http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (4): 904-907, 2011
ISSN 1812-5638 / DOI: 10.3923/itj.2011.904.907

A Novel Method to Construct Undirected Network

Lijun Tian, Chaoqun Liu and Guang Sun
Department of Information Technology, Hunan University of Finance and Economics,
No.139, Fenglin 2nd Road, Changsha, 410205, China

Abstract: This study presented an Undirected Network Construction Method. UNCM initializes a network with one vertex, zero edge and adds vertices one by one. Every time a vertex was added, a sequence of random variables $\xi_{n,r}$ was given to determine the presence or absence of edges. $\xi_{n,r}$ has a (0-1) distribution with probability $p_{n,r}$ independently. There are three types of probability sequence $p_{n,r}$: specified, completely random and constrained random. By using a specified sequence, regular networks or other specific network can be constructed and by completely random sequence, some types of networks such as Poisson random graphs can be constructed. To construct more complex networks, a constrained random sequence is needed. As an example, configuration model was discussed. Theoretical analysis shows that any types of undirected networks can be constructed by UNCM, which means the research on how to construct such networks can be converted to the research on the characteristics of probability sequences $p_{n,r}$.

Key words: UNCM, network construction, complex network, undirected network, random graph

INTRODUCTION

In the past decade, complex network, leading to a new and emerging scientific discipline called as network science, has become a hot research field which stimulated a great deal of interest in studying of it and brought together researchers from many areas including computer science, sociology, physics, biology and others (Wang and Chen, 2003; Newman *et al*., 2006).

Euler's solution of the Königsberg bridge problem is often cited as the first true proof in the theory of networks. The research of random graphs by Paul Erdös and Alfréd Rényi is considered to the beginning of complex network study (Albert and Barabasi, 2002; Strogatz, 2001). However, the really cause of dramatic advantages in this field is the discovery of small-world (Watts and Strogatz, 1998) and scale-free (Barabasi and Albert, 1999) properties of complex networks.

The research on networks mainly aims to: find statistic properties of real-world networks such as computer networks and social networks; create models of networks; predict what the behavior of networked systems (Newman, 2003).

Although it has made remarkable progress, the research of networks is still on its early stages of development and there are many problems to be solved. Lack of effective and common models of networks that can simulate real-world networks exactly is one of these problems.

In this study, we proposed a method to construct undirected complex networks such as Poisson random graphs, generalized random graphs, small-world networks, scale-free networks, and any types of other undirected networks.

CONSTRUCTION METHOD

A network (or graph) is a set of vertices interconnected by edges (Fig. 1). Figure 1a shows a network with different type of vertices (1, 3, 4, 5 or 2, 6), different type of edges (solid or dotted, directed or undirected line), self-cycle (L), multiple edge (H and I).

Figure 1b shows a simple network which has only one type of vertices interconnected by one type of edges, without self-cycle and multiple edges, not to mention hyperedges (that join more than two vertices together).

Network showed in Fig.1b is often called as undirected graph or undirected network which is the most

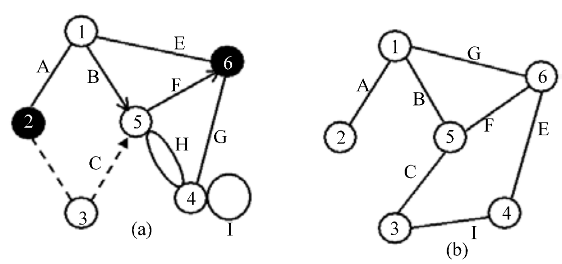

Fig. 1: Example of various types of network

Corresponding Author: Lijun Tian, Department of Information Technology, Hunan University of Finance and Economics, No.139, Fenglin 2nd Road, Changsha, 410205, China

important type of complex networks. Many well-known models, such as Poisson random graph, configuration model (Newman *et al*., 2001), WS model (Watts and Strogatz, 1998), BA model (Barabasi and Albert, 1999) etc., all are model of undirected networks. Here, our method deals only with undirected networks, so it is called as Undirected Networks Construction Method (UNCM).

