Research Journal of Applied Sciences, Engineering and Technology 5(2): 568-573, 2013
ISSN: 2040-7459; E-ISSN: 2040-7467
© Maxwell Scientific Organization, 2013
Submitted: May 18, 2012 Accepted: June 15, 2012 Published: January 11, 2013

A Method for Measuring Node Importance in Hypernetwork Model

Quan Xiao
Department of Information Technology, Jiangxi Key Laboratory of Data and Knowledge Engineering, Jiangxi University of Finance and Economics, China

Abstract: This study is aiming at the central challenge for network analysis and application of how to identify key nodes in hypernetwork representations. Although there are many measuring methods based on traditional network, most of them use single criteria, which is often confronted with the problem of incomplete or inaccurate information on the structure of hyper network. In this study, a new multi-criteria measuring method with adjustable parameters is proposed, which takes the node degree, star degree and betweenness into considerations. According to the experimental results with a numerical example of hypernetwork, our method shows good performance on discrimination and precision to evaluate the importance of nodes as compared with several existing measuring methods for nodes in traditional network.

Keywords: Hypernetwork, multi-criteria, node importance measuring

INTRODUCTION

It is no exaggeration to say that we are living in kinds of networks every minute due to the development of information technology. Social network, computer network, traffic network, knowledge network and even contagion transmission network are surrounding us, making us unable to escape (Newman, 2010). Based on different network, researchers have been studying gradually from the shallower to the deeper and there occurred a lot of important developments in many fields including mathematics, physics, computer and information sciences, biology and the social sciences. For instance, "how does knowledge spread and deliver between humans", "why rumors diffuses faster than the truth", "how does regional or global financial crisis emerge and sprawl", "what causes traffic jams in metropolises and how to avoid them" are some of the hot topics recently discussed. From these topics we can see that on the one hand, humans are thinking how to utilize network to promote the flowing of persons, goods, capitals, information and finally the efficiency of production and management, while on the other hand, it is widely considered how to avoid and reduce hazards of pandemic diseases or big electricity cut caused by the out-of-control of networks.

Along with the rapid development of network theories, the constructed networks specific to different application background are becoming progressively more complex over time (Slotte-Kock and Coviello, 2010): the number of nodes increases, their types vary and the connections between nodes are diversified. Additionally, nodes and edges may have complicated dynamic behaviors such as aggregation, nesting, or reduction etc. In this situation, the relations between objects described by nodes are not limited to be binary, but might be in the form of multi-elemental subset clusters. While sometimes, the isomerism, hierarchy and granularity of objects in network need to be considered. The increasing complexity of network makes it difficult to describe the network feature of real world by ordinary graph in more and more cases. Taking scientific collaboration network for example, by ordinary graph, we can easily represent whether two authors have co-author relationships by two nodes and a line join them up. But, it is difficult to describe the situation that more than three authors write one paper collaboratively. Bipartite graph may be a feasible representation, but it breaks the homogeneity of nodes through the distinction between "author nodes" and "paper nodes", which will result in the ineffectiveness of some parameters by the generated ambiguity when calculating connectivity, centrality and other topological properties (Estrada and Rodriguez-Velazquez, 2005).

Hypernetwork makes it possible for researching on the relationships in and between complex idiosyncratic networks, the concept of which was first proposed by Sheffi (1985) and hypernetwork is regarded as "network above the existing network". By extending network to hypernetwork, it is able to describe the multi-elemental, multilayer, multilevel and multiattribute of objects and significantly enriches the represent ability of the traditional network. For the present, supply chain hypernetwork, financial hypernetwork, power supply hypernetwork, population migration hypernetwork, interpersonal hypernetwork

and knowledge hypernetwork has been proposed and studied by many researchers (Wu and Cai, 2008).

Whether in the traditional network or in hypernetwork, it has been an issue worth studying in the fields of network analysis and system science to evaluate the importance of nodes and discover the critical ones. For example, who takes the authoritative position in a research field in scientific collaboration (hyper) network? How to identify and protect the hidden danger points in (hyper) network security? How to distinguish the person (s) whose dimission may bring about great losses of knowledge in knowledge sharing (hyper) network and then make policies to retain the talented person (s). In the control process of virus and disease, the infection of which individuals or groups will lead to pandemic diseases? The theoretical solutions of the problems above can be equivalent to the designing of the method for measuring node importance in (hyper) network aiming at different system features. The objective of this study is to investigate the method for measuring node importance in hypernetwork model, which may help answering the questions above and furthermore make some contributions to hypernetwork theories and applications.

