A Dynamic Random Graph Model for Diameter-Constrained Topologies in Networked Systems

Luigi Alfredo Grieco, Mahdi Ben Alaya, Thierry Monteil, Khalil Drira

A Dynamic Model for Diameter Constrained Networked Systems

Luigi Alfredo Grieco, *Senior Member, IEEE,* Mahdi Ben Alaya, Thierry Monteil, and Khalil Drira

***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 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 $\sqrt[D]{2 \cdot N \cdot lnN}$ 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.

L. A. Grieco is with the Department of Electrical and Information Engineering, Politecnico di Bari, Bari, Italy, e-mail: a.grieco@poliba.it

M. B. Alaya and T. Monteil are with CNRS, LAAS, and Univ. de Toulouse, Toulouse, France. K. Drira is with CNRS, LAAS, Toulouse, France. e-mail: {ben.alaya, monteil, khalil}@laas.fr

of vertice 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 vertice (n_{i_t}, n_{j_t}). 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 vertice (n_{i_t}, n_{j_t}) 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 vertice 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_t}, n_{j_t}).

Accordingly, our model is grounded on the following equation:

$$l_{t+1} = l_t + P_t \tag{1}$$

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

$$k_{t+1} = k_t + \frac{2}{N} P_t \tag{2}$$

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

TABLE I
NOTATION.

Symbol	Meaning
N	Number of vertice
k_t	Average degree at time t
n_q	q-th vertex
D	Maximum diameter
(n_{i_t}, n_{j_t})	t-th couple of vertice wishing to establish a path
l_t	Number of edges at time t
w_t	Number of path with less than $D+1$ edges between a couple of vertice at time t
P_{t-1}	Probability that no path exists shorter than $D+1$ edges between the vertice (n_{i_t}, n_{j_t})
A_t^{NxN}	Symmetric binary adjacency matrix at time t
$Pr\{x\}$	Probability of event x
$\hat{x}$	Upper bound on x

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 $N = 5$ vertice and constrained by $D = 2$ max path length. In this example, the ordered sequence of vertice (n_{i_t}, n_{j_t}) 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.

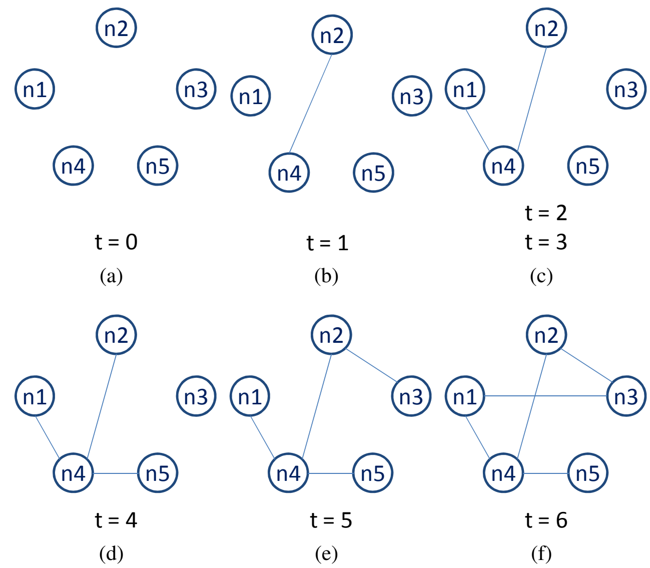


Fig. 1. Evolution of a random graph ($N = 5$, $D = 2$): a) initial state; b) the first edge is added at $t = 1$ to connect vertice n_2 and n_4; c) the second edge is added at $t = 2$ to connect vertice n_1 and n_4, but no edge is added at $t = 3$ because the vertice n_1 and n_2 (asking for a connection) are already connected by a path of 2 hops $\leq D$; d) the third edge is added at $t = 4$ to connect vertice n_4 and n_5; e) the fourth edge is added at $t = 5$ to connect vertice n_2 and n_3; e) the fifth edge is added at $t = 6$ to connect vertice n_1 and n_3, for which the only existing path was longer than D hops.

