9

Techniques for Analyzing Random Graph Dynamics and Their Applications

Ali Hamlili
ENSIAS, Mohamed V–Souissi University, Rabat, Morocco

1. Introduction

Graph theory is birth in 1736 with the publication of the work of the Swiss mathematician Leonhard Euler on the problem of finding a round trip path that would cross all the seven bridges of the city of Königsberg exactly once (Euler, 1736). Since then, this theory has known many important developments and has answered to a lot of practical issues. Today, the graph theory is considered as an essential component of discrete mathematics. It aims at analyzing the structure induced by interactions between a set of elements and to study the resulting fundamental properties. Graph theory occurs as a fundamental and theoretical framework for analyzing a wide range of the so-called *real-world networks* in biology, computer sciences, multi-agent systems, chemistry, physics, economy, knowledge management, and sociology. In many works, graph models are employed as constructive descriptions to represent and understand the behavior of different complexe systems (Molloy and Reed, 1998; Mieghem et al., 2000; Newman, 2003; Kawahigashi et al. 2005; Jurdak, 2007). In such models, the graph vertices stand for the components (nodes) of the network that encode information about the values of the state variables of the dynamical system and the edges represent the mutual relationships between the correspondent end-nodes. In practice, random graph theory has become increasingly important for modeling networks whose behaviors exhibit nondeterministic looks. In recent years, many significant results have used random graph models to explain, replicate and simulate the behavior of dynamic real-world networks (Hekmat and Van Mieghem, 2003; Kawahigashi et al., 2005; Durrett, 2006; Onat et al., 2008; Hewer et al., 2009; Hamlili, 2010; Trullols et al., 2010).

To provide a convenient way to represent and analyze dynamic networks by dynamic random graphs, it is very important to clarify how the model of random graphs should explain the behavior of change in the topology of the network. Thus, we introduce some stochastic processes (times of graph change, graph configurations, degree number at a chosen vertex …) in order to attempt to account for the observed statistical properties in graph dynamics. Therefore, we will try to highlight the basic mathematical operations that transform a graph into other one to make possible describing the dynamic change of graph configurations. In this objective, different concepts and notation are introduced in the preliminary sections and will be used throughout the chapter. A reader familiar with the common topics in general graph theory may skip ahead. However, he may use it as necessary to refer to unknown definitions or unusual notations. Also, a particular attention is agreed to Erdös-Rényi's random graph model (Erdös and Rényi, 1960; Bollobàs, 2001).

2. Preliminary concepts

This section is a short introduction on graph theory. It will review the basic definitions and notation used throughout all this chapter.

2.1 Graphs

A classical graph is a static structure of a set of objects where some pairs of these objects are connected by one or several directed or undirected links. In this chapter, we assume that there is no multiple links between a pair of objects and the orientation of the links doesn't play a decisive role.

Definition 1

An undirected simple graph or simply a graph G can be defined as a pair $G=(V,E)$ *of two sets: a nonempty set V of elements called vertices, and a set* $E\subset\left\{(u,v)\in V^2 \,/\, v\neq u\right\}$ *of unordered pairs of vertices. The elements of E are called edges.*

Since the graph is undirected, (u,v) and (v,u) designate the same edge which we write simply uv. Furthermore, the assumption "simple" states the fact that between two given vertices, we cannot pass more than a single edge.

Definition 2

If $|S|$ *denotes the cardinality of a set S, the number of vertices* $N=|V|$ *and the number of edges* $M=|E|$ *of a graph* $G=(V,E)$ *define respectively the order and the size of this graph.*

Furthermore, we assume in this chapter that graphs can be finite or infinite according to their order and such that the sets of vertices and edges can't be jointly or separately empty.

Definition 2

Let $G=(V,E)$ *and* $G^*=\left(V^*,E^*\right)$ *be two graphs. We say that* G^* *is a host for G or equivalently G is a subgraph of G*, if and only if* $V\subset V^*$ *and* $E\subseteq E^*$.

Definition 3

A graph $G=(V,E)$ *is called a weighted graph if and only if a positive function (or weights) can be defined on the set of edges E.*

Depending on the underlying area of application, such weights might represent probabilities, costs, lengths, capacities, or other positive quantities having a particular meaning. In the general weighted graph version, both vertices and edges can be weighted.

2.2 Neighbors, neighborhood and connectivity

Routing problems are among the oldest problems in graph theory. They are generally based on the hypothesis of connectivity. Let us note that the concepts of neighborhood and path are the most typical ideas associated to the connectivity assumption.

Definition 4

A neighbor of a vertex u is any vertex v such that $uv \in E$.

We note $N_G(u)$ the set of direct neighbors of *u* (also called the *neighborhood* of *u*). e.g.

$$N_G(u) = \{ v \in V \mid uv \in E \} \tag{1}$$

Inversely, two vertices *u* and *v* of a graph *G* are said to be *adjacent*, if $v \in N_G(u)$. The *closed neighborhood* of *u* is denoted $\overline{N}_G(u) = N_G(u) \cup \{u\}$ and for a set $S \subset V$, the closed neighborhood of *S* can be defined as $\bar{N}_G(S) = \bigcup_{u \in S} \bar{N}_G(u)$. By analogy, two edges are called neighbors if they have an end-vertex in common. In addition, pairwise non-adjacent vertices or edges are called independent. If all vertices of a subset $S \subseteq V$ are pairwise adjacent, *S* is called a complete subset or a clique.

Definition 5

On a given graph $G = (V, E)$ *we can define a path* $\Pi_{u,v}$ *between a pair of vertices u and v if and only if there is a sequence of vertices (or walk)* $u_i = u, u_{i+1}, \cdots, u_{j-1}, u_j = v$ *such that*

$$u_k u_{k+1} \in E, \forall k = i..j-1$$

where the vertices u and v are called end-vertices of the path.

An *elementary path* is a path such that when all the vertices are sequentially visited, a same vertex is never met twice. A path such that the end-vertices coincide is called *cycle*. An *elementary cycle* is a cycle such that all the vertices have exactly two neighbors. The concept of path is behind the notion of connectivity. In the rest of this chapter, we note $\mathcal{P}_{u,v}$ the set of all paths between the vertices *u* and *v*.

Definitions 6

Let G be a simple graph,

i. *A pair of vertices* (u,v) *of G is called connected if G contains a path connecting u to v. Otherwise, they are called disconnected.*
ii. *A graph G is called connected if any pair of vertices of G is connected. Otherwise, it is called non-connected.*

To achieve a *fully connected* graph *G*, there must exist a path from any vertex to each other vertex in the graph.

The path between the source vertex and the destination vertex may consist of one hop when source and destination are neighbors or several hops if they aren't directly connected by an edge of *G*. The *hopcount* specifies the number of hops through a path between two vertices. This measure is meaningful only when there is a path between the source and the destination. The *average hopcount* of a graph is the average value of the hopcount between the end-vertices of all the possible paths.

Furthermore, in a non-connected graph there is no path between at least one source-destination pair of vertices. Hence, a non-connected graph consists of several disconnected clusters and/or vertices. Thus, routing is only possible between the different vertices of a same cluster.

Definitions 7

Consider a weighted graph G, (u,v) *a pair of vertices of G and let w be the weight function defined on G, the weight assigned to a path* $\Pi_{u,v}$ *can be computed as the sum of weights assigned to its edges. i.e.*

$$\Pi_{u,v} = \langle u_i = u, u_{i+1}, \cdots, u_{j-1}, u_j = v \rangle \Rightarrow w(\Pi_{u,v}) = \sum_{k=i}^{j\_1} w(u_k u_{k+1}) \tag{2}$$

2.3 Matrix representation of graphs and degree function of vertices

The topological structure of the graph $G = (V,E)$ can alternatively be described by a $|V| \times |V|$ adjacency matrix $A_G = (a_{u,v})_{u,v \in V}$ such that each entry is either 0 or 1

$$a_{u,v} = \begin{cases} 1 & \text{if} \quad v \in N(u) \\ 0 & \text{else} \end{cases} \tag{3}$$

where $a_{u,v} = 1$ signifies that *uv* is an edge of *G*. i.e. $uv \in E$.

Definition 8

The degree of a vertex u is the number of its direct neighbors

$$d_G(u) = |N_G(u)| \tag{4}$$

Proposition 1

Consider an undirected graph $G(V,E)$ *and* $A_G = (a_{u,v})_{u,v \in V}$ *its adjacency matrix, then*

$$d_G(u) = \sum_{v \in V} a_{u,v} \tag{5}$$

The two last equations (4) and (5) are equivalent by definition of the matrix A_G. They induce a function d_G from *V* to **N** (the set of nonnegative integers) called the *degree function*. Particularly, a vertex of degree 0 is called an *isolated* vertex and a vertex of degree 1 is called a *leaf*.

3. Random graphs

Another theory of graphs began in the late 1950s. It was baptized *random graph theory* in several papers by Paul Erdös and Alfréd Rényi. As a real-world network model, the Erdös-

Rényi's random graph model has a number of attractive properties (Bollobàs, 2001). This model is exceptionally quantifiable; it allows an easy calculation of average values of the graph characteristics (Janson et al., 2000; Hamlili, 2010).

In this section, we want introduce in first a generalization of the concepts of the theory of random graphs. This generalization is intended to describe the issues in applicative frameworks where the number of the graph vertices can vary randomly (number of communicating entities in a wireless ad hoc network, number of routers in the Internet, etc). Also, we will show that most classical models, such as those of Erdös-Rényi random graphs and geometric random graphs can be derived as special cases of the model that we put forward as a generalized alternative.