UNCM's basic ideas are using a growth model, adding vertex one by one. As showed in Fig. 1b, we sort all vertices one by one and construct the network by this order. For a network with N vertices, there are N! orders (6! in Fig.1b). But for the homogeneity of the network, all the networks constructed by N! orders are the same one, so we can choose any one of it. Every time a vertex was added, a sequence of random variables is set to determine the presence or absence of edges.

More formally, UNCM is described as follows (to construct a network with a given number N of vertices):

Step 1: Starting with a vertex, without any edges

Step 2: For a number n (initial value set to 2), if $2 \le n \le N$, add a vertex

Step 3: Select a sequence of random variables: $\xi_{n,1}$, $\xi_{n,2}$, ..., $\xi_{n,n-1}$, which values is given by the following formula:

$$\xi_{n,r} = \begin{cases} 1 & p_{n,r} \\ 0 & 1-p_{n,r} \end{cases} \qquad 2 \le n \le N, 1 \le r \le n-1 \tag{1}$$

That is, random variable $\xi_{n,r}$ obeys (0-1) distribution with probability $p_{n,r}$ independently. In mathematics, there is $0<p_{n,r}<1$. In particular, we specified if $p_{n,r} = 1$, there is $\xi_{n,r} = 1$, and if $p_{n,r} = 0$, there is $\xi_{n,r} = 0$. $\xi_{n,r} = 1$ means that the edge between vertex n and vertex r is present, and $\xi_{n,r} = 0$ means the edge is absent

Step 4: n plus 1, if $n \le N$ go to Step 2, else finish the construction

Some properties of network constructed by UNCM such as degree of vertex, mean degree, etc. are given by formulas as follows.

Degree of vertex i while network have n vertices in total can be computed by formula (2):

$$\deg_{n,i} = \sum_{r=1}^{i-1} \xi_{i,r} + \sum_{r=i+1}^{n} \xi_{r,i} \qquad 1 \le i \le n, 2 \le n \le N \tag{2}$$

Total number of edges |E| can be computed by formula (3):

$$|E| = \sum_{j=2}^{n} \sum_{r=1}^{j-1} \xi_{j,r} \qquad 2 \le n \le N \tag{3}$$

Because total degree of all n vertices equals to 2×|E|, the mean degree z_n can be computed by formula (4):

$$z_n = \frac{2 \times |E|}{n} = \frac{2}{n} \sum_{j=2}^{n} \sum_{r=1}^{j-1} \xi_{j,r} \qquad 2 \le n \le N \tag{4}$$

Set M = n(n-1)/2, where M is the maximum possible number of edges, the mean probability P_n is defined as formula (5):

$$P_n = \frac{|E|}{M} = \frac{|E|}{n(n-1)/2} = \frac{2 \times |E|}{n(n-1)} = \frac{z_n}{n-1} \tag{5}$$

UNCM is similar to BA model (B) in some respects, but there are differences between them:

- UNCM start with a vertex, while BA start with a small number (m_0) of vertices
- UNCM add a vertex n with i ($1 \le i \le n-1$) edges which determined by the probability sequence p(n, r), while BA with a fixed number $m(\le m_0)$
- The most important difference is UNCM add edges by a sequence of random variables, while BA use a mechanism called preferential attachment. By selecting a specific sequence, any types of networks can be constructed by UNCM, which will be illustrated in the remaining sections

SPECIFIED SEQUENCE

According to UNCM, if $p_{n,r}$ (=1 or =0), then $\xi_{n,r}$ (=1 or =0). Actually, $\xi_{n,r}$ is no longer a random variable in this case. But we can construct any types of regular graphs or other specific types of networks by using specified sequence. Some examples are described as follows:

- By specifying $p_{n,r}$ as formula (6), we can construct a zero graph with N vertices and zero edges

$$p_{n,r} = 0 \quad 2 \le n \le N, 1 \le r \le n-1 \tag{6}$$

- By specifying $p_{n,r}$ as formula (7), we can construct a complete graph with N vertices and edges

$$p_{n,r} = 1 \quad 2 \le n \le N, 1 \le r \le n-1 \tag{7}$$

- By specifying $p_{n,r}$ as formula (8), we can construct a star graph with N vertices and N-1 edges

N-1 edges

$$p_{n,r} = \begin{cases} 1 & 2 \le n \le N, r = 1 \\ 0 & 3 \le n \le N, 2 \le r \le n-1 \end{cases} \quad (8)$$