B. Main result

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

$$k_{t+1} \approx k_t + \frac{2}{N} \cdot exp\left(-\frac{1}{N} \cdot \frac{k_t^{D+1} - k_t}{k_t - 1}\right) \tag{3}$$

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 vertice $(n_{i_{t+1}}, n_{j_{t+1}})$, 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:

$$P_t = \prod_{c=1}^{D} Pr\{A_t^c(i_{t+1}, j_{t+1}) = 0\} \tag{4}$$

Now, given that $2 \cdot l_t$ elements are equal to one in A_t, it yields $Pr\{A_t(i,j) = 1\} = \frac{2 \cdot l_t}{N^2}$. 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:

$$Pr\{A_t^c(i_{t+1}, j_{t+1}) = 0\} \approx \left[1 - \left(\frac{2 \cdot l_t}{N^2}\right)^c\right]^{N^{c-1}} \tag{5}$$

Now, recalling that $(1+\frac{1}{x})^x \rightarrow e$, when $x \rightarrow \infty$, for a sufficiently large N, Eq. (5) can be written as:

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

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

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

which, after a little algebra, becomes

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

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

□

III. Numerical Validation

To validate the model (3), we have considered a complex scenario composed of N vertice (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).

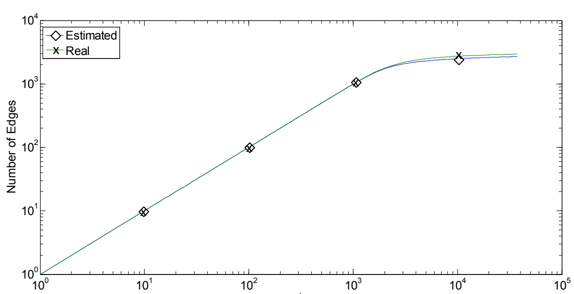


Fig. 2. Evolution of the number of edges over the time t, ($N = 1000$, $D = 5$).

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

Algorithm 1 Validation Code

procedure Simulate Graph Formation(N,D,$stop$) ▷ $stop$: stop condition
  $l(0) \leftarrow 0$ ▷ Initial number of edges
  $est_l(0) \leftarrow 0$ ▷ Estimated initial number of edges
  $counter \leftarrow 0$ ▷ Counter initial value
  $A \leftarrow 0^{NxN}$ ▷ Adjacency matrix initial value
  while !$stop$ **do** ▷ The procedure is run until the $stop$ condition is verified
    $i \leftarrow randi(N)$ ▷ An integer random number between 1 and N is assigned to i
    $j \leftarrow randi(N)$
    if $i \neq j$ **then**
      $counter \leftarrow counter + 1$
      $est_l(counter) \leftarrow est_l(counter) + P_{counter}$
      $d \leftarrow shortestpath(A, i, j)$ ▷ d is the length of the shortest path from n_i to n_j
      if $d < D + 1$ **then**
        $A(i, j) \leftarrow 1;$
        $l(counter) \leftarrow l(counter - 1) + 1$
      else
        $l(counter) \leftarrow l(counter - 1)$
      end if
    end if
  end while
end procedure

of P_t abruptly decreases after a certain t, meaning that the topology reached the steady state.

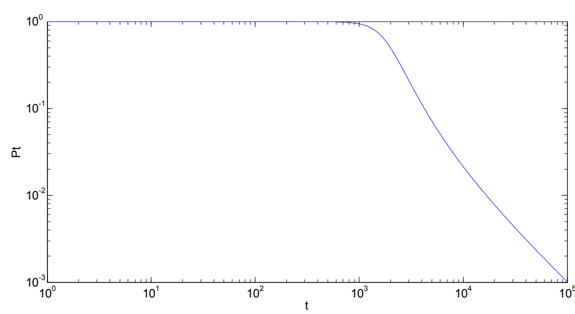


Fig. 3. Evolution of P_t, ($N = 1000$, $D = 5$).

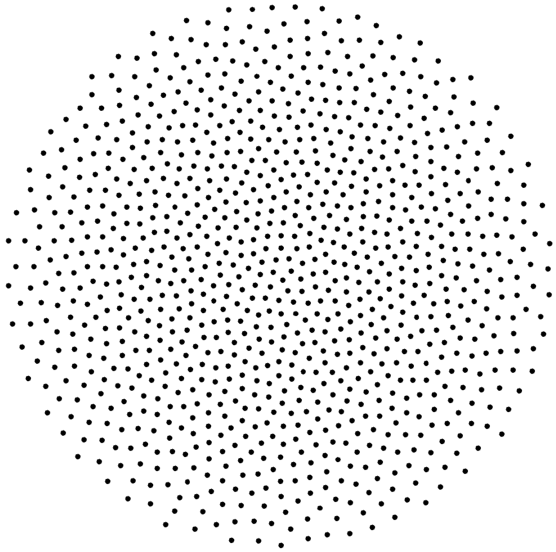
To provide a further insight, Fig. 4 pictures the evolution of a random graph ($N = 1000$, $D = 5$): (a) at the beginning of the simulation, when no edge is present; (b) during the transient, when a few edges have been created and a new edge is added; and (c) at steady state, when all required paths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