3.1 Generalized random graph model

Intuitively, a generalized random graph representation can be defined in a simple way using the fully weighted graphs.

Definition 9

Consider a non-empty set Ω, *called the set of possible vertices and* $\mathbf{P}=\left(p_{u,v}\right)_{u,v\in\Omega}$ *a symmetric* $|\Omega|\times|\Omega|$ *matrix of probabilities. We call generalized random graph a graph G where each vertex u is generated with the probability* $p_{u,u}\in\left]0,1\right[$, *and where for all two existing vertices u and v, the edge uv is built with a probability* $p_{u,v}\in\left]0,1\right[$.

Let $\mathcal{G}_{\mathbf{P}}(\Omega)$ be the collection of all possible graphs made on the set of possible vertices Ω such that the graph vertices *u* are generated independently with the probabilities $p_{u,u}$ and edges *uv* are built independently in $\Omega\times\Omega$ with the respective probabilities $p_{u,v}$. i.e.

$$\mathcal{G}_{\mathbf{P}}(\Omega)=\left\{\ G\ \middle|\ \forall\, u\in\Omega:\Pr[u\in V]=p_{u,u}\wedge\forall\, u,v\in\Omega:v\neq u\Rightarrow\Pr[uv\in E]=p_{u,v}\right\}$$

This definition of random graph models is very general. We should note that, if *G* is a generalized random graph,

$$G=(V,E)\Rightarrow V\subseteq\Omega\wedge E\subseteq V\times V \tag{6}$$

As will be discussed later, this way of modeling a random graph will represent opportunities for characterizing complex situations where classical models such as Erdös-Rényi model are not satisfactory.

Definition 10

The extended adjacency matrix $A=\left(a_{u,v}\right)_{u,v\in\Omega}$ *associated to a graph* $G=(V,E)$ *of the model* $\mathcal{G}_{\mathbf{P}}(\Omega)$ *is a* $|\Omega|\times|\Omega|$ *matrix such that* $a_{u,v}$ *is independent equal to 1 if the pair of vertices uv belongs to E knowing that u and v belong to V and 0 otherwise.*

3.2 Practical examples

Different particular cases can be identified and as stated above, in different contexts it may be useful to define the term random graph with different degrees of generality. Hence, the generalized model can describe random geometric graphs (Steele, 1997; Barabasi and Albert, 1999; Penrose, 1999). It suffices to consider *G* such that $V = \Omega$ and the edge probabilities

$$p_{u,v} = p\left(\|u - v\|\right) = \begin{cases} 1 & \text{if } 0 < \|u - v\| \le R \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

In this model, $p_{u,v}$ depends on the Euclidean distance $\|u - v\|$ between the geometric points locating the vertices *u* and *v*.

$$\mathcal{G}_R(V) = \left\{ G(V,E) \,\middle|\, E \subseteq V \times V \wedge \forall\, u,v \in V :\ \Pr(vu \in E) = 1 \text{ if } \|u - v\| \le R \text{ and } \Pr(vu \in E) = 0 \text{ if } \|u - v\| > R \right\}$$

This model is very interesting and as such it can formalize the framework of mobile wireless ad hoc networks where the connectivity of the network depends on the geometric positions of the communicating nodes and a radio coverage range which is generally supposed the same for all the network nodes. In such networks, the random dynamics of the associated graph is induced by the mobility of nodes.

Another example is the random graph model initiated by Erdös and Rényi in the 1950s. This kind of graphs can be represented by a generalized model where the set of vertices is not random (it is constantly the same $V = \Omega$) and all the graph pairs of vertices are connected with the same probability $p_{u,v} = p$. Thus, let *V* a nonempty set of vertices, we can define the collection $\mathcal{G}_p(V)$ of all possible graphs made on the set of vertices *V*, such that the graph edges are built independently in $V \times V$ with a probability *p*. i.e. formally, we can write

$$\mathcal{G}_p(V) = \left\{ G(V,E) \,\middle|\, E \subseteq V \times V \wedge \forall\, v,w \in V :\ v \ne w \Rightarrow \Pr[vw \in E] = p \right\}$$

This model was be used in most areas of science and human activities in biology, chemistry, sociology, computer networks, manufacturing, etc. In a random Erdös-Rényi graph with *N* vertices, the edges are independently and randomly built with a probability *p* between the $N(N-1)/2$ possible edges of the full mesh graph. This definition builds the binomial model $\mathcal{G}_p(V)$ of random graphs, also referred as Erdös-Rényi model (Bollobàs, 2001).

Proposition 2

Consider a nonempty set V, a real p in $]0,1[$ *and a random graph G on V. Let* $N = |V|$ *be the order of G and M be the random order of G, then*

$$\Pr\left(M = m \,\middle|\, G \in \mathcal{G}_p(V)\right) = p^m (1-p)^{\binom{N}{2} - m} \tag{8}$$

and

$$E[M] = p\,\frac{N(N-1)}{2} \tag{9}$$

Proof

The set $\mathcal{G}_p(V)$ has $\binom{N}{2}$ random graphs with equal probabilities. In consequence, each one can be chosen with a probability equal to $\binom{N}{2}^{-1}$. Thus, on one hand,

$$\Pr\left(M = m \,|\, G \in \mathcal{G}_p(V)\right) = \binom{N}{2}^{-1}\binom{N}{2} p^m (1-p)^{\binom{N}{2}-m}$$
$$= p^m (1-p)^{\binom{N}{2}-m}$$

On the other hand, the number of edges M in a random Erdös-Rényi graph is a random variable with an average value equal to

$$E[M] = \sum_{m=0}^{N(N-1)/2} m\, p^m (1-p)^{\binom{N}{2}-m} = p\,\frac{N(N-1)}{2}$$

3.3 Degree distribution

As defined above, the degree function d_G on the graph vertices returns the number of vertices directly connected to the considered vertex. Thus, the *degree distribution* measures the local connectivity relevance of the graph vertices. In a generalized random graph model, the degree distribution can be set depending on the wished point of view. Its general form is defined by

$$p_k = \Pr\left[d_G(u) = k\right] \tag{10}$$

3.3.1 Binomial model and Poisson approximation

In particular, in the Enrdös–Rényi representation $\mathcal{G}_p(V)$, the theoretical degree distribution of any vertex u is defined by a binomial distribution (Bollobàs, 2001)

$$p_k = \binom{|V|-1}{k} p^k (1-p)^{|V|-1-k} \tag{11}$$

when the number of vertices $|V|$ is small. Otherwise, from the limit central theorem (LCT), the binomial distribution is approximately Gaussian with parameters

$$\begin{cases} \mu = \left(|V|-1\right) p & \left(\text{the mean}\right) \\ \sigma^2 = \left(|V|-1\right) p\left(1-p\right) & \left(\text{the variance}\right) \end{cases} \tag{12}$$

But the Gaussian distribution is continuous, while the binomial distribution is discrete. Thus, sometimes when $|V|$ is large and under certain special assumptions we prefer the Poisson approximation to the LCT approximation.

Proposition 3

When the graph order is large, the degree distribution is in the order of

$$p_k \approx e^{-\zeta} \frac{\zeta^k}{k!} \tag{13}$$

where p is small and $|V|$ is sufficiently large.

Proof

When the graph oder $|V|$ is large, we can also write the degree density as

$$\Pr\left[d(u)=k\right] = \binom{|V|-1}{k} \left[\frac{\zeta}{|V|-1}\right]^k \left(1-\frac{\zeta}{|V|-1}\right)^{|V|-1-k} \approx e^{-\zeta}\frac{\zeta^k}{k!}$$

where p is small and $|V|$ is sufficiently large. This shows the equation (13).

More generally, this approximation is known in probability theory as the De Moivre's Poisson approximation to the Binomial distribution. Indeed, the Poisson distribution can be applied whenever it is dealing with systems with a large number of possible events such that each of which is rare.

Thus, it is in the order of things to note that the main advantage of the Enrdös–Rényi models comes from the traceability of calculations and the simplicity of parameter estimation. A form or another of this model will be applied as and when it's required.

3.3.2 Power-law models

The traditional model of Erdös and Rényi is not a universal representation for all the random graph behaviors. Sometimes, in many natural scale-free networks (such as World Wide Web pages and their links, Internet, grid computer networking, etc), despite the randomness of the resulting graphs, experimental studies have revealed that the degree distribution can have a pure power-law tail (Albert et al., 1999; Barabàsi and Albert, 1999; Faloutsos et al. 1999)

$$p_k = \alpha\, k^{-\gamma} \tag{14}$$

where $\alpha > 0$ is a scaling constant and $\gamma > 0$ is a constant scaling exponent. Scale-free Barabàsi-Albert random graphs are generally built through a growth process combined to a

profile of preferential attachment to existing vertices (Barabàsi and Albert, 1999). Although such graphs have a large number of vertices, the degree distribution deviates significantly from the Poisson law expected for classical random graphs (Barabàsi and Albert, 1999). Other forms of degree distribution with a power-law tail have been studied (Amaral et al., 2000; Newman et al., 2001)

$$p_k = C\ k^{-\gamma} \exp\left(-\frac{k}{\kappa}\right);\ \text{for}\ k > 1 \tag{15}$$

where C is a constant fixed by the requirement of normalization, γ is the constant scaling exponent, and κ is a typical degree size from which the exponential adaptation becomes significant.

In scale-free graphs, the parameters estimation of the degree distribution has to be obviously adapted to each specific case of power-law. In general, the distribution coordinates have to be converted into the logarithmic scales and then apply a method such as least-square method.