LITERATURE REVIEW

Corley and Sha (1982) firstly propose the problem of "most vital node" (MVN). They define the MVN between two nodes as the node that results in the maximum increment of the distance between the two nodes when removed. But regrettably, this method is incapable of the evaluation of node importance in the whole network globally. In recent years, the blossom of empirical study in network, especially the discovery of "small world" and "scale free", makes it reasonable to evaluate the importance of network nodes on the basis of their otherness. The method of measuring node importance in network arisen at present can be classified into two types as follows.

Social Network Analysis (SNA) based methods: Major methods for measuring node importance based on SNA are under the assumption that the importance of a node is equivalent to its notability when connected to other nodes (Knoke and Burt, 1983). These methods generally study the metrics while maintaining the integrity of network, i.e., do not break the connectivity. The basic idea is to find some useful attributes (such as quantity of information contain in degree or shortest path) to highlight diversity between nodes, in other words, to adequately display the positional traits of nodes in a network and to "magnify" the significance to define the importance. Proposed metrics based on SNA include two categories as "centrality" and "prestige", while indicators include degree, closeness, betweenness, eigenvector, cumulated nomination etc. Many representative works have been done about the SNA based methods by Bonacich (1987), Altmann (1993) and Poulin *et al*. (2000). Typically, Lee *et al*. (2010) studied the suitability of degree and betweenness as indicators for the influence of individual customers on the behavior of the entire customer base. Gneiser *et al*. (2010) analyzed different centrality measures and propose a measure which is based on the Page Rank algorithm. Jun *et al*. (2010) proposed a unify multiple metrics for evaluating framework of node importance with non-conflict equivalent classes. In a word, these methods can be considered as topology-based.

Removing based methods: To evaluate the importance of a node in network by measuring the destructiveness to the network performance when remove it is another thinking that is "destructiveness equals to importance". The greater the extent of severity when we remove a node is, the more important it is. Because the maintenance of network connectedness or system function depends on its existence. Network efficiency and contraction are typical removing based metrics for node importance measuring. Li *et al*. (2004) took the opinion that to destroy node with different distance would bring different losses to network and assumed the losses to vary inversely with the distance between a pair of node. They took the inverse of distance between node pair as weight value and then calculated the weighted sum of all disconnected node pairs to measure the destructiveness to the network connectivity. This study made a quantized contribution to the "removing based" thinking. An *et al*. (2006) analyzed the requirements of an effective importance measurement for nodes in a network and modeled the comprehensive importance measure for nodes in a node-weighed network based on the approach of deleting nodes. With the help of the procedure of calculating the distances among nodes on a graph, a new algorithm for this model is designed. Leung and Tanbeer (2012) proposed a novel algorithm to mine social networks for significant friend groups, which first constructs a significant friend-tree (SF-tree) to capture important information about linkage between users in the social networks and then uses the SF-tree to find significant friend groups among all friends of users in the social networks by removes friends with low climp values. In general, it is necessary for removing based methods to settle the problems of topological changes even the segmentation of network arise by the nodes removed.

Literature analysis: In different methods for measuring node importance, common metrics include: degree, betweenness, cohesion, subgraph, spanning tree count and average shortest distance etc. Each indicator focuses on a particular feature of network, e.g., the

position of node in network structure, the influence and control capability of node to information spreading, or the contribution of node in substructure. Degree is the simplest method, but it only emphasizes the number of edges connected with adjacent nodes, which ignores the indirect effects from them. Betweenness is used to describe the informational control capability of a node to the connection between other nodes, but it is unable to mediate its partial contribution. Cohesion prefers the geometrical position in a network, but the distinction between nodes in the same hierarchy is not very obvious. Sub graph based methods reflect the contribution to the partial network structure, but the stability of sub graph under different weighting cannot be guaranteed. Removing based methods is difficult to measure the node importance when the influence of removing a node to the network performance is too little or too much.

Through the analysis above, we conclude that measuring the node importance under a single indicator will rely seriously on certain aspect of the network characteristics, which ignores the role of other indicators. Moreover, the measuring results on the basis of static network are unlikely to take account of the dynamic relationship between network nodes. Different evaluation methods in some cases provide considerably different results in terms of the importance of nodes (Landherr *et al*., 2010). Consequently, during the process of practices and studies, usually multiple approaches are adopted to measure the importance of the same node corporately.