IV. Example Applications

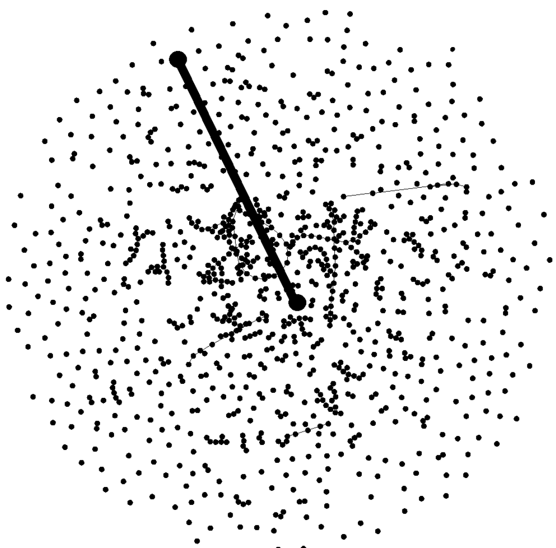
A. Rank bound and convergence

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

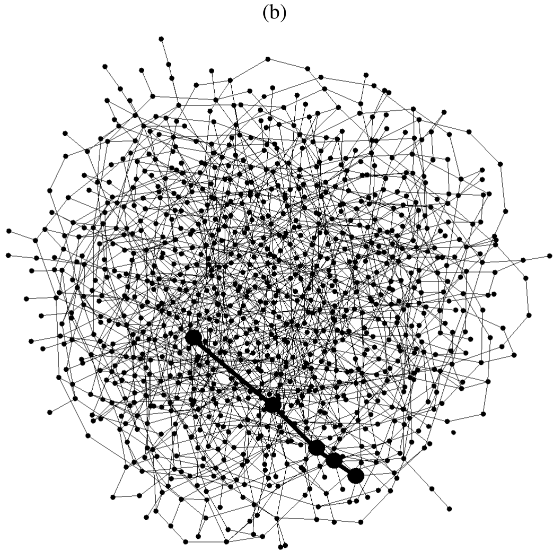
$$\hat{k} \leq \sqrt[D]{2 \cdot N \cdot lnN}. \quad (9)$$



(a)



(b)



(c)

Fig. 4. Evolution of a random graph ($N = 1000$, $D = 5$): a) initial state; b) during the transient a new edge is added; c) at steady state no edge is added because paths no longer than D hops already exist.

Proof. The way we are building the overlay is so that an edge between a couple of vertice (n_{i_t}, n_{j_t}) 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 vertice will not sort any effect.

Based on this consideration, we extract from the sequence of equi-probable couples of vertice considered in the brief, the sequence of instants in which any couple of vertice 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 proof this Theorem is chosen in such a way that no one edge path exists at time t for (n_{i_t}, n_{j_t}). The avoid ambiguity, we will refer to P'_t to refer to this new probability. Accordingly, it follows:

$$P'_t = \prod_{c=2}^{D} Pr\{A_t^c(i_{t+1}, j_{t+1}) = 0\} \tag{10}$$

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

$$P'_t \approx exp\left(-\frac{1}{N} \cdot \frac{(\frac{2 \cdot l_t}{N})^{D+1} - (\frac{2 \cdot l_t}{N})^2}{\frac{2 \cdot l_t}{N} - 1}\right) \tag{11}$$

Therefore to estimate an approximated upper bound on $\hat{k}$, it is sufficient to find any value of k_t so that $\frac{1}{P'_t} \geq N(N-1)/2$, which gives $\sum_{D}^{c=2} k_t^c \geq 2 \cdot N \cdot lnN - N \cdot ln2$. Notice that the latter inequality is satisfied if $k_t^D \geq 2 \cdot N \cdot lnN$, 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 $\sqrt[D]{2 \cdot lnN}$ 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:*

$$w_t = \frac{1}{N} \cdot \frac{k_t^{D+1} - k_t}{k_t - 1} \tag{12}$$

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

$$w_t = \frac{1}{N} \cdot \sum_{s=1}^{D} k_t^s = \frac{1}{N} \cdot \frac{k_t^{D+1} - k_t}{k_t - 1} \tag{13}$$

