A Dynamic Random Graph Model for Diameter-Constrained Topologies in Networked Systems
Luigi Alfredo Grieco, Mahdi Ben Alaya, Thierry Monteil, Khalil Drira

To cite this version:

HAL Id: hal-01228346
https://hal.science/hal-01228346
Submitted on 16 Nov 2015

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.
Abstract—Random graphs have been widely investigated in literature because of their relevance to many scientific domains. In this brief, the attention is focused on diameter constrained random graphs, useful to analyze unstructured overlays for delay bounded network applications and systems. To this end, a general process of arrivals is considered to describe the sequence of vertex couples (i.e., node couples) among which a path composed of no more than \( D \) edges (i.e., links) has to be established. Accordingly, a topology formation mechanism \( M \) is formulated, expressing the rules that drive the addition of new edges, obeying to the constraint on the maximum diameter \( D \). Third, using graph theoretic arguments, an original discrete time model is proposed that describes the evolution of the average network degree (i.e., the number of edges per node) subject to \( M \) and \( D \). Fourth, the model is successfully validated using computer simulations in a wide range of scenarios (with up to \( 2^{16} \) nodes). Finally, concrete examples are provided to illustrate useful applications of the proposed approach, also in the presence of link failures.

Index Terms—Graph theory, Topology, Networks, Overlay.

I. INTRODUCTION

Graph-based models are fundamental tools to assess, predict, and control the performance of complex systems, made of interacting dynamic units. In these systems, vertice are usually associated to the dynamic units whereas edges represent interactions. The application domains of graph-based models include coupled biological/chemical systems, social networks, software applications, and communication protocols [1]–[3].

In many real systems, unfortunately, the properties of their interacting units cannot be deterministically known in advance. In these cases, random graphs [4] can be fruitfully used to infer the characteristics of the topology, based on the probabilistic behavior of vertice and edges (see also [5] for a comprehensive overview on the subject).

With reference to communication issues, random graphs have been mainly adopted to describe unstructured overlays [6]–[8], web properties [9], and Internet topology [10].

In this brief, we focus on diameter constrained overlays, i.e., virtual network topologies having a diameter no larger than a predefined threshold \( D \). This kind of overlay is very useful to support delay sensitive applications, such as in Peer-to-Peer (P2P) TV [11] and emerging Machine-to-Machine (M2M) systems [12]. In fact, the higher the diameter \( D \) the higher the end-to-end communication delay [13]. The problem of building diameter constrained graphs has been thoroughly afforded in [14] with reference to structured overlays, built upon distributed hash tables (DHT). Unfortunately, to the best of the authors’ knowledge, no theoretical contribution has been formulated yet, able to describe with closed form expressions the dynamics of an unstructured evolving overlay, subject to a constraint of the maximum diameter \( D \). This kind of model could be very useful to enable closed loop autonomic management strategies as well as to characterize, in a tractable form, both transient and steady state properties of network topologies in M2M scenarios [12], [15] and beyond.

Starting from this premise, a theoretical model based on random graph is formulated herein, which considers a discrete time process of arrivals to describe the sequence of vertex couples among which a path composed of no more than \( D \) edges has to be established. Accordingly, a general topology formation mechanism \( M \) is formulated, expressing the rules that drive the addition of new edges, obeying to the constraint on the maximum diameter \( D \). Then, exploiting the properties of the binary adjacency matrix \( A \) in graph theory [16], an original and tractable discrete time model is proposed that describes the evolution of the average network degree (i.e., the average number of edges per node) subject to \( M \) and \( D \). The model is successfully validated using computer simulations in a wide range of scenarios (with up to \( 2^{16} \) nodes). Finally, concrete examples are provided to illustrate useful applications of the proposed approach. They include: (i) the derivation of an approximated upper bound \( \sqrt{2 \cdot N \cdot \ln N} \) on graph average degree (i.e., the average number of edges per vertex); (ii) the comparison with respect to delay optimal de Bruijn graphs [14]; (iii) the analysis of the graph robustness; (iv) the derivation of system dynamics in presence of edge failures.

The rest of the brief is organized as follows: the main theoretical achievement is presented in Sec. II and validated in Sec. III. Useful examples of its applications are described in Sec. IV. The last Sec. V closes the brief and draws future research.

II. MODEL

A. Target Scenario and Notation

The target scenario considered in this brief consists of a graph of \( N \) vertice, \( n_q \) being the \( q \)-th vertix (\( q \in [1,N] \)). Furthermore, an ordered sequence of equi-probable \(^1\) couples

---

\(^1\)It is worth to note that equi-probable arrivals (i.e., homogeneous conditions) are usually assumed in the current literature dealing with diameter constrained graphs [14]. In fact, if the vertice of the graph represent the gateway through which the service of a large number of nodes are made available (as in M2M systems [17]), it is not unlikely that the overlay that inter-connect such gateways actually reflects this assumption.