- It is more complex to construct a regular ring lattice, which is a network of N vertices, each connected to its 2K nearest neighbours by undirected edges. A lattice can be constructed by specifying $p_{n,r}$ as formula (9).

$$p_{n,r} = \begin{cases} & 2 \le n \le K, 1 \le r \le n-1 \\ 1 & K+1 \le n \le N, n-K \le r \le n-1 \\ & N-K+1 \le n \le N, 1 \le r \le n+K-N \\ 0 & \text{others} \end{cases} \quad (9)$$

COMPLETELY RANDOM SEQUENCE

Structures of regular graphs constructed by UMCN above are completely determined, without any randomness. We can construct random graphs by using completely random sequence which is defined as follows: for a specific number $n(2 \le n \le N)$, $p_{n,r}$ independent of the network with n-1 vertices constructed in the previous steps.

As an example, we construct a completely random graph by specifying $p_{n,r}$ as formula (10):

$$p_{n,r} = P \quad 2 \le n \le N,\ 1 \le r \le n-1,\ 0 < P < 1 \quad (10)$$

According to Eq. 10, all $\xi_{n,r}$ have the same probability P. That is, any vertices pair i and j ($i \le j$) connect each other with probability P independently, which is completely equivalent to the definition of ER model $G_{n,p}$.

In this case, p_k, the probability of a vertex having degree k obey the binomial distribution and approximately Poisson distribution while N tends to infinity. So the network constructed by this way is also known as Poisson random graph.

We can construct other types of random graphs by using completely random sequence, e.g., a non-equal probability sequence. However, not all types of random graphs such as those by the configuration model can be constructed by this way.

CONSTRAINED RANDOM SEQUENCE

The configuration model networks cannot be constructed by UNCM using completely random sequence, because the sequence of it is not completely random, which is limited to some constraints.

Constrained random sequence is defined as follows: for a specific number $n(2 \le n \le N)$, $p_{n,r}$ constrained by the network with n-1 vertices constructed in the previous steps.

Usually, it is constrained by the degree of vertices. So, constrained random sequence can be also defined as: $p_{n,r}$ is a function of vertex degree $deg_{n-1,i}$.

As an example, we construct a network with N vertices by the configuration model using constrained random sequence. The procedure is described as follows:

- According to the configuration model (Newman *et al*., 2001), specify a degree distribution p_k, then choose a degree sequence, which is a set of N values of the degrees k_i of vertices i=1...N. Each vertex i is given a degree, which is number of "stubs" or "spokes" in the configuration model
- If a network with $n-1(2 \le n \le N)$ vertices was constructed, compute $deg_{n-1,i}$, the degree of each vertex i by formula (2). means that vertex i can add edges, while means cannot add any more
- Set a variables sequence which means if vertex r can add edges or not, is defined as formula (11):

$$\zeta_{n,r} = \begin{cases} 1 & deg_{n-1,r} < D_{k_r} \\ 0 & deg_{n-1,r} = D_{k_r} \end{cases} \quad 2 \le n \le N, 1 \le r \le n-1 \quad (11)$$

Then we can compute $p_{n,r}$ as formula (12):

$$p_{n,r} = \begin{cases} p_n & \zeta_{n,r} = 1 \\ 0 & \zeta_{n,r} = 0 \end{cases} \quad 2 \le n \le N, 1 \le r \le n-1 \quad (12)$$

Probability p_n is no longer equal to p_k, because some of the vertices cannot add any edges any more. Set s is the number of vertices that can add edges and t is the number of edges possible to add according to probability p_k.

Clearly, there is

$$s = \sum_{r=1}^{n-1} \zeta_{n,r},\ t = \sum_{r=1}^{n-1} \xi_{n,r} \quad 2 \le n \le N$$

then p_n can be computed as formula (13):

$$p_n = \binom{t}{s} (p_k)^t (1-p_k)^{s-t} \quad 2 \le n \le N \quad (13)$$

Using a similar approach, we can compute sequences of probability $p_{n,r}$ to construct any other type of networks such as NS model, BA model, etc.

THEORETIC ANALYSIS AND DISCUSSION

Why any types of networks can be constructed by UNCM? Here, we analyze its mechanism.