This end 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., $k_t = \sqrt[D]{2 \cdot N \cdot lnN}$. Under this assumption, we can obtain:

$$\hat{w} = \frac{1}{N} \cdot \frac{(2 \cdot N \cdot lnN)^{\frac{D+1}{D}} - 2 \cdot N \cdot lnN}{2 \cdot N \cdot lnN - 1} > 2 \cdot lnN \tag{14}$$

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 vertice, 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:

$$l_{t+1} = l_t + P_t - \lambda_o \cdot l_t \tag{15}$$

where λ_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 uniquness 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:

$$P_\infty = \lambda_o \cdot l_\infty \tag{16}$$

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. □

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

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwanga, “Complex networks: Structure and dynamics,” *PHYSICS REPORTS*, vol. 424, no. 4-5, pp. 175–308, 2006.

[2] M. Babaei, H. Ghassemieh, and M. Jalili, “Cascading failure tolerance of modular small-world networks,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 58, no. 8, pp. 527–531, Aug 2011.

[3] Y. Chen, W. Yu, F. Li, and S. Feng, “Synchronization of complex networks with impulsive control and disconnected topology,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 60, no. 5, pp. 292–296, May 2013.

[4] P. Erdos and A. Renyi, “On the evolution of random graphs,” *Mat Kutato Int. Kozl*, vol. 5, no. 17, pp. 17–60, 1960.

[5] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, “Random graphs with arbitrary degree distributions and their applications,” *PHYSICAL REVIEW E*, vol. 64, 2001.

[6] L. Massoulie, A.-M. Kermarrec, and A. J. Ganesh, “Network awareness and failure resilience in self-organising overlay networks,” in *22nd IEEE Int. Symposium on Reliable Distributed Systems*, 2003.

[7] T. Bonald, L. Massoulie, F. Mathieu, D. Perino, and A. Twigg, “Epidemic live streaming: Optimal performance trade-offs,” in *SIGMETRICS*. ACM, 2008.

[8] R. J. Lobb, A. P. C. da Silva, E. Leonardi, M. Mellia, and M. Meo, “Adaptive overlay topology for mesh-based P2P-TV systems,” in *NOSSDAV*. ACM, 2009.

[9] W. Aiello, F. Chung, and L. Lu, “A random graph model for massive graphs,” in *32nd ACM Symposium on Theory of Computing*. ACM, 2000.

[10] H. Reittu and I. Norros, “On the power-law random graph model of massive data networks,” *Performance Evaluation*, vol. 55, pp. 3–23, 2004.

[11] P. Veglia and D. Rossi, “Performance evaluation of P2P-TV diffusion algotihms under realistic settings,” *Peer-to-Peer Networking and Application*, 2013.

[12] L. A. Grieco, M. B. Alaya, T. Monteil, and K. Drira, “Architecting information centric ETSI-M2M systems,” in *Proc. of IEEE PerCom*, 2014.

[13] J. Kurose and K. Ross, *Computer Networking: A Top-Down Approach*. Pearson Education, Limited, 2012.

[14] D. Loguinov, J. Casas, and X. Wang, “Graph-theoretic analysis of structured peer-to-peer systems: Routing distances and fault resilience,” *IEEE Transactions on Networking*, vol. 13, no. 5, 2005.

[15] M. B. Alaya, S. Matoussi, T. Monteil, and K. Drira, “Autonomic computing system for self-management of machine-to-machine networks,” in *Proceedings of the 2012 international workshop on Self-aware internet of things*, ser. Self-IoT ’12. ACM, 2012, pp. 25–30.

[16] S. Noel and S. Jajodia, “Understanding complex network attack graphs through clustered adjacency matrices,” in *21st IEEE Annual Computer Security Applications Conference*, 2005.

[17] A. Dhraief, A. Belghith, K. Drira, T. Bouali, and M. A. Ghorbali, “Autonomic management of the HIP-based M2M overlay network,” *Procedia Computer Science*, vol. 19, no. 0, pp. 98 – 105, 2013.