4. Random Graph Dynamics (RGD)

In classic graph theory a graph is simply a collection of objects connected to each other in some manner. This description is very restrictive. In fact, the notions of random graph theory have been introduced in the objective to produce better models and more complete tools to represent non deterministic looks of configurations of a dynamic network. Here again, the language of random graphs is used simply to relate the graph structure of the different network situations. This language must be completed by defining a number of indispensable operations in order to introduce a basic mathematical framework for graph dynamics modeling.

Definition 11

A dynamic graph is a graph such that its configuration (or topology) is subject to dynamic changes with time.

Hence, it goes without saying that the topology changes induced by the random graph dynamics are made through a number of fundamental operations that affect only the set of edges E.

4.1 Random change process

Generally, we can define several kinds of graph dynamics. The following contexts are the basic ways of defining this kind of behavior:

- Vertex-dynamic graph model: the set of vertices varies with time. In this context vertices may be added or deleted. However, we must be careful in this case to the edges such that an end-vertex is removed. They should just be deleted too.
- Edge-dynamic graph model: the set of edges E varies with time. Thus, edges may be added or deleted from the graph. In this case, there are no consequences to fear on the set of vertices.

- Vertex-weighted dynamic graph model: the weights on the vertices vary with time.
- Edge-weighted dynamic graph model: the weights on the edges vary with time.
- Fully-weighted dynamic graph model: the weights on both vertices and edges vary with time.

Thus, we consider at first a model of dynamic graphs that combines all these aspects together. A such random dynamic graph G can be defined as a stochastic graph process i.e. a collection of independent random graphs $\mathbf{G} = \{G_t | t \in I\}$ where the parameter t is usually assumed to be time and which take values in a set I which can be continuous or countable (finite or infinite).

In the widest sense, each graph G_t can belong to a different model $\mathcal{G}_{\mathbf{P}_t}(\Omega)$. The set of all possible states is called the state space. If the state space is discrete, we deal with a discrete state stochastic graph process, which is called a chain of graphs. The state space can also be continuous; we then deal with a continuous-state stochastic process. A similar classification can be made regarding whether the index set I is continuous leading to a continuous-time stochastic process or countable leading to a discrete-time stochastic process. Thus, a dynamic graph **G** is a representation which assumes that at any time t, there exists an instance $G_t = (V_t, E_t)$ that belongs to a model $\mathcal{G}_{\mathbf{P}_t}(\Omega)$.

To explain the changes in a dynamic random graph, we must often refer to events of presence, absence, addition, deletion, birth, death and structure of objects; where the term "object" refers to both vertices and edges of the graph. The following definitions clarify what do these events really mean and how can they be mathematically defined in the context of graph dynamics.

Definition 12

The weighted graphs $(G_t; \mathbf{P}_t)$ *thus defined are called instantaneous configurations of the random dynamic graph* **G**.

The random graph dynamics can be characterized by the ordered stochastic time process $T_0 = 0 < T_1 < T_2 < \cdots < T_{k-1} < T_k < \cdots$ where the configuration changes of the random dynamic graph G are operated. Thus, a marked random graph process is defined such that, at the k^{th} time of change $T_k = t_k$, a $|\Omega| \times |\Omega|$ symmetric matrix $\mathbf{P}_k = \left(p_{u,v}^{(k)}\right)_{u,v \in \Omega}$ of probabilities is selected and a graph (configuration) $G_k = (V_k, E_k)$ is chosen according to the random graph model $\mathcal{G}_{\mathbf{P}_k}(\Omega)$.

In this model, the diagonal entries of the matrix $\mathbf{P}_k$ represent the probabilities to select, at time t_k, the vertices of V_k individually in Ω and the off-diagonal entries of this symmetric matrix represent the probabilities to activate the edges between two given vertices of V_k. Both the sets of vertices and edges in the graph configuration process are chosen as temporary. Thus, between two existing end-vertices u and v an edge is selected with the temporary probability $p_{u,v}^{(k)}$. Both the vertices and the edges of a dynamic random graph can be seen as subject to dynamic random changes.

This type of model is very interesting for modeling dynamic networks where the parameters are all temporary. As examples, we can mention routers and links of the Internet, friends in web social networking, communicating nodes in mobile wireless mobile ad hoc networks that use essentially the methods of broadcast.

4.2 Change operations

In this section, the studied approaches are classified according to their own definition in the context of graph dynamics. The RGD is said low when only few elements (number of vertices, number of edges, building edges probability, vertices clustering …) change over time. Otherwise, the dynamics is strong. Furthermore, there are several operations that build new graphs from old ones. They might be characterized through a number of descriptor events and basic transforms.

Definition 13 (Presence)

In a dynamic graph **G**, *an object is present at time t if it belongs to the instance* G_t *of* **G**.

Thus, a vertex u is present at time t, if and only if $u \in V_t$. Similarly, an edge uv is present at time t, if and only if $uv \in E_t$ knowing that both u and v belong to V_t.

Definition 14 (absence)

In a dynamic graph **G**, *an object is absent at time t, if it does not belong to the instance* G_t *of* **G**.

Operating dynamic changes on a random graph consists of a series of graph modifications (weight, vertex or edge adaptations). In contrast, the dynamicity analysis is more effective when all combination of additions and deletions of edges and vertices are taken into account.

Definition 15 (Addition)

An object appears in the dynamic graph **G**, *if it transits from the state absent to the state present.*

From an operational perspective, this definition should be clarified. Let u and v be two vertices of G such that uv is not an edge of G, the addition of the edge uv to the graph G is defined by the operation

$$G + uv = \left(V, E \cup \{uv\}\right) \tag{16}$$

Now, there are two ways to add a new vertex to a graph. One way is to add an isolated vertex to the graph G, i.e.

$$G \tilde{\oplus} u = \left(V \cup \{u\}, E\right) \tag{17}$$

and the other is to add a new vertex u which will be connected to an existing vertex v of the graph G, i.e.

$$G \underset{v}{\hat{\oplus}} u = \left(V \cup \{u\}, E \cup \{uv\}\right) \tag{18}$$

From the algorithmic point of view this last operation (18) can be seen as a composition of the two previous ones (16) and (17). Also, note that any instance $G=(V,E)$ of the model $\mathcal{G}_{\mathbf{P}}(\Omega)$ has a host $G^*=(\Omega,E)$ which is defined by adding to the graph *G* isolated vertices taken in $\Omega - V$

$$G^* = G \underset{u\in\Omega-V}{\breve{\oplus}} u \tag{19}$$

Thus, we conceive that the adjacency matrix of G^* is obtained by completing the adjacency matrix of *G* by 0. This matrix will be called in the rest of the chapter *extended adjacency matrix of G*. The advantage of working with the host graph G^* can be viewed rather as freeing from the assumption that the set of vertices of a random dynamic graph varies with time. But the downside of this alternative is that the graph model can exhibit needlessly an excessively large order or incomplete information.

Proposition 4

Let $V \subset \Omega$ *be two nonempty sets such that* $G=(V,E)$ *be a graph and* $G^*=(\Omega,E)$ *be a host for G. Then*

$$G \underset{v}{\hat{\oplus}} u = \left(G \breve{\oplus} u\right) + uv \tag{20}$$

Proof

This property results trivially from the definitions corresponding to the operations $\hat{\oplus}$ and $\breve{\oplus}$.

Definition 16 (Deletion)

An object disappears (or is deleted) from the dynamic graph **G**, *if it transits from the state present to the state absent.*

In these terms, the edge deletion can be formalized as follow. Let *u* and *v* be two vertices of a graph $G=(V,E)$ such that *uv* is an edge of *G*, the deletion of the edge *uv* from the graph *G* is defined by the operation

$$G - uv = (V, E - \{uv\}) \tag{21}$$

On another hand, the vertex deletion can be defined

$$G\Delta u = \left(V - \{u\}, E - \bigcup_{v\in N(u)} \{uv\}\right) \tag{22}$$

Indeed, the deletion of the vertex *u* induces the deletion of all the edges having *u* as end-vertex.

Definition 17 (Birth)

The birth of an object is the date of its first appearance in the dynamic graph ***G****.*

Formally, the birth of a vertex *u* is the date τ_u^+ of its first occurrence

$$\tau_u^+ = \min\left\{t \in \mathbf{R}+ \middle| u \in V_t\right\} \tag{23}$$

and the birth of an edge *e* is the date τ_u^+ of its first occurrence

$$\tau_e^+ = \min\left\{t \in \mathbf{R}+ \middle| e \in E_t\right\} \tag{24}$$

Definition 18 (Death)

Death of an element is the date for the last deletion of this element from the dynamic graph **G**.

Thus, the death of a vertex is the date τ_u^- of its last occurrence

$$\tau_u^- = \max\left\{t \in \mathbf{R}+ \middle| u \in V_t\right\} \tag{25}$$

and the birth of an edge *e* is the date τ_e^- such that

$$\tau_e^- = \max\left\{t \in \mathbf{R}+ \middle| e \in E_t\right\} \tag{26}$$

Definition 19 (Structure)

A structure S of a graph consists in a dynamical set of elements that satisfies a given property.

A path, a click and a cluster are all examples of structures. Let us note that from the random viewpoint the succession of graph transforms and structure updates that convert a configuration of the dynamic graph in another one are not known in advance.

