Phase Transition in Opportunistic Mobile Networks

Augustin Chaintreau, Laurent Massoulié
Thomson, Paris, France. email: firstname.lastname@thomson.net

***Abstract*—Opportunistic mobile networks take advantage of local opportunities of wireless communication between devices (cellphones, etc.) to construct a path over time between a source and a destination. This paper uses a model of random temporal network to study the existence of those paths that use a small number of time slots and a small number of steps. It establishes that a phase transition occurs as time and hops are jointly increase according to the logarithm of the network size. For a given intensity of contact, as time grows the network abruptly change from a regime where almost surely no path exists to a regime where paths exist with a positive probability. Our proof illustrates a strong correlation close to the critical point between nearby paths (those who share a prefix and suffix term), which explain the relatively high value of the variance. We identify combinatorial properties specific to temporal paths, which are critical to characterize the phase transition and impact the estimation of the probability of success. We believe that it is the first rigorous proof that a phenomenon recalling the small world effect may be found in dynamic random graphs.**

I. Introduction

The proliferation of mobile devices with increasing capacity for storage and communication brings new networking opportunities. For delay tolerant applications, one could consider using the opportunistic bandwidth available while two devices are co-located, and some of their local data storage, to transport data over time between devices that are otherwise disconnected. This forwarding method, which allows to communicate possibly without using any infrastructure, relies on the construction of a path over time, which should implicitly uses the diversity and fluctuation of mobility to find a way towards a particular point. Several research works recently consider such networking option as an alternative to infrastructure or capacity limitation of traditional networks which use a contemporaneous path. ”Routing over time”, in contrast with routing at a fixed time, appears as a problem with new constraints which in turn call for different solutions than existing routing protocols, in particular when some uncertainty is found in the network and links that are difficult to predict.

Following a recent research direction [1], [2], [3], this paper studies the properties of the paths available between nodes in an opportunistic context where links created by mobility follows unpredictable pattern. In contrast with other works in the area, we do not focus on the performance of one or several forwarding algorithms. We would rather like to identify important properties generally present in such context, and which are relevant to design and understand how any forwarding technique behaves. To make the problem easy to capture, we consider a nominal case corresponding to a uniform random temporal network, where we will apply combinatorial techniques. We assume that, at a given point in time, each pair of nodes may met, independently of others, according to a fixed probability. We focus on large networks, where the number of nodes grows arbitrarily large, while keeping the amount of local contacts seen on average by each node fixed.

In this paper, we establish that the presence of successful paths, or, in other words, the chance that solutions exist for the “routing over time” problem is tightly coupled to a phase transition occurring around the logarithm of the size of the network. Preliminary results about this phase transition have been reported in [2], where similar results have been compared with empirical findings. This work confirms rigorously that the “small world phenomenon” applies to random temporal networks in the following sense: paths exist according to a positive probability between two arbitrary nodes when the time and the number of hops are scaled properly as the logarithm of the network size.

Our work sheds new light on the nature of dynamic graphs or temporal networks that include random links. Temporal networks, which are graphs featuring a time dimension, where paths should follow a chronological property, have been studied from an algorithmic standpoint for several decades. Finding a maximum flow, a shortest path or checking connectivity were shown to be significantly different problem than their static counterpart [2]. However, we are not aware of any work studying properties of random temporal networks. As we demonstrate rigorously for the first time in this paper, phase transition occurs in random temporal networks according to a threshold that depends on both time and hops. To our belief, it proves that this model we introduced is worth being considered for future investigation on the properties of random structure with time.

The proof of the main result relies on an application of the well known second moment method, which shows that, with a positive probability, the number of paths remains close enough to its expectation to be positive. However, a few technical results found in the proof are surprising: First, it proves that the variation of these number of paths remains of the same order as the expectation. Hence we can only conclude thanks to an advanced second moment method deduced by Janson, Łuczak and Ruciński from the Fortuin-Kasteleyn-Ginibre inequality. Second, this variation comes only from the correlation of a small neighborhood around each path: those paths sharing a prefix and suffix. It indicates that the solutions to the “routing over time” problem are to be found in small clusters of paths which are locally highly correlated, whereas the realization of solutions outside these clusters occur almost independently.