DEFINITIONS AND METHODS

Related definitions and notations: To facilitate the description of our method, we give the definitions to be appeared in the measuring model first as bellows:

Hyper network: There are two different definitions to hyper network. The first one defines hyper network as "network in network", while the other one considers hyper network as "network represented by hypergraph". We take the latter one in this study and refer Berge's basic definition of hypergraph (Berge, 1987) to define hyper network as follows:

Definition 1: A hypernetwork HN = (V, E) on a set V is a family e_j of non-empty subsets of V called hyperedges, where $V = \{V_1, V_2, \dots, V_n\}$ is a finite set of nodes, $E = \{e_1, e_2, \dots, e_m\}$ is a family of subset of V, s.t. $e_j \neq \emptyset$, $j = 1, 2, \cdots, m$ and $\bigcup_j e_j = V$.

Fig. 1 shows an example of hypernetwork, where,

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8\}$$

$$E = \{\{v_1, v_2, v_3\}, \{v_2, v_3\}, \{v_2, v_4, v_5\}, \{v_1, v_7\}, \{v_8\}\}$$

In the hypernetwork diagram, the edge e_j is represented by a solid line surrounding its nodes; if $|e_j| = 2$ e_j can be represented by a solid line connecting the two nodes. The hypernetwork (V, E) can also be denoted by $(V; e_1, e_2, \dots, e_m)$. The hypernetwork discussed in this study is undirected hypernetwork without isolated nodes.

Besides the forms of set and diagram, in scientific or engineering calculation tasks, hypernetwork is often handled as matrix. Adjacency matrix $AM_{n\times n}$ reflects the adjacent relations between hypernetwork nodes, where the element am_{ij} in AM is the adjacent value of nodes v_i and v_j. Incidence matrix $IM_{m\times n}$ describes the relations between nodes and hyperedges, where $im_{ij} = 1$ if node v_j is an element in e_i, or else $im_{ij} = 0$.

Node degree and star degree: We give the definition of node degree and star degree as follows:

Definition 2: The node degree $d_N(v_i)$ of node $v_i \in V$ in hypernetwork $HN = (V, E)$ is the number of nodes connected with v_i. The definition is the same as that in traditional network and node degree can be calculated as: $\overrightarrow{d_N} = AM_{n\times n} \times \overrightarrow{e_{n\times 1}}$, where $\overrightarrow{e_{n\times 1}} = \{1,1,\cdots,1\}^T$ and each element $d_N(v_i)$ in $\overrightarrow{d_N}$ is the node degree of v_i, which is normalized as:

$$d'_N(v_i) = \frac{d_N(v_i)}{n-1}$$

Definition 3: For node $v_i \in V$ in hypernetwork $HN = (V, E)$, the star $S(v_i)$ of v_i is the partial hypergraph derived from all hyperedges $e_j \in E$ that contain v_i. The star degree $d_S(v_i)$ of v_i is the number of hyperedges in $S(v_i)$, which can be calculated as:

$$\overrightarrow{d_S} = \overrightarrow{e_{1\times m}} \times IM_{m\times n}$$

where, $\overrightarrow{e_{1\times m}} = \{1,1,\cdots 1\}$ and element $d_S(v_i)$ in $\overrightarrow{d_S}$ denotes the star degree of v_i, which is normalized as:

$$d'_S(v_i) = \frac{d_S(v_i)}{n-1}$$

For instance, the star of node v_2 in Fig. 1 is $S(v_2) = \{\{v_1, v_2, v_3\}, \{v_2, v_3\}, \{v_2, v_4, v_5\}\}$ and its star degree $d_S(v_2) = 3$.