---
of vertex is considered, among which a path composed of no more than \(D\) edges has to be established. The \(t\)-th couple is described by the vertex \((n_i, n_j)\). For sake of simplicity, the variable \(t\) will be referred to as time from now on. Knowing the \(t\)-th couple, a new edge is established in the graph if and only if the two vertex \((n_i, n_j)\) are not reciprocally reachable in no more than \(D\) edges. Knowing, the number of edges \(l_{t-1}\) at time \(t-1\), the probability that a couple of vertex at time \(t\) will not be reciprocally reachable in no more than \(D\) edges is defined as \(P_{t-1}\). Notice that, since we are assuming homogeneous conditions, \(P_{t-1}\) is the same for all the possible couples \((n_i, n_j)\).

Accordingly, our model is grounded on the following equation:

\[
l_{t+1} = l_t + P_t
\]

which, considering that the average degree \([4]\) (i.e., the number of edges per vertex) is \(k_t = \frac{2l_t}{N}\), can be also expressed as:

\[
k_{t+1} = k_t + \frac{2}{N}P_t
\]

The presence of an edge between any couple of vertex at time \(t\) will be expressed (as usual in graph theory) using the binary symmetric adjacency matrix \(A_t^{N \times N}\), so that \(A_t(i, j) = A_t(i, j) = 1\) if and only if an edge between \(n_i\) and \(n_j\) exists at time \(t\) (otherwise \(A_t(i, j) = A_t(i, j) = 0\)).

### TABLE 1

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>Number of vertex</td>
</tr>
<tr>
<td>(k_t)</td>
<td>Average degree at time (t)</td>
</tr>
<tr>
<td>(n_q)</td>
<td>(q)-th vertex</td>
</tr>
<tr>
<td>(D)</td>
<td>Maximum diameter</td>
</tr>
<tr>
<td>((n_i, n_j))</td>
<td>(t)-th couple of vertex wishing to establish a path</td>
</tr>
<tr>
<td>(l_t)</td>
<td>Number of edges at time (t)</td>
</tr>
<tr>
<td>(w_t)</td>
<td>Number of path with less than (D+1) edges between a couple of vertex at time (t)</td>
</tr>
<tr>
<td>(P_{t-1})</td>
<td>Probability that no path exists shorter than (D+1) edges between the vertex ((n_i, n_j))</td>
</tr>
<tr>
<td>(A_t^{N \times N})</td>
<td>Symmetric binary adjacency matrix at time (t)</td>
</tr>
<tr>
<td>(Pr{x})</td>
<td>Probability of event (x)</td>
</tr>
<tr>
<td>(\hat{x})</td>
<td>Upper bound on (x)</td>
</tr>
</tbody>
</table>

To provide an illustrative example of the networked system we are modeling as a random graph, Fig. 1 shows the evolution of a graph made of \(N = 5\) vertex and constrained by \(D = 2\) max path length. In this example, the ordered sequence of vertex \((n_i, n_j)\) is: \((n_2, n_4), (n_1, n_4), (n_1, n_2), (n_4, n_5), (n_2, n_3), \) and \((n_1, n_3)\). Accordingly, for each of them, a new edge is added if and only if a path shorter than 3 edges is not already available among corresponding vertice. In the sequel of the contribution, we will derive a law that rule the evolution of this kind of graphs in a general case.

**B. Main result**

**Proposition 1.** For a sufficiently large \(N\), the following expression describes the dynamics of the average graph degree:

\[
k_{t+1} = k_t + \frac{2}{N} \cdot \exp\left(-\frac{k_t}{\hat{x}}\right)
\]

**Proof.** The model considered here is based on Eq. (1), or equivalently on finding an accurate approximation for the probability \(P_t\). The latter expresses the probability to find a path at time \(t\) (no longer than \(D\) edges) between a generic couple of vertex \((n_i, n_j)\), knowing that the number of already existing edges is \(l_t\).

To fulfill this objective, we first leverage a well known property of the matrix \(A_t\): \(A_t^{c}(i, j) = 0\), \(c \in N^+\), if and only if no path composed of \(c\) edges exist between \(n_i\) and \(n_j\) at time \(t\) [16].