II. Random Temporal Networks

Our model of random temporal network is a variant of the model described in [1]. It follows simplifying assumptions about mobility of human that are in general not met in practice. However, the qualitative conclusion can be expressed with regard to a new definition of diameter in temporal network, and it appears valid in several empirical findings. More about this point may be found in [1].

a) Nodes, contacts: We consider a network made of N nodes, generally denoted by index i. We assume that the network lasts for $t(N)$ time slots. We define the set of all possible opportunistic contacts:

$$\mathcal{C} = \{ \ (i,j,t) \ \mid 1 \leq i,j \leq N \, , \, i \neq j \, , \, 1 \leq t \leq t(N) \ \} \ .$$

Note that we consider here directed contacts where i is able to transmit to j at time t although j may not necessarily be able to transmit to i at the same time. The undirected case can be considered as well in the same framework, but we will ignore it in this paper.

We assume that all contacts occur independently and with the same probability p. We denote by λ the average number of contacts made by each node in a single time slot. We call this number the intensity of the network, and we will assume that it remains constant as N grows. In other words, we will choose $p = \frac{\lambda}{N}$.

Let us denote by Γ_p the subset which contains all contacts actually occurring in the network. It is a binomial random subset capturing all the information about the network. For any subset of contacts $A \subseteq \mathcal{C}$, according to a usual notation, we denote by I_A the indicator that this subset of contacts are all occurring:

$$I_A = \mathbb{I}_{\{A \subseteq \Gamma_p\}} \ .$$

b) Paths: A path of length k, for any $k \geq 1$ is simply defined as a subset of $\mathcal{C}$ with some property: it should contain sequence of contacts forming a paths among the nodes (no nodes is allowed to appear twice), and it should satisfy a chronological property.

$$A = \{ \ (i_0, i_1, t_1) \, , \, (i_1, i_2, t_2) \, , \, \ldots \, , \, (i_{k-1}, i_k, t_k) \ \} \ ,$$

such that $s, i_1, \ldots, i_{k-1}, d$, are all distinct nodes and $t_1 < t_2 < \cdots < t_k$.

Hence a path of length k is characterized by a subset of k transition times chosen among $\{ \ 1, 2, \ldots, t(N) \ \}$ and a sequence of $k+1$ distinct nodes. If the source and destination have been already fixed, there remains $k-1$ distinct nodes to choose in the sequence.

Let us fix two distinct nodes s and d arbitrarily. In the following of this article, we denote by S, the set of paths leading from s to d, between time $t = 1$ and $t = t(N)$, and that are of length $k(N)$. Based on the remark above, one can deduce:

$$|S| = (N-2) \ldots (N - k(N)) \times \binom{t(N)}{k(N)} \ .$$

c) Number of successful paths: Finally we define by X the number of paths in S which are occurring in the network. This is a random variable that may be written as follows:

$$X = \sum_{A \in S} I_A \ .$$

It enumerates exactly the number of successful paths that may be constructed from s to d if $t(N)$ time slots can be used and $k(N)$ hops are allowed. The goal of the next section is to show that, when t and k are of the order of a logarithm, the variable X can be in two possible regimes depending on the constant involved. This characterizes a phase transition occurring at this scale jointly with time and hops.

III. Phase transition

The main result of this paper is the following:

Theorem 1: Let $t(N) = \lfloor \tau \cdot \ln(N) \rfloor$ and $k(N) = \lfloor \gamma\tau \cdot \ln(N) \rfloor$, where $\gamma < \lambda$, then as N grows large

$$\begin{cases} \mathbb{P}\left[X = 0\right] \to 1 & \text{if } \ 1/\tau > \gamma \ln(\lambda) + h(\gamma) \, , \\ \mathbb{P}\left[X = 0\right] < e^{-\frac{1}{1 - \frac{1}{(1-\frac{\gamma}{\lambda})^2}}} & \text{if } \ 1/\tau < \gamma \ln(\lambda) + h(\gamma) \, , \end{cases}$$

where h is the entropy function that is defined on $[0;1]$ by $h : \gamma \mapsto -\gamma \ln(\gamma) - (1-\gamma) \ln(1-\gamma)$

Before proving the theorem, we would like to make some important remarks:

- The condition $\gamma < \lambda$ is not restricting, as the maximum value of the RHS in the inequality of the phase transition is obtained for $\gamma = \frac{\lambda}{1+\lambda}$. In other words, this theorem proves that the delay optimal path occurs with probability at least $e^{1 - \frac{(\lambda+1)^2}{\lambda^2}}$. Moreover, whenever the condition $\gamma < \lambda$ is not verified, we can still deduce that paths exist with a positive probability and a *shorter* length.
- The proof of the subcritical regime is not new. In this regime, almost surely no path exists, as it was already shown in [2] from an estimation of the expectation $\mathbb{E}\left[X\right]$. The proof of the supercritical regime (where a path exist with positive probability) is significantly more involved as it requires to bound the variation and hence capture the correlation between the different terms summing in X.
- The bound on the probability in the supercritical regime may not be tight. In particular, it is possible that the transition is sharp and this probability is 1. However, we prove that such result cannot be simply derived from the second moment method, as shown by a lower bound on the variance of X.