Any network has a form of adjacency matrix $A=[a_{ij}]_{N\times N}$, which is completely equivalent to each other. For a network as Fig. 1b showed, its adjacency matrix is A(1b). Because the network is undirected without self-cycle, A(1b) is a symmetric matrix, diagonal elements $a_{ii}=0$, so the network is equivalent to an upper triangular matrix B(1b).

$$A(1b)=\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} B(1b)=\begin{bmatrix} & 1 & 0 & 0 & 1 & 1 \\ & & 0 & 0 & 0 & 0 \\ & & & 1 & 1 & 0 \\ & 0 & & & 0 & 1 \\ & & & & & 1 \\ & & & & & \end{bmatrix}$$

We use random variables $\xi_{n,r}$ to replace 1 or 0, so any network G with N vertices equivalent to a upper triangular matrix of random variables B(G). In B(G), column n indicates the vertex n, where $\xi_{n,r}$ indicates the presence or absence of edge between vertex n and r ($1 \le r \le n-1$).

$$B(G)=\begin{bmatrix} \xi_{2,1} & \xi_{3,1} & \cdots & \xi_{N-1,1} & \xi_{N,1} \\ & \xi_{3,2} & \cdots & \xi_{N-1,2} & \xi_{N,2} \\ & & & \vdots & \vdots \\ & 0 & & \xi_{N-1,N-2} & \xi_{N,N-2} \\ & & & & \xi_{N,N-1} \\ & & & & \end{bmatrix}$$

$$P(G)=\begin{bmatrix} p_{2,1} & p_{3,1} & \cdots & p_{N-1,1} & p_{N,1} \\ & p_{3,2} & \cdots & p_{N-1,2} & p_{N,2} \\ & & & \vdots & \vdots \\ & 0 & & p_{N-1,N-2} & p_{N,N-2} \\ & & & & p_{N,N-1} \\ & & & & \end{bmatrix}$$

Because $\xi_{n,r}$ obey (0-1) distribution with probability $p_{n,r}$ independently, there is a matrix of probability P(G).

So, any type of network can be constructed by UNCM using a sequence of probability $p_{n,r}$. That is, to create any model of networks, the only thing we have to do is that select a sequence $p_{n,r}$, which is easier to understand and study.

CONCLUSION

Undirected network is one of most important type of complex network. In this study, we proposed a method called as UNCM to construct such networks. Firstly, we defined UNCM. UNCM uses a growth model, which initializes a network with one vertex and adds one vertex at each step. Every time a vertex was added, a sequence of random variables $\xi_{n,r}$ was given to determine the presence or absence of edges. $\xi_{n,r}$ has a (0-1) distribution with probability $p_{n,r}$ independently. Then, we constructed several networks as examples by using three types of $p_{n,r}$: specified, completely random and constrained random. Finally, we analyzed UNCM with adjacency matrix of networks which showed it can construct any types of networks.

That is, the research on model of networks can be converted to the research on the probability sequences $p_{n,r}$. In the future, we will study the characteristics of the sequences to construct networks of existing model, and to discover new model.

ACKNOWLEDGMENTS

This study was support by Innovative Research Project of Hunan (2010244429).

REFERENCES

Albert, R. and A.L. Barabasi, 2002. Statistical mechanics of complex networks. Rev. Mod. Physi., 74: 47-97.

Barabasi, A.L. and R. Albert, 1999. Emergence of scaling in random networks. J. Sci., 286: 509-512.

Newman, M.E.J., S.H. Strogatz and D.J. Watts, 2001. Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E, 64: 026118-026134.

Newman, M.E.J., 2003. The structure and function of complex networks. Soc. Indust. Applied Mathe. Rev., 45: 167-256.

Newman, M.E.J., A. Barabasi and D.J. Watts, 2006. The Structure and Dynamics of Networks. Princeton University Press, Princeton, New Jersey, ISBN-13: 978-0-691-11357-9.

Strogatz, S.H., 2001. Exploring complex networks. Nature, 410: 268-276.

Wang, X.F. and G. Chen, 2003. Complex networks: Small-world, scale-free and beyond. IEEE Circ. Syst. Magaz., 3: 6-20.

Watts, D.J. and S.H. Strogatz, 1998. Collective dynamics of small-world networks. Nature, 393: 440-442.