In this way, without lack of generality, \(P_t\) can be expressed as follows:

\[
P_t = \prod_{c=1}^{D} Pr\{A_t^c(i_{t+1}, j_{t+1}) = 0\}
\]

Now, given that \(2 \cdot l_t\) elements are equal to one in \(A_t\), it yields \(Pr\{A_t(j, j) = 0\} = \frac{2l_t}{N^2}\). Also, since any element of \(A_t^c\) is no other than the sum of \(N^c-1\) elements, each one being a product of \(c\) coefficients belonging to \(A_t\), we can approximately write:

\[
Pr\{A_t^c(i_{t+1}, j_{t+1}) = 0\} \approx \left[ 1 - \left(\frac{2 \cdot l_t}{N^2}\right)^c \right]^{N^c-1}
\]
Now, recalling that \((1 + \frac{1}{x})^x \to e\), when \(x \to \infty\), for a sufficiently large \(N\), Eq. (5) can be written as:

\[
Pr \{ A_t^k(i_{t+1}, j_{t+1}) = 0 \} \approx \exp \left( -N^{c-1} \cdot \left( \frac{2 \cdot l_t}{N^2} \right)^c \right) \tag{6}
\]

Accordingly, by substituting (6) in (4), it is obtained:

\[
P_t \approx \exp \left( -\sum_{c=1}^{D} N^{c-1} \cdot \left( \frac{2 \cdot l_t}{N^2} \right)^c \right) \tag{7}
\]

which, after a little algebra, becomes

\[
P_t \approx \exp \left( -\frac{1}{N} \cdot \left( \frac{2 \cdot l_t}{N^2} \right)^D \cdot \frac{1}{2N^2} + \frac{2 \cdot l_t}{N} \right) \tag{8}
\]

From [4], the average degree can be expressed as \(k_t = \frac{2l_t}{N}\), so that we obtain the proof by substituting (8) in (2).

\[\square\]

III. NUMERICAL VALIDATION

To validate the model (3), we have considered a complex scenario composed of \(N\) vertices (with \(N\) ranging from \(2^5\) to \(2^{16}\)) and \(D\) ranging from 3 to 10. Using an ad hoc simulator we developed in Matlab, the relative error between the real evolution of the degree \(k(t)\) and the one estimated using (3) for all \(t\) is evaluated. In any case, we found that the average relative error is below 10\% if we consider the entire evolution of \(k(t)\). Also, the relative error at steady state (once the graph is completely formed), for \(D \leq 5\), is below 10\%, whatever \(N\). Finally, we notice a slight increase in the relative error as \(D\) increases; (in any case) it remains smaller than 25\% and it falls below 20\% for \(N > 2^{12}\).

To provide an illustrative example, Fig. 2 plots the dynamic evolution of the number of edges versus the time \(t\) (similar results have been obtained for different values of \(N\) and \(D\) so that the average relative error is below 10\% in all cases).

![Fig. 2. Evolution of the number of edges over the time \(t\), \((N = 1000, D = 5)\).](image)

It is worth to note that the number of edges linearly increases with \(t\) till a saturation point is reached. From that moment on, \(l\) exhibits a very slow rise. This can be explained by plotting also the values of the probability \(P_t\). Fig. 3 shows that \(P_t\) is almost equal to one for some time during the network formation, meaning that, since the number of edges is low, it is highly likely to add a new edge as soon as a new couple of nodes needs to establish a path. At the same time, the values of \(P_t\) abruptly decreases after a certain \(t\), meaning that the topology reached the steady state.

![Fig. 3. Evolution of \(P_t\), \((N = 1000, D = 5)\).](image)

To provide a further insight, Fig. 4 pictures the evolution of a random graph \((N = 1000, D = 5)\): (a) at the beginning of the simulation, when no edge is present; (b) during the transient, when a few edges have been created and a new edge is added; and (c) at steady state, when all required paths (no longer than \(D\) edges) have been already created.

IV. EXAMPLE APPLICATIONS

A. Rank bound and convergence

**Theorem 1.** Being \(\hat{k}\) the maximum degree \(k\) in system (3), it can be bounded as follows:

\[
\hat{k} \leq \sqrt{2N \cdot \ln N}. \tag{9}
\]
Proof. The way we are building the overlay is so that an edge between a couple of vertex \((n_i, n_j)\) is formed or not based only upon the first time that couple issues a request for a path. If a path shorter than \(D + 1\) hops already exists the edge is not established otherwise it is established. From that moment on, the next requests for a path issued by the same couple of vertex will not sort any effect.

Based on this consideration, we extract from the sequence of equi-probable couples of vertex considered in the brief, the sequence of instants in which any couple of vertex appears for the first time. Of course, the length of such a sequence of time instants will be composed of at most \(N(N - 1)/2\) elements.

In order to estimate an upper bound on the steady state average degree \(\hat{k}\), it is necessary to consider that at time \(t\), \(1/P_t\) expresses the average time required to establish the next edge in the overlay.