Figure 1 plots as a function of γ the value of the RHS in the condition, for three values of λ. As can be observed on this plot, a maximum exist that correspond to the path of minimum delay. For any value of τ above a minimum value, there exists a range of possible value for γ where paths exist with positive probability.

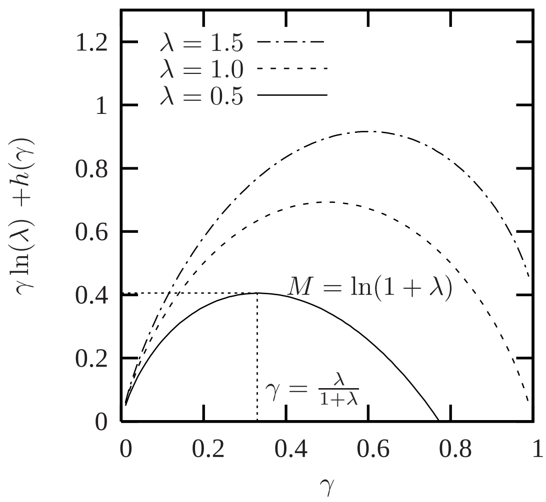

Fig. 1. Phase transition (short contact case)

A. First and second moment method

1) Expectation, and the subcritical regime: Let us denote by μ the expected number of successful paths $\mu = \mathbb{E}[X]$. Note that a given path A occurs with probability given by $(\lambda/N)^{k(N)}$, so that as N grows

$$\mu \sim \lambda^{k(N)} \frac{1}{N} \binom{t(N)}{k(N)} \tag{1}$$

It is a rather easy application of Stirling formula (see [2]) that, as N grows,

$$\mu \sim \frac{\lambda^{k(N)} N^{-1} e^{t(N)h(\gamma)}}{\sqrt{2\pi\tau\gamma(1-\gamma)\ln(N)}} \sim \frac{N^{-1+\tau(\gamma\ln(\lambda)+h(\gamma))}}{\sqrt{2\pi\tau\gamma(1-\gamma)\ln(N)}} .$$

This proves that, under the condition of the subcritical regime, μ goes to zero as N grows. Since X takes only integer values, an immediate application of Markov inequality shows that $P[X > 0]$ vanishes as N grows, proving the first half of the theorem.

2) Variance, and the second moment inequality: We define

$$\Delta = \sum_{A,B \text{ distinct, } A\cap B \neq \emptyset} \mathbb{E}[I_A I_B]$$

The following inequality applies to all variables defined with a binomial random subset, like the variable X. It combines the FKG inequality and Chernoff bound to prove that an upper bound on the ratio Δ/μ^2 is sufficient to prove that a variable is non-null with positive probability. The proof may be found in [4] (Theorem 2.18 p.33):

Lemma 1: $P[X = 0] < \exp\left(-\frac{\mu^2}{\mu + \Delta}\right)$.

B. The weight of fixed intersection

During all this section, we fix a path A chosen in $\mathcal{S}$, and $C \subseteq A$ a subset of A that we suppose non-empty. To avoid unnecessary notation, we denote $t(N)$ by t and $k(N)$ by k.

The goal of this section is to upper bound the contribution in the sum Δ of the terms where B is chosen in $\mathcal{S}$ such that $A \cap B$ is exactly C. Note first that under this condition,

$$\mathbb{E}[I_A I_B] = \mathbb{P}[A \cup B \subseteq \Gamma_p] = \left(\frac{\lambda}{N}\right)^{2\cdot k - |C|} . \tag{2}$$

This comes from the fact that $|A \cup B| = |A| + |B| - |A \cap B|$ and that A and B contains k contacts each.