Hyperpath:

Definition 4: Declare hyperpath $HP(v_1, v_{q+1}) = \{v_1, e_1, v_2, e_2, \cdots, e_q, v_{q+1}\}$ as an alternating sequence of node and hyperedge in hypernetwork $HN = (V, E)$ with length q when the conditions below are satisfied:

- $\{v_1, v_2, \dots, v_{q+1}\}$ is different from each other in HN = (V, E)

- $\{e_1, e_2, \dots, e_q\}$ is different from each other in HN = (V, E)
- $v_k, v_k+1 \varepsilon e_k, k = 1, 2, \dots, q$

Betweenness:
Definition 5: Betweenness is the fraction of shortest hyperpaths between node pairs that pass through the node of interest. The betweenness $B(v_i)$ of node v_i in hypernetwork HN = (V, E) is calculated as follows:

- For each pair of nodes (v_j, v_k), compute all shortest hyperpaths between them
- For each pair of nodes (v_j, v_k), determine the fraction of shortest hyperpaths that pass through v_i
- Sum this fraction over all pairs of nodes (v_j, v_k)

More compactly the normalized betweenness can be represented as:

$$B(v_i) = \frac{\sum_j^n \sum_k^n \frac{g_k(i)}{g_k}}{n(n-1)/2}, j \neq k \neq i, j < k$$

where,
g_k = Total number of shortest hyperpaths from node v_j to v_k
$g_k(i)$ = The number of those hyperpaths that pass through node v_i

Measuring method: From a social capital view, the amount of social capital owned by individual is determined by two factors:

- The first one is the scale of connections accessible for actor.
- The second one is the occurrence frequency of an individual in the social exchange activities between different peoples (Bourdieu, 1986).

Regard Fig. 1 as a social hypernetwork, the importance of node v_2 is not only directly bound up with its neighbors as v_1, v_3, v_4, v_5, but as well closely related to its associated hyperedges as e_1, e_2, e_3. Besides, if v_4, v_5 intent to communicate with v_1, v_3, or v_7, it is inevitable to ask v_2 for agency. Hence, there is indirect relevancy between the importance of v_2 and the number of individuals that exchange frequently via v_2. When we measuring the importance of individual v_2 in the whole hypernetwork, it is insufficient to count on only one indicator, but should take its adjacent nodes, associated hyperedges and its betweenness in the exchange of hypernetwork into account.

For these reasons, we propose a method for measuring the node importance in hypernetwork with adjustable parameters, which include node degree, star degree and betweenness. Denote I(v) as the importance of node v in hypernetwork and it is calculated as:

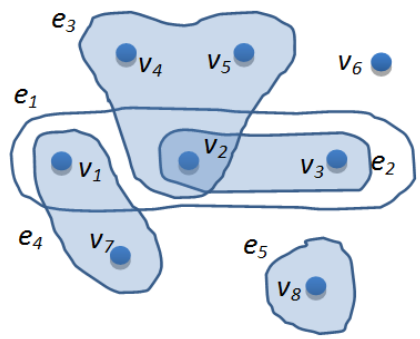

Fig. 1: An illustrative example of hypernetwork

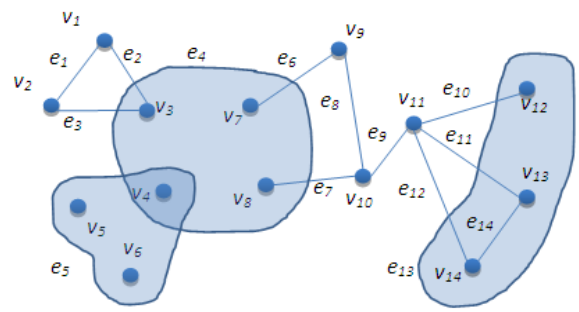

Fig. 2: A numerical example

$$I(v) = \alpha \cdot d_N(v) + \beta \cdot d_S(v) + \gamma \cdot B(v)$$

where, α, β, γ are adjustable parameters, which satisfy $\alpha + \beta + \gamma = 1$.

In order to facilitate contrast between different hypernetworks and eliminate the influence of hypernetwork size on numerical values, it is necessary to normalize indicators to make them in the domain of (0, 1). The normalize importance value of node v is computed as:

$$I'(v) = \frac{I(v)}{\sum_{i \in N} I(v_i)}$$

where,

$$\sum_{i \in N} I'(v_i) = 1$$

NUMERICAL EXAMPLE AND DICUSSION

To better illustrate the differentiation between the typical methods for measuring node importance in network, we take the hypernetwork in Fig. 2 for example to discuss the effectiveness and applicability of our measuring method proposed in this study. The experiment is conducted in the environment of Java, in April, 2012 at Jiangxi University of Finance and Economics.