4.3 Graph topology changes in RGD

Remark that between two consecutive changes of the graph configuration recorded at T_k and T_{k+1}, the extended adjacency matrix $A^{(k)}$ of G_k remains unchanged all through the interval $[T_k, T_{k+1}[$. Thus, the dynamicity of the graph topology can be characterized by the variation of the extended adjacency matrix between T_{k-1} and T_k

$$\Delta A^{(k)} = A^{(k)} - A^{(k-1)} \tag{27}$$

where $A^{(k)}$ is the extended adjacency matrix of the current configuration of the graph at time $T_k = t_k$ and $A^{(k-1)}$ is the extended adjacency matrix of the previous configuration at time $T_{k-1} = t_{k-1}$ (with $T_{k-1} < T_k$).

4.3.1 Characterizing the change of the number of vertices

We can define for successive configurations of the graph, the number of new vertices, the number of lost vertices and the number of maintained vertices respectively at $T_k = t_k$ by

$$\alpha^{(k)} = \sum_{u\in\Omega} \mathbf{1}_{\left\{\Delta a_{u,u}^{(k)} > 0\right\}},\ \beta^{(k)} = \sum_{u\in\Omega} \mathbf{1}_{\left\{\Delta a_{u,u}^{(k)} = 0\right\}} \quad \text{and} \quad \gamma^{(k)} = \sum_{v\in\Omega} \mathbf{1}_{\left\{\Delta a_{u,u}^{(k)} < 0\right\}} \tag{28}$$

where $\mathbf{1}_S$ indicates the characteristic function of the set S and $\Delta a_{u,u}^{(k)}$ symbolizes the diagonal term of the matrix $\Delta A^{(k)}$ associated to the eventual vertices of the dynamic graph.

Proposition 5

Let $\mathbf{G}$ *be a random dynamic graph following the model* $\mathcal{G}_{\mathbf{P}_t}(\Omega)$ *such that the configuration change is characterized by the stochastic time process* $T_0 = 0 < T_1 < T_2 < \cdots < T_{k-1} < T_k < \cdots$. *Consider the extended adjacency matrices* $A^{(k)} = \left(a_{u,v}^{(k)}\right)_{u,v\in\Omega}$ *of* $G_k = (V_k, E_k)$ *and* $\Delta A^{(k)}$ *the matrix defined by the equation (27) . Then,*

$$\begin{cases} \Pr\left(\Delta a_{u,u}^{(k)} = 1\right) = p_{u,u}^{(k)}\left(1 - p_{u,u}^{(k-1)}\right) \\ \Pr\left(\Delta a_{u,u}^{(k)} = 0\right) = p_{u,u}^{(k)}\, p_{u,u}^{(k-1)} + \left(1 - p_{u,u}^{(k)}\right)\left(1 - p_{u,u}^{(k-1)}\right) \\ \Pr\left(\Delta a_{u,u}^{(k)} = -1\right) = \left(1 - p_{u,u}^{(k)}\right) p_{u,u}^{(k-1)} \end{cases} \tag{29}$$

Proof

Since all trials are independent, each of the three probabilities can be decomposed as follows

$$\begin{cases} \mathbf{Pr}\left(\Delta a_{u,u}^{(k)} = 1\right) = \mathbf{Pr}\left(\left\{a_{u,u}^{(k)} = 1\right\} \wedge \left\{a_{u,u}^{(k-1)} = 0\right\}\right) = \mathbf{Pr}\left(a_{u,u}^{(k)} = 1\right)\mathbf{Pr}\left(a_{u,u}^{(k-1)} = 0\right) \\ \mathbf{Pr}\left(\Delta a_{u,v}^{(k)} = 0\right) = \mathbf{Pr}\left[\left(\left\{a_{u,u}^{(k)} = 1\right\} \cap \left\{a_{u,u}^{(k-1)} = 1\right\}\right) \cup \left(\left\{a_{u,u}^{(k)} = 0\right\} \cap \left\{a_{u,u}^{(k-1)} = 0\right\}\right)\right] = \\ \qquad = \mathbf{Pr}\left(\left\{a_{u,u}^{(k)} = 1\right\} \cap \left\{a_{u,u}^{(k-1)} = 1\right\}\right) + \mathbf{Pr}\left(\left\{a_{u,u}^{(k)} = 0\right\} \cap \left\{a_{u,u}^{(k-1)} = 0\right\}\right) = \\ \qquad = \mathbf{Pr}\left(a_{u,u}^{(k)} = 1\right)\mathbf{Pr}\left(a_{u,u}^{(k-1)} = 1\right) + \mathbf{Pr}\left(a_{u,u}^{(k)} = 0\right)\mathbf{Pr}\left(a_{u,u}^{(k-1)} = 0\right) \\ \mathbf{Pr}\left(\Delta a_{u,u}^{(k)} = -1\right) = \mathbf{Pr}\left(\left\{a_{u,u}^{(k)} = 0\right\} \cap \left\{a_{u,u}^{(k-1)} = 1\right\}\right) = \mathbf{Pr}\left(a_{u,u}^{(k)} = 0\right)\mathbf{Pr}\left(a_{u,u}^{(k-1)} = 1\right) \end{cases}$$

and because $\Pr\left(a_{u,u}^{(k-1)} = 1\right) = p_{u,u}^{(k-1)}$, $\Pr\left(a_{u,u}^{(k-1)} = 0\right) = 1 - p_{u,u}^{(k-1)}$, $\Pr\left(a_{u,u}^{(k)} = 1\right) = p_{u,u}^{(k)}$ and $\Pr\left(a_{u,v}^{(k)} = 0\right) = 1 - p_{u,v}^{(k)}$ we have the result (29).

Corollary 6

Under the assumptions of the above proposition, and by supposing that the configuration process of the dynamic graph is homogeneous and that all the element of Ω *can appear with the same probability* π *in a configuration of the dynamic random graph* $\mathbf{G}$, *the triplet* $\left(\alpha^{(k)}, \beta^{(k)}, \gamma^{(k)}\right)$ *follows a trinomial distribution of parameters*

$$\left(|\Omega|;\pi\left(1-\pi\right),\pi\left(1-\pi\right),\pi^2+\left(1-\pi\right)^2\right) \text{ i.e.}$$

$$\Pr\left(\alpha^{(k)}=n_1,\beta^{(k)}=n_2,\ \gamma^{(k)}=n_3\right)=\frac{|\Omega|!}{n_1!\,n_2!\,n_3!}\left[\pi\left(1-\pi\right)\right]^{n_1+n_2}\left[\pi^2+\left(1-\pi\right)^2\right]^{n_3} \tag{30}$$

Proof

On one hand, there are only three possible outcomes $\left\{\Delta a_{u,u}^{(k)}=1\right\}$, $\left\{\Delta a_{u,u}^{(k)}=0\right\}$ and $\left\{\Delta a_{u,u}^{(k)}=-1\right\}$ such that $\Pr\left(\Delta a_{u,u}^{(k)}=1\right)+\Pr\left(\Delta a_{u,u}^{(k)}=0\right)+\Pr\left(\Delta a_{u,u}^{(k)}=-1\right)=1$ and all trials are independent. On the other hand, we have $|\Omega|$ discrete trials (relating the different situations in the set Ω) and because the occurrence measures of graph vertex situations $\alpha^{(k)}$, $\beta^{(k)}$ and $\gamma^{(k)}$ are interrelated by the equation $\alpha^{(k)}+\beta^{(k)}+\gamma^{(k)}=|\Omega|$, this means that we deal with a multinomial distribution of $|\Omega|$ trials and from the previous proposition (proposition 5) three outcomes with respective probabilities

$$\begin{cases}\Pr\left(\Delta a_{u,u}^{(k)}=1\right)=\pi\left(1-\pi\right)\\ \Pr\left(\Delta a_{u,u}^{(k)}=0\right)=\pi^2+\left(1-\pi\right)^2\\ \Pr\left(\Delta a_{u,u}^{(k)}=-1\right)=\left(1-\pi\right)\pi\end{cases}$$

This shows the required result in (30).

Following this line of thinking, these results can be applied only if the number of the observed vertices varies with time. That is when the dynamicity affects the vertices of the dynamic graph. Otherwise, while the number observed vertices remain unchanged, we will show similar results in the next subsection under the assumption of dynamicity of edges. These two approaches are complementary.

4.3.2 Characterizing the change of the number of edges

First and foremost let us remain under the assumption of the general model $\mathcal{G}_{\mathbf{P}_t}(\Omega)$ of dynamic graphs. In the context of connectivity, the number of edges connected to the same vertex defines its degree. Thus, from the uncertain connectivity viewpoint of dynamic random graphs, we can define locally for each graph vertex the number of new neighbors, the number of lost neighbors and the number of maintained neighbors respectively at $T_k=t_k$ by

$$\nu_u^{(k)}=\sum_{v\in V_k}\mathbf{1}_{\left\{\Delta a_{u,v}^{(k)}>0\right\}},\ \mu_u^{(k)}=\sum_{v\in V_k}\mathbf{1}_{\left\{\Delta a_{u,v}^{(k)}=0\right\}} \text{ and } \lambda_u^{(k)}=\sum_{v\in V_k}\mathbf{1}_{\left\{\Delta a_{u,v}^{(k)}<0\right\}} \tag{31}$$

where $\mathbf{1}_S$ indicates the characteristic function of the set S and $\Delta a_{u,v}^{(k)}$ indicates a generic entry of the matrix $\Delta A^{(k)}$. Essentially, the knowledge, step by step, of the evolution of the local metrics $\nu_u^{(k)}$ and $\lambda_u^{(k)}$ will allow us to determine a propagation equation of degree function at each vertex of the observed dynamic graph.