Since all paths B as defined above contributes the same expectation in Δ, we only need to count the amount of such subset. The originality of the proof is on the method used to control these subsets. Let us start with a simple case, assuming C contains a single element, that is a given transition of A. One can see that there are two possible cases:

- If, on the one hand, C contains an extremal transition (i.e. C contains (s, i_1, t_1) or (i_{k-1}, d, t_k)), then there exist approximately
$$N^{k-2} \binom{t-1}{k-1}$$
paths B such that $A \cap B = C$. Indeed, since s,d and either i_1 or i_{k-1} are all already fixed, it remains $k+1-3 = k-2$ vertexes to be chosen to construct B. We also have to choose $(k-1)$ times for transition, in at most $t-1$ possible times (since the transition time selected by C is now forbidden).
- If, on the other hand, C contains a internal transition of A (i.e. $C = \{(i_{j-1}, i_j, t_j)\}$ with $2 \le j \le k-1$), then the number of possible paths B such that $A \cap B = C$ is at most
$$N^{k-3} \binom{t-1}{k-1} .$$
Note that it is a coefficient strictly less than the one above, since the vertexes already fixed by C in B are s, d, i_{j-1}, i_j, and there remain only $k+1-4 = k-3$ to be chosen.

As a conclusion, the amount of possible paths intersecting A exactly on C is smaller if intersection are avoiding extremal case. This following generalizes this remark.

We denote by $e_j = (i_{j-1}, i_j, t_j)$ the j_th edge of the path A. We can decompose $C \subseteq A$ as: there exists $2 \times l$ integers

$$1 \le a_1 < b_1 < a_2 < b_2 < \cdots < a_l < b_l \le k\,, \quad \text{such that}$$

- $e_j \in C$ if and only if $\exists m\ j \in [a_m, b_m]$,
- $[b_{m-1}+1, a_m-1]$ is non-empty (i.e. $a_m - b_{m-1} \ge 2$), and any index j contained in this interval verifies $e_j \notin B$.

The number of interval l above depends on C and is the number of fragments made of contiguous index of transitions in A. We define the number of fragments that are internal as the one not containing neither the index 1 nor k. It may be defined as

$$l_{\mathtt{int}}(C) = l - \mathbb{I}_{\{a_1=1\}} - \mathbb{I}_{\{b_l=k\}} .$$

Lemma 2: For any $A \in \mathcal{S}$ and $C \subseteq A, |C| = c$, we have

$$|\{\ B \in \mathcal{S}\ \mid A \cap B = C\ \}| \le N^{k-1-c-l_{\mathtt{int}}(C)} \binom{t-c}{k-c} .$$

Proof: We have by definition of the decomposition

$$\sum_{m=1}^{l} |[a_m, b_m]| = \sum_{m=1}^{l} (b_m - a_m + 1) = c\,.$$

The result comes from the following observation. Given that A is fixed, and that B share all contacts in C with A, a number of vertexes appearing in B are already chosen. First, it is obvious that s and d has been fixed in advance to be included in B.

Any fragment $[a_m, b_m]$ involves the vertexes given by $i^A_{a_m-1}, i^A_{a_m}, \ldots, i^A_{b_m}$, a total of $b_m-(a_m-1)+1 = b_m-a_m+2$. Note that these subset of vertexes are disjoint for two different fragments, because by definition $[b_m+1, a_{m+1}-1]$ should not be empty. Hence, the union of these subsets of vertexes contain

$$\sum_{m=1}^{l} (b_m-a_m+2) = \sum_{m=1}^{l} (b_m-a_m+1)+l = c+l \quad \text{vertexes.}$$

We should add s (respectively, d) to the union of this subset, if it has not already been included before, which is the case if and only if $a_1 > 1$ (resp. $b_l < k$). Hence the total number of vertexes already determined in B is

$$c + l + (1 - \mathbb{I}_{\{a_1=1\}}) + (1 - \mathbb{I}_{\{b_l=k\}}) = c + 2 + l_{\texttt{int}}\,.$$

The lemma follows the fact that the path B involves exactly $k+1$ vertexes in total (when we include s and d). In addition, $k-c$ transition times remain to be chosen, and none of them can take any of the c transition times already contain in C. ■

Lemma 2 and Eq.(2) implies for any A, C

$$\sum_{B\in\mathcal{S}, A\cap B=C} \mathbb{E}\left[I_A I_B\right] \le \lambda^{2k-|C|} N^{-k-1-l_{\texttt{int}}(C)} \binom{t-|C|}{k-|C|}\,. \tag{3}$$