Table 1: Comparison among node importance measuring methods for Fig. 2

ND		SD		BC		CC		PR		This study	
Node	Value	Node	Value	Node	Value	Node	Value	Node	Value	Node	Value
3	0.3846	11	0.3077	10	0.4679	8	0.5200	3	0.0852	10	0.3494
4	0.3846	3	0.2308	8	0.4167	10	0.4815	4	0.0852	11	0.3462
7	0.3077	10	0.2308	11	0.3846	3	0.4643	7	0.0750	8	0.3237
8	0.3077	13	0.2308	3	0.2821	4	0.4643	11	0.0750	3	0.2949
11	0.3077	14	0.2308	4	0.2821	7	0.4483	13	0.0732	4	0.2756
10	0.2308	4	0.1538	7	0.0833	9	0.4194	14	0.0732	7	0.1571
12	0.2308	7	0.1538	9	0.0321	11	0.3939	8	0.0728	13	0.1154
13	0.2308	8	0.1538	13	0.0000	1	0.3333	10	0.0705	14	0.1154
14	0.2308	12	0.1538	14	0.0000	2	0.3333	12	0.0668	12	0.0962
1	0.1538	1	0.1538	12	0.0000	5	0.3333	1	0.0651	9	0.0929
2	0.1538	2	0.1538	1	0.0000	6	0.3333	2	0.0651	1	0.0769
5	0.1538	9	0.1538	2	0.0000	13	0.3023	5	0.0651	2	0.0769
6	0.1538	5	0.0769	5	0.0000	14	0.3023	6	0.0651	5	0.0577
9	0.1538	6	0.0769	6	0.0000	12	0.3023	9	0.0630	6	0.0577

There are 14 nodes and 14 hyperedges in the hypernetwork of Fig. 2 and then we utilize 6 different methods, which include Node Degree (ND), Star Degree (SD), Betweenness Centrality (BC), Closeness Centrality (CC), PageRank (PR) and method of this study (take $\alpha = 0.25$, $\beta = 0.25$, $\gamma = 0.5$), to measure the importance of the 14 nodes in the hypernetwork. The calculated results and the corresponding sort of each node are listed in Table 1.

From the calculated results above we find that there are slight differences in the importance value and sort of nodes among these methods, due to their different emphasis on evaluation. Actually, there are certain correlations between them. In traditional network, a node has large degree usually has larger betweenness centrality, closeness centrality, or page rank values as well. However, in hypernetwork, the situation to be considered should make some difference.

In the hypernetwork of Fig. 2, nodes v_3 and v_4 are adjacent to 5 nodes separately, which have the largest node degree value, yet they locate in relatively marginal positions. The node degree of v_7, v_8 and v_{11} are all 4, but structurally v_{11} is obviously more important than v_7 and v_8 are the alternative to each other to some extent. Consequently, measuring node importance in hypernetwork solely relying on node degree is impracticable.

Taking the star degree into accounts, the star degree of node v_{10} and v_3 is 3 equally and this can be considered as they appear in three different groups respectively. But for the reasons of adjacent nodes and intermediary positions, the importance of v_{10} should be more than v_3.

Page Rank is a link analysis algorithm, named after Larry Page and used by the Google Internet search engine, which assigns a numerical weighting to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of measuring its relative importance within the set (WikiPedia, 2012). In traditional network, PageRank is an effective method for measuring the authority of web pages and individuals. But in hypernetwork, the calculated result by PageRank method shows that v_3 and v_4 are the most important nodes and the distinctions between node importance values are not obvious. That means PageRank method might not be appropriate for hypernetwork.

Betweenness centrality reflects the significance of node in social/information exchanges. The result shows that the importance value measured solely by betweenness centrality differs over-obviously and the BC value of 7 nodes are 0. For this reason, betweenness centrality should be utilized together with other metrics.

Closeness centrality is also a frequently-used method which indicates the distance between a node to others. But the calculated value shows insufficient differentiations.

According to the method proposed in this study, the sorting of node importance value in Fig. 2 from high to low is: v_{10}, v_{11}, v_8, v_3, v_4, v_7, v_{13}, v_{14}, v_{12}, v_9, v_1, v_2, v_5 and v_6. In contrast to the different measuring method above, the calculated result indicates that the proposed method for measuring node importance in hypernetwork model with adjustable parameters is able to highlight the differentiation between nodes, while reflecting the influence of node on the entire hypernetwork objectively. Moreover, the adjustable parameters expand the applied scope of the measuring method, to achieve instructive value to practical application under different hypernetwork modeling circumstances.