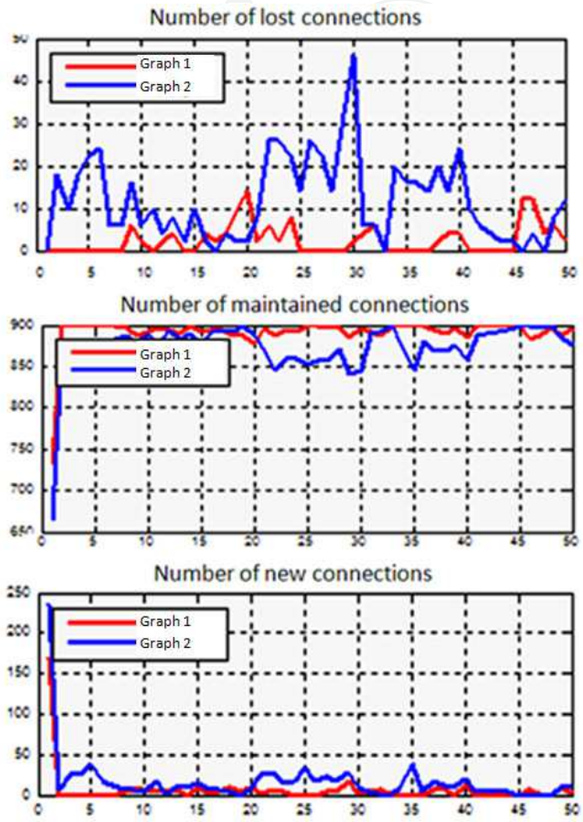


Fig. 1. Connectivity metrics change for two simulated random dynamic graphs

Proposition 7

Let $\mathbf{G}$ *be a random dynamic graph following the model* $\mathscr{G}_{\mathbf{P}_t}(\Omega)$ *and consider the configurations* $(G_k)_{k\in\mathbf{N}}$ *of* $\mathbf{G}$ *associated to the stochastic time process* $T_0 = 0 < T_1 < T_2 < \cdots < T_{k-1} < T_k < \cdots$. *If* d_{G_k} *denotes the underlying degree function to* G_k

$$\forall u \in \Omega:\ d_{G_k}(u) = d_{G_{k-1}}(u) + \nu_u^{(k)} - \lambda_u^{(k)} \tag{32}$$

Proof

The proof of (32) results directly from the definitions of $\nu_u^{(k)}$ and $\lambda_u^{(k)}$.

We can also define global graph indices which reflect the global configuration changes of the dynamic random graph (see figure 1)

$$\nu^{(k)} = \sum_{v \in V_k} \nu_u^{(k)}\ ,\ \mu^{(k)} = \sum_{v \in V_k} \mu_u^{(k)} \text{ and } \lambda^{(k)} = \sum_{v \in V_k} \lambda_u^{(k)} \tag{33}$$

Under the condition of conservation of the number of vertices from a configuration to a successor one, we can establish a number of results for dynamic graphs under Erdös-Rényi model constrains.

Proposition 8

Let $\mathbf{G}$ *be a random dynamic graph following the model* $\mathcal{G}_p(V)$ *such that the configuration change is characterized by the stochastic time process* $T_0 = 0 < T_1 < T_2 < \cdots < T_{k-1} < T_k < \cdots$. *Consider the extended adjacency matrices* $A^{(k)} = \left(a_{u,v}^{(k)}\right)_{u,v \in V_k}$ *of* $G_k = (V, E_k)$ *and* $\Delta A^{(k)}$ *the matrix defined by the equation (27) and such that all the terms of its principal diagonal are equal to zero, then*

$$\begin{cases} \Pr\left(\Delta a_{u,v}^{(k)} = 1\right) = p(1-p) \\ \Pr\left(\Delta a_{u,v}^{(k)} = 0\right) = p^2 + (1-p)^2 \\ \Pr\left(\Delta a_{u,v}^{(k)} = -1\right) = (1-p)p \end{cases} \tag{34}$$

Proof

Since all trials are independent, each of the three probabilities can be decomposed such as follows

$$\begin{cases} \Pr\left(\Delta a_{u,v}^{(k)} = 1\right) = \Pr\left(\left\{a_{u,v}^{(k)} = 1\right\} \cap \left\{a_{u,v}^{(k-1)} = 0\right\}\right) = \Pr\left(a_{u,v}^{(k)} = 1\right) \Pr\left(a_{u,v}^{(k-1)} = 0\right) \\ \Pr\left(\Delta a_{u,v}^{(k)} = 0\right) = \Pr\left[\left(\left\{a_{u,v}^{(k)} = 1\right\} \cap \left\{a_{u,v}^{(k-1)} = 1\right\}\right) \cup \left(\left\{a_{u,v}^{(k)} = 0\right\} \cap \left\{a_{u,v}^{(k-1)} = 0\right\}\right)\right] = \\ \qquad = \Pr\left(\left\{a_{u,v}^{(k)} = 1\right\} \cap \left\{a_{u,v}^{(k-1)} = 1\right\}\right) + \Pr\left(\left\{a_{u,v}^{(k)} = 0\right\} \cap \left\{a_{u,v}^{(k-1)} = 0\right\}\right) = \\ \qquad = \Pr\left(a_{u,v}^{(k)} = 1\right) \Pr\left(a_{u,v}^{(k-1)} = 1\right) + \Pr\left(a_{u,v}^{(k)} = 0\right) \Pr\left(a_{u,v}^{(k-1)} = 0\right) \\ \Pr\left(\Delta a_{u,v}^{(k)} = -1\right) = \Pr\left(\left\{a_{u,v}^{(k)} = 0\right\} \cap \left\{a_{u,v}^{(k-1)} = 1\right\}\right) = \Pr\left(a_{u,v}^{(k)} = 0\right) \Pr\left(a_{u,v}^{(k-1)} = 1\right) \end{cases}$$

and because $\Pr\left(a_{u,v}^{(k)} = 1\right) = \Pr\left(a_{u,v}^{(k-1)} = 1\right) = p$ and $\Pr\left(a_{u,v}^{(k)} = 0\right) = \Pr\left(a_{u,v}^{(k-1)} = 0\right) = 1 - p$, we have the result (34).

Corollary 9

Under the assumptions of the above proposition, and by supposing that the random dynamic graph belongs to the $\mathcal{G}_p(V)$ *model, the triplet* $\left(\nu^{(k)}, \mu^{(k)}, \lambda^{(k)}\right)$ *follows a trinomial distribution of parameters* $\left(|V|(|V|-1)/2; p(1-p), p(1-p), p^2 + (1-p)^2\right)$. *i.e.*

$$\Pr\left(\nu^{(k)}=n_1,\,\mu^{(k)}=n_2,\,\lambda^{(k)}=n_3\right)=\frac{\binom{|V|}{2}!}{n_1!\,n_2!\,n_3!}\left[p\left(1-p\right)\right]^{n_1+n_2}\left[p^2+\left(1-p\right)^2\right]^{n_3} \tag{35}$$

Proof

First, note that all pairs of vertices u and v, there are only three possible independent outcomes $\left\{\Delta a_{u,v}^{(k)}=1\right\}$, $\left\{\Delta a_{u,v}^{(k)}=0\right\}$ and $\left\{\Delta a_{u,v}^{(k)}=-1\right\}$. Furthermore, we have

$$\Pr\left(\Delta a_{u,v}^{(k)}=1\right)+\Pr\left(\Delta a_{u,v}^{(k)}=0\right)+\Pr\left(\Delta a_{u,v}^{(k)}=-1\right)=1$$

and all trials are independent. Subsequently, we have $|V|$ discrete trials and since there are no other situations than those described above, the three measures of occurrence of graph edges $\nu^{(k)}$, $\mu^{(k)}$ and $\lambda^{(k)}$ are interrelated and verify

$$\nu^{(k)}+\mu^{(k)}+\lambda^{(k)}=|V|\left(|V|-1\right)/2 \tag{36}$$

Each of the $|V|\left(|V|-1\right)/2$ trials (off-diagonal generic entries of the symmetric matrix $\Delta A^{(k)}$) matches to one of the three possible outcomes with respective probabilities

$$\Pr\left(\Delta a_{u,v}^{(k)}=1\right)=p\left(1-p\right),\ \Pr\left(\Delta a_{u,v}^{(k)}=0\right)=p^2+\left(1-p\right)^2 \text{ and } \Pr\left(\Delta a_{u,v}^{(k)}=-1\right)=\left(1-p\right)p \tag{37}$$

Thus, by taking into account the results established by the equation (36) and the system (37), we can conclude that, *the triplet* $\left(\nu^{(k)},\,\mu^{(k)},\,\lambda^{(k)}\right)$ *follows a trinomial distribution of parameters* $\left(|V|\left(|V|-1\right)/2;p\left(1-p\right),p\left(1-p\right),p^2+\left(1-p\right)^2\right)$.

Corollary 10

Under the assumptions of the proposition 8, the off-diagonal entries of the correlation matrix associated to $\left(\nu^{(k)},\,\mu^{(k)},\,\lambda^{(k)}\right)$ *are such that*

$$\rho\left(\nu^{(k)},\mu^{(k)}\right)=\rho\left(\lambda^{(k)},\mu^{(k)}\right)=-\sqrt{\frac{\left(1-p\right)^2+p^2}{2\left[\left(1-p\right)^2+p\right]}} \text{ and } \rho\left(\lambda^{(k)},\nu^{(k)}\right)=-\frac{p\left(1-p\right)}{\left(1-p\right)^2+p} \tag{38}$$

Proof

The corollary results from the general properties of the correlation matrix of a trinomial distribution of parameters $\left(|V|\left(|V|-1\right)/2;p\left(1-p\right),p\left(1-p\right),p^2+\left(1-p\right)^2\right)$.