Under this assumption, the expression of \(P_t\) to consider is slightly different from that in Eq. (8) because the sequence of vertix couples we are considering to prove this Theorem is chosen in such a way that no one edge path exists at time \(t\) for \((n_i, n_j)\). The avoid ambiguity, we will refer to \(P_t'\) to refer to this new probability. Accordingly, it follows:

\[
P_t' = \prod_{c=2}^{D} Pr\{A_t'(i_{t+1}, j_{t+1}) = 0\} \tag{10}
\]

which, following the same passages reported in the proof of Proposition 1, can be also written as:

\[
P_t' \approx \exp \left( -\frac{1}{N} \cdot \left( \frac{2\cdot l_N}{N} \right)^{D+1} - \left( \frac{2\cdot l_N}{N} \right)^2 \right) \tag{11}
\]

Therefore to estimate an approximated upper bound on \(\hat{k}\), it is sufficient to find any value of \(k_t\) so that \(\sum k_t^D \geq N(N - 1)/2\), which gives \(\sum_{c=2}^{D} k_t^c \geq 2 \cdot N \cdot \ln N - N \cdot \ln 2\). Notice that the latter inequality is satisfied if \(k_t^D \geq 2 \cdot N \cdot \ln N\), from which the proof follows.

Remark 1. It is worth remarking that, the bound \(\hat{k}\) can be fruitfully exploited to compare the properties of the graph under study with respect to well known graph. To this end, if we consider as ground for comparison the delay optimal de Bruijn graphs, representing one of the most compact topologies discovered so far [14], we will find that, for the same diameter \(D\), the degree of the unstructured overlay considered in this brief is only \(\sqrt{2} \cdot \ln N\) times larger at most, even if it is based on a much simpler construction mechanism.

B. Robustness

Knowing that the graph construction model adopted herein ensures, at steady state, at least one path shorter than \(D + 1\) edges between any couple of nodes, it is worth investigating how many paths (composed of less than \(D + 1\) edges) are present between any couple of nodes. This metric is intimately related to the topology robustness: the higher the number of
paths the higher the number of alternative solutions to route messages in case of failures.

**Theorem 2.** Defined as \( w_t \) the number of paths composed of less than \( D + 1 \) edges between any couple of nodes at time \( t \), it holds:

\[
w_t = \frac{1}{N} \cdot \left( k_i^{D+1} - k_i \right) \quad (12)
\]

**Proof.** Also in this case the properties of the adjacency matrix are exploited. In particular, \( A^c_i(j, j) \) indicates the number of paths composed of \( c \) edges between the vertex \( n_i \) and \( n_j \). Thus, considering that the average value of any element if \( A \) is \( \frac{2l_i}{N^2} \), it holds:

\[
w_t = \frac{1}{N} \cdot \sum_{s=1}^{D} k_t^s = \frac{1}{N} \cdot k_t^{D+1} - k_t \quad (13)
\]

This ends the proof.

Based of this Theorem, we can derive an approximate assessment of the level of redundancy, if we consider for \( k_t \), the bound derived in Theorem 2, i.e., \( k_t = \sqrt{2 \cdot N \cdot \ln N} \). Under this assumption, we can obtain:

\[
\dot{w} = \frac{1}{N} \cdot \left( \frac{(2 \cdot N \cdot \ln N) \cdot \lambda_o}{2 \cdot N \cdot \ln N - 1} - 2 \cdot N \cdot \ln N \right) > 2 \cdot \ln N \quad (14)
\]

This result indicates that the overlay investigated in this brief is able to provide at steady state at least two paths among any couple of vertices, for \( N > e \).

**C. Link failures**

To include also possible link failures and dynamics in the model, it is necessary to modify Eq. (1) as follows:

\[
l_{t+1} = l_t + P_t - \lambda_o \cdot l_t \quad (15)
\]

where \( \lambda_o \) is the probability that an edge is removed during one time step. The resulting equations could be very useful to design topology management algorithms using control theoretic arguments. Its utility in finding the uniqueness of the equilibrium point is shown in the following Theorem.

**Theorem 3.** The system (15) admits one and only one equilibrium point \( l = l_\infty \).

**Proof.** In order to find the equilibrium point of system (15), we impose \( l_{t+1} = l_t = l_\infty \) in (15). Accordingly, the following equality is obtained:

\[
P_\infty = \lambda_o \cdot l_\infty \quad (16)
\]

Eq. (16) admits only one solution because its leftmost member monotonically decreases with \( l \), starting from the value one at \( l = 0 \) whereas the rightmost member monotonically increases, starting from zero at \( l = 0 \).

This ends the proof.

\[\square\]

V. CONCLUSION

A novel tractable model for describing the dynamics of a diameter constrained random graph is proposed, validated, and analyzed in this brief. Useful examples of its adoption have been also provided in order to demonstrate its real utility. Future research will encompass: (i) the evaluation of gravity based features and heterogeneous conditions; (ii) the characterization of M2M overlays; (iii) the study of the properties of the equilibrium point found in Theorem 3; and (iv) the formulation of a control theoretic framework for overlay topologies built upon the model proposed herein.

ACKNOWLEDGMENTS

This work was supported by the INSA of Toulouse (FR).

REFERENCES