CONCLUSION

In this study, we propose a method for measuring node importance in hypernetwork model with adjustable parameters, which takes node degree, star degree and betweenness into considerations. Through comparing among the different node importance measuring methods, our measuring method has the advantages as follows: Firstly, it is capable of depicting

nuance of different nodes meticulously. Secondly, the parameters in the computational model can be changed according to different hypernetwork environment, which make our method more flexible. Finally, the method synthetically takes three typical indicators into account that avoids the one-sidedness of measuring based on only one indicator.

Although we hold the point that the method for measuring node importance in hypernetwork model proposed in this study is feasible and effective, there are still many issues that need further researches. Such as: the influence of different values selected for the adjustable parameters on the measuring result and how to judge and weigh the relations of these parameters, etc. All in all, for different hypernetwork application background, to choose suitable parameter combination rationally, this method will develop a greater value in practical hypernetwork applications

ACKNOWLEDGMENT

This study has been supported by Science and Technology Project of Education Department of Jiangxi Province under Grant No. GJJ11734 and Research Project of Jiangxi University of Finance and Economics under Grant No. 05462015.

REFERENCES

Altmann, M., 1993. Reinterpreting network measures for models of disease transmission [J]. Soc. Network., 15(3): 1-17.

An, S., Y. Du and J. Qu, 2006. A comprehensive importance measurement for nodes [J]. Chinese J. Manage. Sci., 14(1): 106-111(in Chinese).

Berge, C., 1987. Graph and Hypergraph [M]. North Holland, Amsterdam.

Bonacich, P., 1987. Power and centrality: A family of measures [J]. Am. J. Soc., 92(5): 1170-1182.

Bourdieu, P., 1986. The Forms of Capital [M]. In: Richardson, J.G. (Ed.), Handbook of Theory and Research for the Sociology of Education. Greenwood Press, New York.

Corley, H.W. and D.Y. Sha, 1982. Most vital links and nodes in weighted networks [J]. Oper. Res. Lett., 1: 157-160.

Estrada, E. and J.A. Rodriguez-Velazquez, 2005. Subgraph centrality in complex networks [J]. Phys. Rev. E, 71(5): 1-9.

Gneiser, M., J. Heidemann, M. Klier, A. Landheer and F. Probst, 2010. Valuation of online social networks taking into account users' interconnectedness [J]. Inf. Syst. E-Bus. Manage., 11: 1-24.

Jun, H., B. Wang and D. Lee, 2010. Evaluating Node Importance with Multi-criteria[C]. Proceeding of IEEE/ACM International Conference on Green Computing and Communications and IEEE/ACM International Conference on Cyber, Physical and Social Computing, pp: 792-797.

Knoke, D., R.S. Burt, 1983. Prominence [A]. In: Burt, R.S. and M.J. Minor (Eds.), Applied Network Analysis, a Methodological Introduction. Sage Publications, Beverly Hills, pp: 195-222.

Landherr, A., B. Friedl and J. Heidemann, 2010. A critical review of centrality measures in scocial networks [J]. Bus. Inf. Syst. Eng., 2(6): 371-385.

Lee, S.H.M., J. Cotte and T.J. Noseworthy, 2010. The role of network centrality in the flow of consumer influence [J]. J. Consum. Psychol., 20(1): 66-77.

Leung, C.K.S. and S.K. Tanbeer, 2012. Mining social networks for significant friend groups [C]. Database Syst. Adv. Appl. Lect. Not. Comput. Sci., 7240: 180-192.

Li, P., Y. Ren and Y. Xi, 2004. An importance measure of actors (set) within a network [J]. Syst. Eng., 22(4): 13-20, (In Chinese).

Newman, M., 2010. Networks: An Introduction [M]. Oxford University Press Inc., New York, USA.

Poulin, R., M.C. Boily and B.R. Masse, 2000. Dynamical systems to define centrality in social networks [J]. Soc. Network., 22(10): 187-220.

Sheffi, Y., 1985. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods [M]. Prentice-Hall Inc., Englewood Cliffs, NJ.

Slotte-Kock, S. and N. Coviello, 2010. Entrepreneurship research on network processes: A review and ways forward [J]. Entrep. Theory Pract., 34(1): 31-57.

Wu, Y. and S. Cai, 2008. Hypergraph Model of prior knowledge in opportunity discovery [C]. International Symposium on Knowledge Acquisition and Modeling, KAM, pp: 216-220.