Note that the graph order $|V|\left(|V|-1\right)/2$ drop out of the off-diagonal entries of the correlation matrix associated to $\left(\nu^{(k)},\,\mu^{(k)},\,\lambda^{(k)}\right)$.

5. Parameter estimation of the random dynamic graph model

The ultimate objective in random graph dynamics analysis is the estimation of the graph model parameters. In some respects, nearly all types of dynamic changes can be interpreted in this way. In the empirical study, for reasons of traceability of the calculations, we restrict ourselves to an homogeneous model $\mathcal{G}_{\mathbf{P}_t}(\Omega)$ such that, there exists a non-negative real $p \in]0,1[$, for all existing pair of vertices u and v $p_{u,v}^{(k)} = p$.

Furthermore, we assume that by some means a vertex may deduct its neighborhood set directly from information exchanged as part of edge sensing. This section aims to show how the model parameters of the random dynamic graph can be estimated.

5.1 Dynamic graphs with known number of vertices and unknown edge probability

The main goal here is to find the maximum-likelihood estimate (MLE) of the edge probability p when the number of the graph vertices $N = |V|$ is known.

Proposition 11

Let **G** *be a random dynamic graph following the model* $\mathcal{G}_p(V)$ *and note* $\zeta = p|V|$ *when p is small and* $|V|$ *is large. Consider D the random variable counting the degree in a fixed vertex u of* **G** *and let* $D_1, D_2, \cdots, D_s$ *be a s-sample formed by the degrees of u for s different configurations of the random dynamic graph at the graph change times* $T_1 < T_2 < \cdots < T_s$ *. Then*

$$\begin{cases} \hat{p}_{MLE} = \dfrac{1}{N-1}\bar{D} & \text{if } N \text{ is small} \\ \hat{\zeta}_{MLE} = \bar{D} & \text{if } N \text{ is large} \end{cases} \tag{39}$$

Proof

If the graph order $N = |V|$ is small, from equation (11), the log-likelihood function of the degree distribution in a given vertex u is

$$\ell(p, N; D_1, \cdots, D_s) = \sum_{i=1}^{s} \log\binom{N-1}{D_i} + \log\left(\frac{p}{1-p}\right) \sum_{i=1}^{s} D_i + s\ (N-1)\ \log(1-p)$$

The MLE of p is obtained by solving the partial differential equation $\dfrac{\partial\, \ell(p, N; D_1, \cdots, D_s)}{\partial\, p} = 0$

i.e.

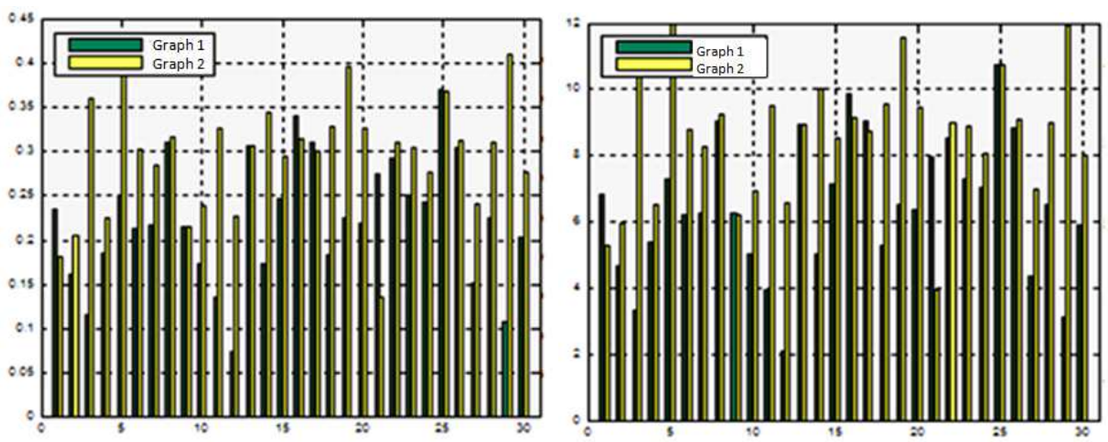
$$\hat{p}_{MLE} = \frac{\sum_{i=1}^{s} D_i}{s(N-1)} = \frac{1}{N-1}\bar{D}$$

Now, if N is large, from equation (13), the log-likelihood function of the degree distribution in a given vertex u is $\ell(\zeta; D_1, \cdots, D_s) = \log(\zeta) \sum_{i=1}^{s} D_i - \sum_{i=1}^{s} \log(D_i!) - s\,\zeta$

The MLE of p is obtained by solving the partial differential equation $\frac{\partial \ell(\zeta; D_1, \cdots, D_s)}{\partial \zeta} = 0$ i.e.

$$\hat{\zeta}_{MLE} = \frac{\sum_{i=1}^{s} D_i}{s} = \overline{D}$$

When s observations of the change of the random dynamic graph topology are accomplished, the variations of $\hat{p}_{MLE}$ and $\hat{\zeta}_{MLE}$ displayed in one vertex or another relates the local knowledge on the values of these parameters taking into account the observed neighborhood behavior at u.



$\hat{p}_{MLE}$ $\hat{\zeta}_{MLE}$

Fig. 2. $\hat{p}_{MLE}$ and $\hat{\zeta}_{MLE}$ evaluations for two simulated random graphs of order 30 vertices

Indeed, since $\hat{p}_{MLE}$ depends only on the observed number of neighbors of a selected vertex for different configurations of the random dynamic graph, this quantity is estimated differently in each vertex (see figure 2).

5.2 Dynamic graphs with unknown number of vertices and edge probability

In the statistical common sense, when the graph order is unknown the method proposed in the previous paragraph cannot be applied. Another resourceful method of point estimation called method of moments (Lehmann E. L. and Casella G., 1998) can be used when the number N of the graph vertices and the edge density parameter p are both unknown. Likewise, these parameters should be estimated on the basis of an s-sample $D_1, D_2, \cdots, D_s$ of the degree of a vertex u for s different instances of the random dynamic graph topology.

Proposition 12

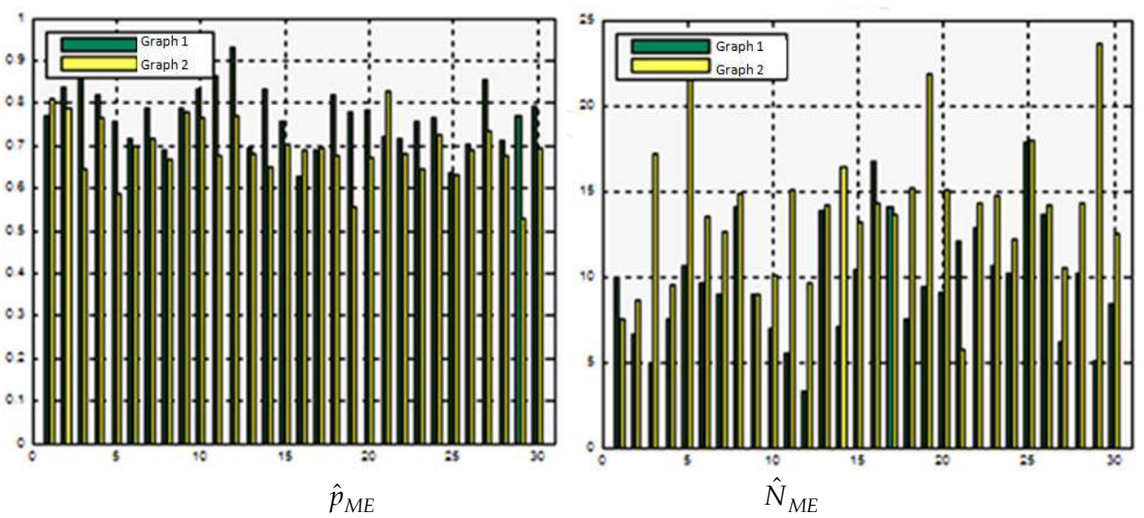
Let ***G*** *be a random dynamic graph following the model* $\mathcal{G}_p(V)$ *and suppose that* $N=|V|$ *is small. Consider D the random variable counting the degree in a fixed vertex u of* ***G*** *and let* $D_1, D_2, \cdots, D_s$ *be a s-sample formed by the degrees of u for s different configurations of the random dynamic graph at the graph change times* $T_1 < T_2 < \cdots < T_s$ *. Then the estimates of moments of the graph parameters are respectively*

$$\begin{cases} \hat{p}_{ME} = 1 - \dfrac{S^2}{\bar{D}} \\ \hat{N}_{ME} = 1 + \dfrac{\bar{D}^2}{\bar{D} - S^2} \end{cases} \tag{40}$$

Proof

In this case, the method of moments supposes that the empirical moment $\bar{D}$ is a natural estimate of the theoretical moment of order 1 $E(D)$ and the theoretical centralized moments of order k $\mu_k = E\left[D - (D)\right]^k$ can be estimated by their respective empirical centralized moments $M_k = \dfrac{1}{s}\sum_{i=1}^{s}\left(D_i - \bar{D}\right)^k$. When $N=|V|$ is small, the degree follows a binomial distribution of parameters p an $(N-1)$ and the method of moments in terms of the average and the variance of the observed degrees leads to the system of equations

$$\begin{cases} (N-1)\,p = \bar{D} \\ (N-1)\,p\,(1-p) = S^2 \end{cases} \tag{41}$$



$\hat{p}_{ME}$ $\hat{N}_{ME}$

Fig. 3. $\hat{p}_{ME}$ and $\hat{N}_{ME}$ evaluations for two simulated dynamic random graphs

It results from the resolution of the system of equations (41) that the moment estimates of the unknown graph parameters

$$\begin{cases} \hat{p}_{ME} = 1 - \dfrac{S^2}{\bar{D}} \\ \hat{N}_{ME} = 1 + \dfrac{\bar{D}^2}{\bar{D} - S^2} \end{cases}$$

Depending on the neighbors met by each of the random dynamic graph vertices, the evaluation of the estimated parameters will be different.

