynamic is the evolution of the number of triangles in time. To evaluate the proportion of links that create triangles when they appear, we compute the number of link creations that leads to an increase of the number of triangles in the graph or that does not change it.

	$P_{+/tri+}$	$P_{+/tri=}$	$f_{+/tri+}$	$f_{+/tri=}$
IMOTE	44 %	56 %	6 %	94 %
RANDOM	10 %	90 %	5 %	95 %

Table 2. Proportion $P_{+/tri+}$ (resp. $P_{+/tri=}$) of links creations that add new triangles (resp. not), and the average proportion $f_{+/tri+}(resp. f_{+/tri=})$ of inactive links that, if created, would add a triangle, (resp. not).

These proportions are given in Table 2. Around 40% of links creations increase the number of triangles in the graph whereas this proportion equals 10% in a random (Erdös-Rényi) graph with the same numbers of vertices and links. The proportion of inactive links that would create a triangle is very low for IMOTE data set and the simple random graph. This emphasizes the fact that
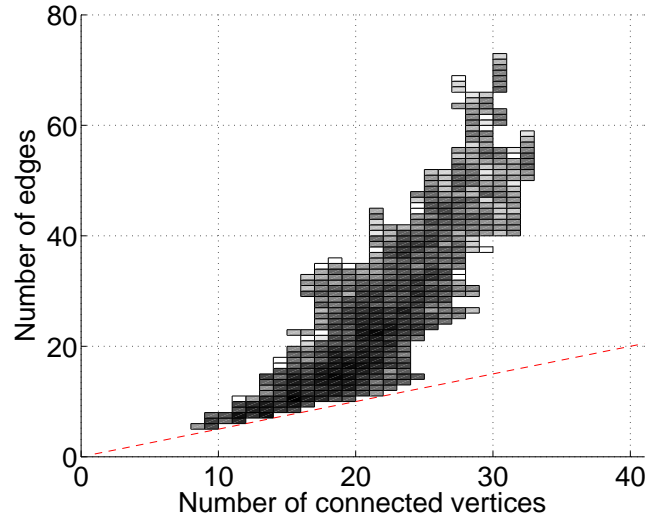
Fig. 3. Joint distribution of the number of connected nodes and the number of links.

this is not because more links can create triangles that the proportion P+/tri+ is higher in experimental data: it is on the contrary an intrinsic property of the dynamics. As the proportions of links that could create a triangle are similar in both graphs ($f_{+/tri+}$), this phenomenon is characteristic of real graphs: links creations tend to create triangles in fairly large proportion.

3.2   Dynamic communities

To describe the graph structure evolution, we isolate "communities", which are commonly considered as large groups of individuals who interact intensively with each other over a long period of time. A community can be seen as a dense connected sub-graph that appears in a large number of time steps (not necessarily consecutive).

We compute the set of connected sub-graphs having more than σ links and that are included in at least τ graphs:

$\mathcal{C} = \{S = (V, E), |\{t\,|\,S \subseteq G_t\}| \geq \tau$ and $|E| \geq \sigma$ and S is connected$\}$. Then, the denser sub-graphs are selected using a density threshold that selects the most important and established ones. Finally, the trajectories of individuals among social groups are inferred: an arc (u,v) represents individuals moving at least once in the data from group u to group v. Fig. 4 shows the identified communities and their dynamic. For example, individual 8 initially belongs to group 13, he/she further moves into group 6, and finally enters group 7.
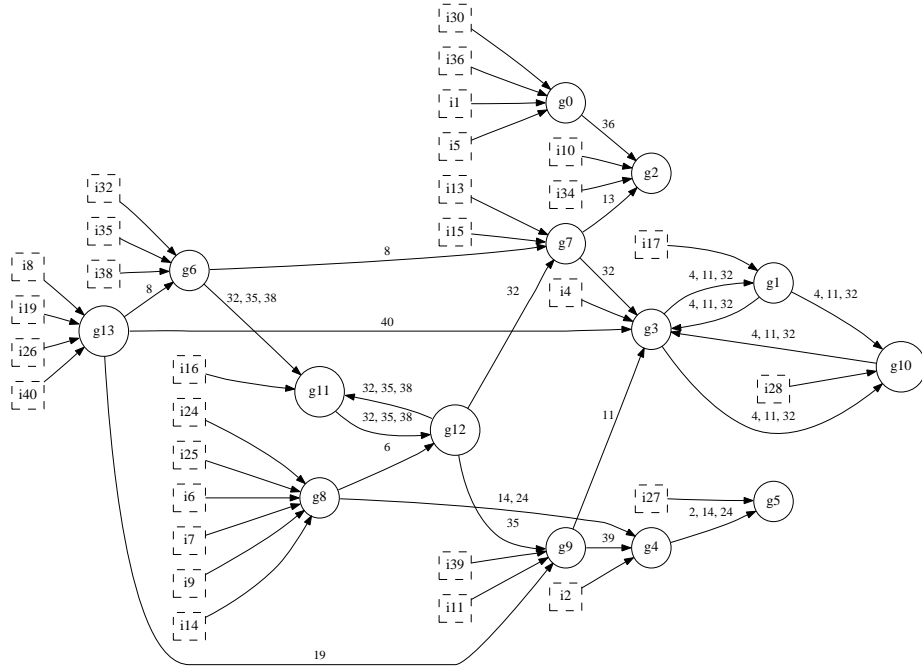
Fig. 4. Individual trajectories in groups ordered by time. i_x (boxes) are individuals while g_x (circles) denotes social groups.

4 Modeling of the dynamics

From the previous set of analysis, we propose generic random dynamic models that allows to generate random dynamic graphs which have a behavior similar to the one observed in experimental data set. Their design is justified by the previous observations. First, as the contact and inter-contact are power-law distributed (as seen in Section 2.1), those non-trivial empirical distributions should be taken into account when constructing a model of the data. Second, computation of empirical times of correlations show that link creation/removal process is less correlated in time that other graphs properties seen in Sect. 2.2. It is also reasonable to consider that links evolutions are uncorrelated in time. These characteristics justify the use of a simple Markovian (memory-less) link creation/removal process. Finally, we observed that for a given number of nodes, the network can have a large number of possible configurations, some of which are very sparse and others more dense. Thus our model should be able to reproduce this property.

The simulation is based on a transition model with Markovian property. For each time step and for each link independently, each link changes its state (active or inactive) using a transition probability depending on the time since the
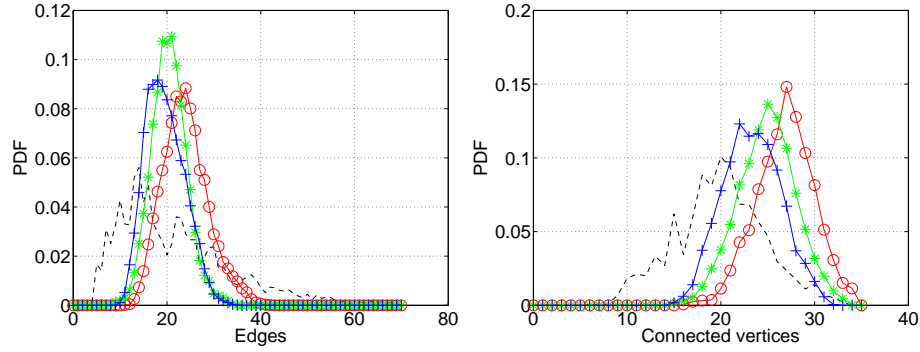
Fig. 5. : Probability distribution function (PDF) for original data and the classical models. On the left we plot the PDF for the number of connected vertices and on the rigth we plot the number of edges. The plots are for original data (−) and for several models: imposing the sole contact and inter-contact duration distribution (-o- on the plot), or adding the statistics of N_C (-*-) or V (-+-).

link is in its current state. In addition, the probability of transition is weighted by a probability of acceptance of the new state depending of the experimental distribution for a property of interest such as $E(t)$, $V(t)$, $N_C(t)$ and $D(t)$. This is implemented by Rejection Sampling [13], based on a Metropolis-Hasting algorithm and take into account the target distribution of the graph property. To take into account in the simulations that the average proportion of link creations that yield triangles is larger than for random graphs, a weight is applied on the transition probability to reproduce the correct dynamical transition process concerning triangles.

In order to justify that the use of properties (like $E(t)$, $V(t)$, $N_C(t)$) within the model in addition to the sole contact and inter contact duration distribution we compare such distribution with the original Data set. A first remark is that the sole contact and inter-contact duration distributions dramatically fail to reproduce the properties. More precisely, the number of connected vertices is strongly over-estimated, the number of connected components is under-estimated, and so is the number of triangles. The non-stationarity in the IMOTE data introduces a much higher variance, yet it does not explain all the differences.

Fig. 5 illustrates the way the models work. It is based on the distribution of $E(t)$ and $N_C(t)$ and produces qualitatively similar characteristics than the real data set. The probability distribution functions of $N_C(t)$ (on the rightmost figure) for the original data (black), and the different models proposed are shown. We can observe that by imposing the distribution of the number of connected vertices improves the accuracy of the simulations.

However, when analyzing more refined characteristics such as the joint distribution of number of links and vertices in connected components, the original data set (shown on Fig 3) and the simulations of the models (shown on Fig. 6)
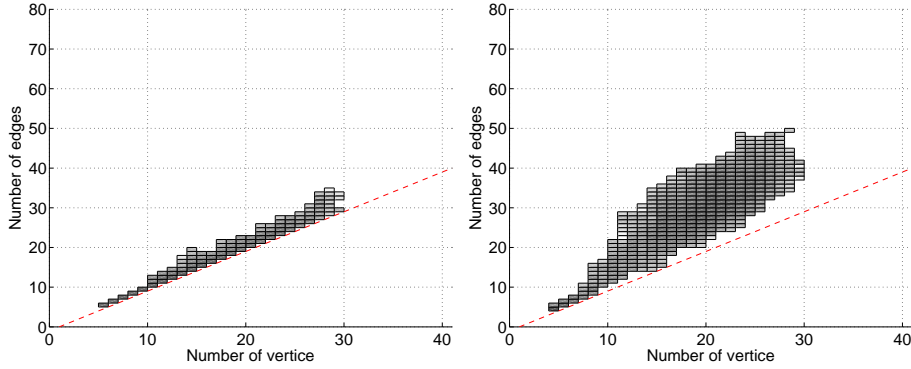
Fig. 6. : Probability distribution function (PDF) for original data and the classical models. On the left and right we plot joint distribution of the number of connected vertices and links in connected components: a model imposing the contact and inter-contact duration distribution and the statistics of N_C (middle plot) fails at reproducing the correct behaviour; a model respecting also the statistics of the triangle dynamics (right plot) reproduces much more the empirical behaviour of the data.

are much different even when we impose number of connected vertices distribution. The connected components in models (sole inter contact and/or with the introduction of graph property PDF) are much less dense. We believe this is of major importance for communication protocol design and realistic models have to reproduce this property and this yield to the introduction of the number of triangle property (the results is depicted on the rightmost figure of Fig 6). This time, the density of connected components is comparable to the original data set. Note that introducing the dynamic of triangles also yield to the creation of "social" groups like the ones depicted in Fig. 4. When using the classical models, the graph is too sparse and the community phenomena is not reproduced in the model.

Our investigations have shown that the model, thanks to the introduction of dynamical characteristics such as the evolution of the number of triangles, manages to generate more realistic simulations. This opens the track to improved models that match the important characteristics of dynamics of mobility networks. This is illustrated by the two last figures that show the dynamic communities in the output of the simulation and the joint probabilities of the number of connected vertices and links in the graph.

This study opens the track to improved models that match the important characteristics of dynamics of mobility networks.

5   Conclusion

By introducing several models, we are able to highlight the diversity of properties that are needed to characterize such networks. Furthermore, our models provide

insight into existing notions of dynamic networks and demonstrate that the structure and the dynamics are complex and are not a direct consequence of the contact and inter-contact durations. Proposing such models is crucial since it enables a validation of the ongoing research conducted in the various areas that deal with dynamic networks. It has also many applications in performance evaluation for instance.

References

1. R. Albert and A.-L. Barabasi. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47, 2002.
2. R. Albert, H. Jeong, and A. Barabasi. The diameter of the World Wide Web. *Nature*, 401:130–131, 1999.
3. A. Chaintreau, J. Crowcroft, C. Diot, R. Gass, P. Hui, and J. Scott. Pocket switched networks and the consequences of human mobility in conference environments. In *WDTN*, pages 244–251, 2005.
4. A. Chaintreau, J. Crowcroft, C. Diot, R. Gass, P. Hui, and J. Scott. Impact of human mobility on the design of opportunistic forwarding algorithms. In *INFOCOM*, 2006.
5. A. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *INFOCOM*, 2005.
6. D. Gruhl, D. Liben-Nowell, R. V. Guha, and A. Tomkins. Information diffusion through blogspace. *SIGKDD Explorations*, 6(2):43–52, 2004.
7. D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *KDD*, pages 137–146. ACM Press, 2003.
8. J. Leskovec, D. Chakrabarti, J. Kleinberg, and C. Faloutsos. Realistic, mathematically tractable graph generation and evolution, using kronecker multiplication. In *ECML/PKDD*, 2005.
9. L. Meyers, B. Pourbohloul, M. Newman, D. Skowronski, and R. Brunham. Network theory and sars: Predicting outbreak diversity. *J. Theor. Biol.*, 232:71–81, 2005.
10. M. Newman. The structure and function of complex networks. *SIAM Review*, pages 167–256, 2003.
11. M. E. Newman and J. Park. Why social networks are different from other types of networks. *Phys Rev E Stat Nonlin Soft Matter Phys*, 68:036122, 2003.
12. R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Let.*, 86:3200–3203, 2001.
13. C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2004.
14. A. Scherrer, P. Borgnat, E. Fleury, J.-L. Guillaume, and R. C. Description and simulation of dynamic mobility networks. *Computer Networks*, 52(15):2842–2858, 2008.
15. D. Watts and S. Strogatz. Collective dynamics of small-world networks. *Nature*, 293:420–442, 1998.