C. Decomposition of Δ

The contribution of the path A to the sum Δ/μ^2, may be written as $\frac{1}{\mu^2}\sum_{B, A\cap B\neq\emptyset} \mathbb{E}\left[I_A I_B\right]$. Combining Eq.(1) and (3) we can show that is less than a function which is equivalent, as N grows, to

$$\frac{1}{N^{k-1}\binom{t}{k}} \sum_{C\subseteq A} \lambda^{-|C|} N^{-l_{\texttt{int}}(C)} \binom{t-|C|}{k-|C|} \binom{t}{k}^{-1}\,. \tag{4}$$

Note that k and t are function of N, although we did not write that explicitly to keep notation readable. The denominator before the sum is equivalent to the number of paths A in $\mathcal{S}$ for large N, and we observe as $\frac{k}{t} = \gamma$,

$$\binom{t-|C|}{k-|C|} \binom{t}{k}^{-1} = \frac{k}{t}\cdot\frac{k-1}{t-1}\cdots\frac{k-|C|}{t-|C|} \le (\gamma)^{|C|}\,.$$

hence, the following lemma is sufficient to conclude

Lemma 3: For any $A \in \mathcal{S}$, we have as N grows,

$$\limsup \sum_{C\subseteq A} (\frac{\gamma}{\lambda})^{|C|} N^{-l_{\texttt{int}}(C)} \le \frac{1}{(1-\frac{\gamma}{\lambda})^2} - 1$$

This lemma shows that a bound which features jointly the size of the intersection and the number of internal fragments may be a very solid tool to counterweight the combinatorial explosion of the number of subset of A.

Note that since A contains $k(N)$ contacts, for a fixed size c and large N there exists approximately $(k(N))^c/(c!)$ many subset C of a given size c, so that each term of this series diverge as N grows if one forgets the coefficient including the number of internal fragments. Accordingly, if one forgets the dependence on the size of c, as one can build an order $k(N)^2$ of subset C with no internal fragment at all, again the series diverge.

Proof: We treat separately the case where $l_{\texttt{int}}(C) = 0$ and $l_{\texttt{int}}(C) \ge 1$: If $l_{\texttt{int}}(C) = j \ge 1$, note first that there exists at most $(k(N))^{2j+2}$ subset which verifies this condition. This is because an intersection with j internal fragment may be characterized by the choice of 2_j integer in $k(N)$ (corresponding to the numbers a_m, b_m of internal fragments, and at most two others (in case of external fragment). Since the left coefficient is upper bounded by 1, we have:

$$\begin{aligned}\sum_{C\subseteq A, l_{\texttt{int}}(C)\ge 1} (\frac{\gamma}{\lambda})^{|C|} N^{-l_{\texttt{int}}(C)} &\le \sum_{j\ge 1} \frac{(\tau\gamma\ln(N))^{2j+2}}{N^j} \\ &\le (\tau\gamma\ln(N))^2 \frac{(\tau\gamma\ln(N))^2/N}{1-(\tau\gamma\ln(N))^2/N}\,.\end{aligned} \tag{5}$$

This term goes to zero as N becomes large.

Assuming now $l_{\texttt{int}}(C) = 0$, let us observe that there exists at most $j+1$ subset of A satisfying this condition and $|C| = j$, hence we have

$$\sum_{l_{\texttt{int}}(C)=0} (\frac{\gamma}{\lambda})^{|C|} \le \sum_{j\ge 1} (1+j)(\frac{\gamma}{\lambda})^j \quad = \tfrac{1}{(1-\gamma/\lambda)^2} - 1\,. \tag{6}$$

■

IV. Conclusion

We prove that random temporal network exhibits a phase transition phenomenon for the existence of solution to route over time between two arbitrary nodes: with a positive probability a path exist within logarithmic number of hops and time. The constants should satisfy a joint condition between length and time, which derive from the entropy function.

Our results demonstrate that random temporal networks feature specific combinatorial properties. We hope that this example will motivate other researcher to investigate the properties of large random structures that include time explicitly.

References

[1] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, J. Scott, and R. Gass, “Impact of human mobility on opportunistic forwarding algorithms,” *IEEE Trans. Mob. Comp.*, vol. 6, no. 6, pp. 606–620, 2007.

[2] A. Chaintreau, A. Mtibaa, L. Massoulié, and C. Diot, “The diameter of opportunistic mobile networks,” in *Proceedings of ACM CoNext*, 2007.

[3] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, “Diversity of forwarding paths in pocket switched networks,” in *Proceedings of ACM IMC*, 2007.

[4] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs*. John Wiley & Sons, 2000.