5.3 Expected degree number of vertices in wide random dynamic graphs

Various random dynamic graph problems (Internet, out vehicle networks, wide ad hoc networks …) are analyzed and interpreted under the assumption that the order of the resulting graph may be relatively large (with some tens, hundreds or even thousands of vertices). Let ζ be the average degree number

$$\zeta = E(D) \tag{42}$$

As a result of scaling, since we use an approximation of the degree distribution, it is very important to work with good estimates. This shows that the resulting computed values are most likely not due merely to chance.

Proposition 13

Let **G** *be a random dynamic graph following the model* $\mathcal{G}_p(V)$ *and suppose that* $N = |V|$ *is large. Consider D the random variable counting the degree in a fixed vertex u of* **G** *and let* $D_1, D_2, \cdots, D_s$ *be a s-sample formed by the degrees of u for s different configurations of the random dynamic graph at the graph change times* $T_1 < T_2 < \cdots < T_s$ *. Let* ζ *be the Poisson distribution parameter of the degree. Then, the empirical moment of order 1* $\hat{\zeta} = \bar{D}$ *is a good estimate of the average degree. i.e.*

i. $\hat{\zeta}$ *is unbiased.*

ii. $\hat{\zeta}$ *realizes the maximum of the likelihood function.*

iii. $\hat{\zeta}$ *is an efficient estimate of* ζ *.*

Proof

The first property results from

$$E(\bar{D}) = \frac{1}{s} E\left(\sum_{i=1}^{s} D_i \right) = \frac{1}{s} \sum_{i=1}^{s} E(D_i) = E(D) = \zeta$$

because $D_1, D_2, \cdots, D_s$ are independent and identically distributed to D.

Furthermore, the log-likelihood function of a sample $D_1, D_2, \cdots, D_s$ representing the degrees of a given vertex *u* of *G* for *s* different configurations of the random dynamic graph is given by

$$\ell(\zeta; D_1, \cdots, D_s) = -\sum_{i=1}^{s} \log D_i! + \sum_{i=1}^{s} D_i \log \zeta - s\,\zeta$$

The MLE is obtained by solving the partial differential equation $\frac{\partial\, \ell(\zeta; D_1, \cdots, D_s)}{\partial\, \zeta} = 0$ i.e

$$\hat{\zeta} = \frac{\sum_{i=1}^{s} D_i}{s} = \bar{D}$$

which shows the proposition *ii*. It follows, on one hand,

$$\mathrm{var}(\hat{\zeta}) = \frac{1}{s^2} \sum_{i=1}^{s} \mathrm{var}(D_i) = \frac{1}{s} \mathrm{var}(D) \quad \text{(because are independent and identically distibuted to } D\text{)}$$

$$= \frac{\zeta}{s} \quad \text{(because } D \text{ is Poisson distributed with parameter } \zeta\text{)}$$

and on the other hand,

$$\frac{\partial\, \ell(\zeta; D_1, \cdots, D_s)}{\partial\, \zeta} = -s + \frac{\sum_{i=1}^{s} D_i}{\zeta} \text{ and } \frac{\partial^2\, \ell(\zeta; D_1, \cdots, D_s)}{\partial\, \zeta^2} = -\frac{\sum_{i=1}^{s} D_i}{\zeta^2}$$

Thus, the Fisher information quantity of the parameter ζ is such that

$$I(\zeta) = E\left[-\frac{\partial^2\, \ell(\zeta; D_1, \cdots, D_s)}{\partial\, \zeta^2} \right] = \frac{\sum_{i=1}^{s} E(D_i)}{\zeta^2} = \frac{s}{\zeta} = \frac{1}{\mathrm{var}(\hat{\zeta})} \tag{43}$$

which is the lower bound of the Fréchet-Darmois-Cramer-Rao (FDCR) inequality.

Then, since $\hat{\zeta}$ is unbiased and verifies the lower bound of the FDCR inequality, it is an efficient estimate.

5.4 Routing performance metrics evaluation

5.4.1 Average hopcount estimation

Proposition 14

Let **G** *be a random dynamic graph following the model* $\mathcal{G}_p(V)$ *and suppose that* $|V|$ *is large and known. Consider D the random variable counting the degree in a fixed vertex u of* **G** *and let* $D_1, D_2, \cdots, D_s$ *be a s-sample formed by the degrees of u for s different configurations of the random dynamic graph at the graph change times* $T_1 < T_2 < \cdots < T_s$ *. Then, the average hopcount MLE is approximately*

$$\hat{\eta}_{MLE} \approx \frac{\log|V|}{\log \bar{D}} \tag{44}$$

Proof

Let ζ be the average degree number. As each vertex in the random graph is connected to about ζ other nodes, after η hops, we expect that ζ^{η} vertices must be reached. Thus all the vertices are reached for the value η_V such that $\zeta^{\eta_V} \approx |V|$. Since $|V|$ and ζ are both known constant parameters, the average hopcount $E[\eta_V] \approx \frac{\log|V|}{\log \zeta}$ in large random graphs is

$$E[\eta_V] \approx \frac{\log|V|}{\log \zeta} \tag{45}$$

From the proposition 11 and the equation (45), and because the MLE of a function of a parameter is equal to the function of the MLE of this parameter, the average hopcount MLE is made by

$$\hat{\eta}_{MLE} \approx \frac{\log|V|}{\log \bar{D}}$$

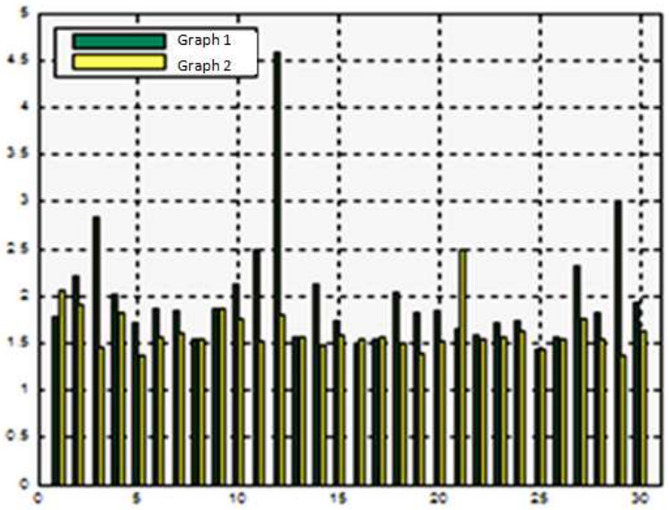


Fig. 4. Hopcount MLE evaluation for two simulated dynamic random graphs

This approximation holds because $\bar{D}$ is a good estimate of ζ in the strict sense of the proposition 13. Under other assumptions than those suggested in this paragraph, different authors have led to further forms of the approximation of the average hopcount number (Mieghem et al., 2000; Bhamidi et al., 2010).

5.4.2 Giant component size estimation

Let us consider **G** a random dynamic graph following the model $\mathcal{G}_p(V)$ and suppose that the graph order $|V|$ is large and known. From the theory of Erdös-Rényi graphs, we know (Molloy and Reed, 1998) that the graph will almost surely have a unique giant component containing a positive fraction of the graph vertices if $p > 1/|V|$ (see figure 5) i.e. the average degree numbet $\zeta > 1$.

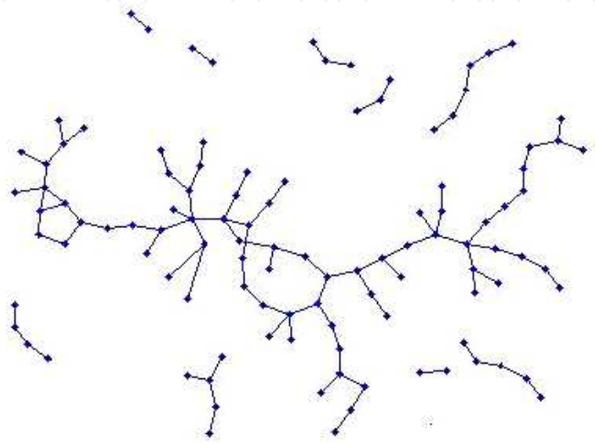


Fig. 5. Existence of the giant component of a graph

Proposition 15

Let **G** *be a random dynamic graph following the model* $\mathcal{G}_p(V)$ *and suppose that* $|V|$ *is large and known. Consider* D *the random variable counting the degree in a fixed vertex* u *of* **G** *and let* $D_1, D_2, \cdots, D_s$ *be a s-sample formed by the degrees of* u *for* s *different configurations of the random dynamic graph at the graph change times* $T_1 < T_2 < \cdots < T_s$. *Then, the average size MLE of the giant component is approximately*

$$\hat{\Theta} = 1 + \frac{1}{\bar{D}} \text{LambertW}\left[-\bar{D}\exp\left(-\bar{D}\right)\right] \tag{46}$$

Proof

Let Θ be the giant component size, for large order random graphs, Θ is a solution of the equation

$$\Theta = 1 - \exp\left(-\zeta\,\Theta\right) \tag{47}$$

But, it is well known that there are two possible solutions to this equation

$$\begin{cases} \Theta_1 = 0 \\ \Theta_2 = 1 + \dfrac{1}{\zeta}\text{LambertW}\left[-\zeta\exp\left(-\zeta\right)\right] \end{cases} \tag{48}$$

where LambertW is the classic "Lambert W " function. Since only the non-zero solution is adequate, the giant component size increases as a function of the average degree ζ of the graph vertices G following the curve $\vartheta : z \mapsto 1 + \frac{1}{z}\text{LambertW}\left[-z\exp(-z)\right]$.

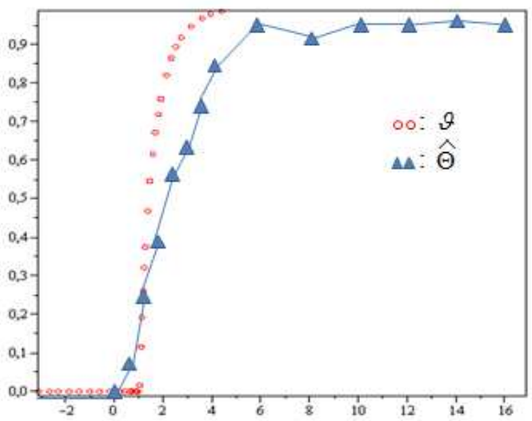


Fig. 6. Comparison of theoretic and empiric solutions ϑ and $\hat{\Theta}$

Thus, from the proposition 11 and because the MLE of a function of an unknown parameter is equal to the function of the MLE of this parameter, when $|V|$ is large the part of the graph occupied by the giant component can be estimated by

$$\hat{\Theta} = 1 + \frac{1}{\bar{D}}\,\text{LambertW}\left[-\bar{D}\exp\left(-\bar{D}\right)\right]$$

Elementary dissimilarities exist between random graph models and real-world graphs. Real-world graphs show strong clustering, but Erdös and Rényi's model does not. Many of the graphs, including Internet and World-Wide Web graphs, show a power-law degree distribution (Albert et al., 1999). This means that only a small part of the graph vertices can have a large degree. In fact, Erdös-Rényi assumptions can imply strong consequences on the behavior of the graph (Newman, 2003).

6. Conclusions

In this chapter we have illustrated several basic tools for representing and analyzing dynamic random graph in a general context and a practical approach to estimate the parameters of classical models of such behaviors. The generalized model that we have proposed can describe not only classical real-world networks models but also situations with more complex constraints. In this model, random graph dynamics is outlined by introducing a random time process where the dynamic graph topology changes are recorded. Among the advantages of this model is the possibility to generate successive independent random graphs with potentially different sets of vertices but all belonging to the same basic set of vertices. Otherwise, the fact that the generalized model allows to consider probabilities which are not necessarily all equal gives the possibility to favor the establishment of certain connections over others. This kind of behavior is often observed in

wireless mobile networks and social networks. Here, there are ways to implement in the generalized model the ability to prioritize connections to closest vertices. These advantages of modeling are generally not possible through the traditional random graphs of Erdös-Rényi. Also, we have described the global behavior between two consecutive configurations by calculating the probability of change. This is done in the case where the dynamic change concerns only the graph edges, as well as in the case where the dynamic change affects also the graph vertices.

At the end of this chapter we have shown how the parameters of a dynamic random graph can be estimated under the assumptions of the Erdös-Rényi model. Thus, the estimation of these parameters has led to the estimates of the average degree of vertices, the average hopcount and the size of the giant component in large dynamic graphs.

7. References

Albert R., Jeong H., and Barabàsi A.L. (1999). Diameter of the Worldwide Web. Nature, vol. 401, pp.130–131.

Amaral L.A.N., Scala A., Barthélémy M., and Stanley H.E. (2000). Classes of small-world networks. Proc. of Nat. Acad. Sci. USA 97, pp. 11149–11152. http://polymer.bu.edu/hes/articles/asbs00.pdf

Bollobàs B. (2001). Random Graphs. Cambridge studies in advanced mathematics.

Barabàsi A. L. and Albert R. (1999). Emergence of scaling in random networks, Science, vol. 286, pp. 509–512.

Bhamidi S., van der Hofstad R. and Hooghiemstra G., (2010). First passage percolation on random graphs with finite mean degrees. The Annals of Applied Probability, Vol. 20, No. 5, pp. 1907–1965, DOI: 10.1214/09-AAP666

Durrett R. (2006). Random Graph Dynamics. Cambridge University Press. Cambridge, U.K.

Euler, L. (1736). Solutio problematis ad geometriam situs pertinentis. (The solution of a problem relating to the geometry of position). Commentarii Academiae Scientiarum Imperialis Petropolitanae, vol. 8, pp. 128–140.

Erdös P. and Rényi A. (1960). On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences, vol. 5, pp. 17–61.

Faloutsos M., Faloutsos P., Faloutsos C. (1999). On the power-law relationship of the internet topology. Proceedings of SIGCOMM '99, New York, USA, pp 251–262 http://www.cis.upenn.edu/~mkearns/teaching/NetworkedLife/power-internet.pdf

Hamlili A. (2010). Adaptive Schemes for Estimating Random Graph Parameters in Dynamic Random Networks' Modeling, Proceedings of IEEE/IFP Wireless days, Venice-Italy, 20 - 22 October, 2010.

Hekmat R. and Van Mieghem P. (2003). Degree distribution and hopcount in wireless ad-hoc networks. Proceedings of the 11th IEEE International Conference on Networks (ICON 2003), Sydney, Australia, pp. 603–609, Sept. 28-Oct. 2003.

Hewer T., Nekovee M. and Coveney P. V. (2009). Parameter exploration in parallel for dynamic vehicular network efficiency, International Conference on Advances in Computational Tools for Engineering Applications Proceedings (ACTEA '09), pp. 16 – 20, 15-17 July 2009.

Janson S., Lucak T., and Rucinski A. (2000). Random Graphs, John Wiley and Sons, New York.

Jurdak R. (2007). Wireless Ad Hoc and Sensor Networks: A Cross-Layer Design Perspective, Springer.

Kawahigashi H., Terashima Y., Miyauchi N. and Nakakawaji T. (2005). Modeling ad hoc sensor networks using random graph theory, Proceedings of Consumer Communications and Networking Conference (CCNC'05), pp.104 – 109.

Lehmann E. L. and Casella G. (1998). Theory of Point Estimation. Second Edition, Springer.

Mieghem P.V., Hooghiemstra G., van der Hofstadh R. (2000). A Scaling Law for the Hopcount, Report 2000125. Delft University of Technology, Delft, The Netherlands. http://wwwtvs.et.tudelft.nl/people/piet/telconference.html

Molloy M. and Reed B. (1998). The size of the giant component of a random graph with a given degree sequence, Combinatorics, Probability and Computing, vol. 7, pp. 295–305.

Newman M. E. J., Strogatz S. H., Watts D. J. (2001). Random graphs with arbitrary degree distributions and their applications. Physical Review E 64, 026118. http://arxiv.org/PS_cache/cond-mat/pdf/0007/0007235v2.pdf

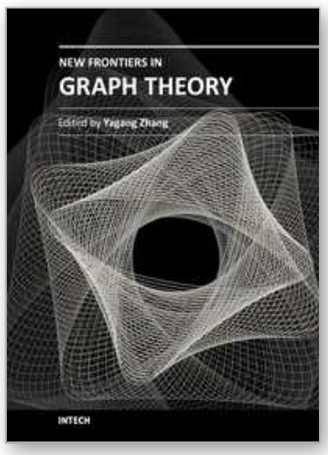
Newman M. E. J. (2003). Random graphs as models of networks, in Handbook of Graphs and Networks, Bornholdt S. and Schuster H. G. (eds.), Wiley-VCH, Berlin.

Onat F. A., Stojmenovic I. and Yanikomeroglu H. (2008). Generating random graphs for the simulation of wireless ad hoc, actuator, sensor, and internet networks, Pervasive and Mobile Computing, Vol. 4, Issue 5, pp. 597-615.

Penrose M. D. (1999). On *k*-connectivity for a geometric random graph, Random Structures and Algorithms, vol. 15, pp. 145–164.

Steele M. (1997). Probability theory and combinatorial optimization. Vol. 69, SIAM.

Trullols O., Fiore M., Casetti C., Chiasserini C. F., and Barcelo Ordinas J. M. (2010). Planning Roadside Infrastructure for Information Dissemination in Intelligent Transportation Systems. Comput. Commun., 33(4):432-442.



New Frontiers in Graph Theory

Edited by Dr. Yagang Zhang

ISBN 978-953-51-0115-4
Hard cover, 526 pages
Publisher InTech
Published online 02, March, 2012
Published in print edition March, 2012

Nowadays, graph theory is an important analysis tool in mathematics and computer science. Because of the inherent simplicity of graph theory, it can be used to model many different physical and abstract systems such as transportation and communication networks, models for business administration, political science, and psychology and so on. The purpose of this book is not only to present the latest state and development tendencies of graph theory, but to bring the reader far enough along the way to enable him to embark on the research problems of his own. Taking into account the large amount of knowledge about graph theory and practice presented in the book, it has two major parts: theoretical researches and applications. The book is also intended for both graduate and postgraduate students in fields such as mathematics, computer science, system sciences, biology, engineering, cybernetics, and social sciences, and as a reference for software professionals and practitioners.

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen