Is a Stable and Connective Network Important to Innovation?

Gupeng Zhang[1], Jianghua Zhou[2], Guoqing Huo[1], Huicong Li[1]
[1]School of Public Policy and Management, University of Chinese Academy of Sciences, Beijing - People's Republic of China
[2]Business School, Beijing Normal University, Beijing - People's Republic of China

***Abstract*--With the patent co-inventing data of top 9 ICT firms with the highest patent application in China, this study establishes the co-inventing network and examines the moderate role of network connectivity, measured by classifying the individuals into two cohorts: inventors in the largest connected component and inventors in other isolated components. The network stability and innovation output demonstrate strong positive interactions, which is significant in not only the largest but also other isolated components. The clustering and centrality demonstrate significant effect on network stability and innovation output in the largest connected component, which is generally the same as that of extant studies. This impact is not significant in the other isolated components, which confirms the moderate role of network connectivity, i.e., fully connected networks constitute the basis for the network structure to be functioning. However, the significantly positive role of the structural hole is not moderated by the network connectivity. We discuss the contributions and implications of our findings.**

I. INTRODUCTION

The effects of individual mobility on knowledge transfer, innovation, and competitive advantage is increasingly becoming an important domain of research [10] [13] [19] [24] [25] [26]. Interorganizational mobility of individuals affects gains or losses in terms of the competitive advantage and performance outcomes (e.g., survival, profitability, effectiveness in head-to-head competition) of organizations that lose individuals [4] [18]. Therefore, most organizations are trying to curb the mobility and keep the stability of their employee groups, particularly the high-performers. Conversely, the employee's performance may also impact their stability. High-performers usually own high satisfaction with the current job, which makes them less likely to leave, while low-performers are more likely to seek outside opportunities. Although there are reciprocal effects [22], direct [11] and indirect [23] evidence suggests that the effect of employee stability on his/her performance is stronger than the reverse, which may be the main cause that most extant studies focused on the former. However, extant studies did not clearly examine to what extent the reciprocal effect is ignorable. Since there is reciprocal effect, the causal analysis of employee stability and performance should take it into account from both empirical and theoretical perspectives. As the employee's performance and stability interact with and function on each other, this study will make a comprehensive examination of the bidirectional causalities, which is one of the main contributions of this study to extant literatures.

In the context of an organizational network, as the network becomes more connected, distance between any two nodes diminishes, it is possible that information can become more democratized [1], information can thereby diffuse more quickly, fostering outcomes such as innovation or creativity [20] [21]. As the inventors' access to the information and knowledge is to a great extent dependent on the links with each other, the moderate effect of the network connectivity on the inventor stability and his/her performance are indispensable. Although the effect of network structure has been widely discussed by extant studies, e.g., [2] [3] [6] [15] [17] [19] [28], they are mostly based on the largest connected component within the whole network. As the disconnected components potentially conflate the influences of small-world structure and simple connection [9] and usually take a relatively small ratio compared with the largest component [5], most studies focused on the largest component, while ignored the methods to develop a weighted average across disconnected components proposed by [21]. However, besides the largest component, other components, e.g., the second and third largest, usually own well structured fabric. These components may also exhibit significant network effect, as the links constitute the base for inventor communication. Inventors with key positions may also have advantages in accessing information, and thereby generate higher innovation output in other smaller components. The specific inventive process may lead to the disconnections, e.g., pharmaceutical researchers are usually assigned to several groups, which are making mutually independent researches; technicians embarking at two different projects within the same firm may also lead to two isolated components. Obviously, inventors in the largest component represent only part of the firm's inventive activity. As the inventors in other components may also be doing important researches, ignoring these components may lead to a bias of the empirical results. In this sense, the network effects on network stability and performance, particularly in the partly connected contexts, deserves a further study. We will compare the differences of the network effect in the fully connected networks with that in partly connected networks, which formulates another main contribution of this study.

Additionally, extant studies provided only evidences that network connectivity is beneficial by proving that a greater ratio of the largest connected component positively impact innovation, e.g., [7] [9] [27]. As the linkages between individuals are the basic element constituting the network, greater extent of connectivity may be the key for the network indicators, e.g., clustering coefficient, centrality, path length, to be functioning on innovation. However, the moderate role of connectivity is not carefully examined by extant studies and will be another main job of this study.

II. DATA AND METHOD

A. Data

Because patent and patent statistics have been included in many research fields [12] and have been treated as the most important output indicator of innovation for their standardized information relating to new ideas and technological development, we use the patent data in constructing the networks. The patent co-inventing networks provide a rich opportunity to study the effect of network connectivity because these networks represent a primary conduit of information for inventors. Therefore, we use patent co-inventing data in establishing R&D cooperation networks. The characteristics of the R&D cooperation network are to a great extent reflected by the patent co-inventing network, which is widely used in studying the flow of information and R&D creativity, e.g., [8] [9].

We use the patents by the top 9 ICT firms that filed the largest number of patents for further analysis. These firms are: Huawei, ZTE, Panasonic, Sony, Intel, Philip, IBM, Samsung and LG. These 9 firms filed about 200 thousand patents in China. Because an inventor may appear in multi patents and it is common to find persons with the same name in China, we identify the unique inventor by checking if two inventors with the same name own the same affiliation. This leads to observation error, as inventors may change their affiliations, or an affiliation may have two inventors with the same name. Similar errors can also be found in [27]. However, this error is extremely small, as there is only a small ratio of inventors with affiliation mobility or the same name. Additionally, part of firms, e.g., Intel, Intel and IBM are western firms with almost no inventors owning identical names. Thus the empirical results could not be changed even if the error is removed.

As the networks are not totally connected, we classify the patents into two cohorts: Inventors in the largest connected component, where any pairs of inventors could reach each other by several intermediates, and inventors not in the largest component, where not all pairs of inventors could reach each other.

Network stability is to a large extent determined by the stability of inventors. In the context of innovation, a high ratio of inventor turnovers from the R&D cooperation network in a short period will lead to an unstable R&D cooperation network. Therefore, we measure the network stability with the inventing life that inventors embarking at innovation (InventLife). The InventLife is measured by the length of period that the inventor first appeared and last appeared in the firm's patents. In detail, the InventLife is measured as follows: As the network to be studied is established with 2003-2005 patents, an inventor is viewed to stay n years in the network if he/she has been absent from and never appeared in the R&D cooperation network since 2004+n+1, e.g., inventor i left the network in 2009 and stayed 4 years in the network since 2004 (2008-2004=4 years). We take the year 2013 and 2014 as the last observation years. Inventors that own patents filed in 2013 or 2014 are viewed to be still in the R&D cooperation network[1], and his inventing life data are set to be censored. Fig. 1 presents the survival curve of the inventing life in the network: Over 60% inventors in the largest component are still in the network one year later, while this ratio is less than 50% for inventors not in the largest component. In other periods, the survival rates of inventors in the largest component are also higher than inventors not in the largest component, which suggests a more stable network relationship in the largest component than in other components.

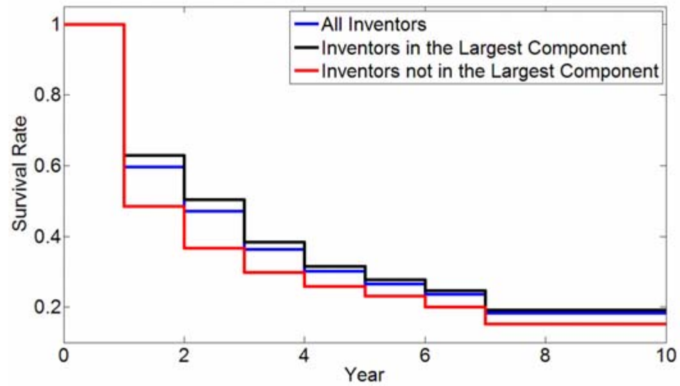

Fig. 1. Survival Curve of the Invent Life in the Inventive Network

B. Variables

We classify the inventors into two cohorts: Inventors in the largest and not in the largest component, so that we could make a comparison and clarify the impact of network structure on network stability in different context. Table 1 presents the summary statistics and correlation matrix of variables of the 9 ICT firms:

Innovation Output: Following most extant studies, e.g.,[9] [28], innovation output is measured by the subsequent patenting in SIPO during 2006-2014. Inventors in the largest component averagely file 10.73 patents, which is twice of the patent output by other inventors (5.01 patents).

InventLife: Table 1 shows that inventors in the largest component averagely stay 3.64 years in the R&D position of the firm (see InventLife), which is longer than that of other inventors (3.07 years).

The network indicators that reflect the inventors' centrality are: Betweenness centrality, i.e., the extent to which an inventor is located 'between' other pairs of inventors; Closeness centrality, i.e., the extent of the closeness to every other inventors; Degree centrality, i.e., the number of inventors that an inventor is directly connected with. The estimation methods of the above three centralities are as follows:

The betweenness centrality of inventor i is defined as

$$Betweenness\_Centrality_i = \sum_{s \neq t \neq i \in I} \frac{\sigma(s,t|i)}{\sigma(s,t)}$$

[1] As not all the patents filed in 2013 and 2014 are in our data, we choose these two years as the last observation years to ensure the accuracy, i.e., inventors absent from the 2013 patents may appear in 2014.

where $\sigma(s,t|i)$ is the total number of shortest paths between s and t that pass through i, and $\sigma(s,t) = \sum_I \sigma(s,t|i)$.

The closeness centrality of inventor i is defined as

$$Closeness\_Centrality_i = \frac{1}{\sum_{j\in U} dist(i,j)}$$

where U is the set of all the inventors excluding inventor i, $dist(i,j)$ is the distance between inventor i and j.

The degree centrality of inventor i is measured by the number of inventors directly connected with inventor i.

ClusterCoefficient: The clustering coefficient of inventor i. Let $\tau_\Delta(i)$ denote the number of triangles, which is a complete subgraph of order three, in I into which inventor i falls, and $\tau_3(i)$ the number of connected triples, which is a subgraph of three vertices connected by two edges, in I for which the two edges are both incident to i. The clustering coefficient of inventor j can be expressed as ([14]):

$$Clustering\_Coefficient_i = \tau_\Delta(j)/\tau_3(j)$$

StructuralHole: The constraint of network connections on inventor, which is measured with the following formula:

$$Structural\_Hole_i = 1 - \sum_{k=1}^{M_i} S_{k,i}^2$$

where M_i is the number of inventors directly connected with inventor i, and

$$S_{k,i} = \begin{cases} \sum_{n=1}^{B_i} \gamma_i\gamma_n & \text{if i has neighbors who are directly connected with } k \\ \gamma_i & \text{if i has no neighbors who are directly connected with } k \end{cases}$$

where n denotes i's neighbor[2] who are directly connected with k, and B_i is the number of i's neighbors who are directly connected with k. γ_i is the inverse of the number of i's neighbors, including k., e.g., i has 4 neighbors, then $\gamma_i = 0.25$, similar explanation applies to γ_n. A higher value of $Structure\_Hole_i$ indicates a low constraint on inventor i, which suggests a greater "freedom" inventor i has to withdraw from existing connections or to exploit structural holes ([16]). This index will have a higher value if inventor owns more structural holes in his/her ego-network.

The characteristics of the firms are controlled by introducing 8 dummy variables that take 1 if the inventor is in the identified firm and 0 otherwise.

In summary, the network structure of the largest component is much different from other components by owning a greater clustering, lower connection constraint and higher centrality. How this would have impact on the patent output and network stability will be further studied.

III. EMPIRICAL RESULTS

We use the two stage regressions: First, we get the instrumental variable by regressing the inventor's post-2003 PatentCount on its pre-2003 value and get the estimated value $\widehat{\text{PatentCount}}$. Similarly, we get the InventLife's estimated value $\widehat{\text{InventLife}}$; Second, we regress the InventLife on $\widehat{\text{PatentCount}}$ and other potential impact factors, and make similar regressions with PatentCount being the dependent variable and $\widehat{\text{InventLife}}$ being the instrumental variable.

As large components usually own well organized structure and may exhibit greater effects, we assign the inventors in larger components with greater weight and apply the weighted negative binomial model to the patent count. We take the edge counts in the component as the weight of its inventors.

The positive impact of inventor stability on patent count can be found in Model 1 – Model 8 in Table 2. This proves that the inventor's network stability positively interacts with his/her innovation performance. However, the significant parameter estimates in both the largest and other components suggest that the positive interaction is not attenuated by the disconnected network.

TABLE 1. SUMMARY STATISTICS AND CORRELATION MATRIX

Inventors in the Largest Component (N=17,808)										
Variable	**Mean**	**Std. Dev.**	**Min**	**Max**	**1**	**2**	**3**	**4**	**5**	**6**
1. PatentCount	10.73	15.39	1	165						
2. InventLife	3.64	2.90	1	10	0.45					
3. ClusterCoefficient	0.63	03786	0	1	0.26	0.15				
4. StructuralHole	0.49	0.27	-0.0069	0.95	0.36	0.15	0.01			
5. BetweenCentrality	0.0014	0.0041	0	0.09	0.45	0.14	-0.31	0.41		
6. CloseCentrality	0.11	0.02	0.06	0.17	0.37	0.10	-0.13	0.53	0.42	
7. DegreeCentrality	4.01	2.65	1	16	0.16	0.02	0.25	0.68	0.23	0.36
Inventors not in the Largest Component (N=4,616)										
Variable	**Mean**	**Std. Dev.**	**Min**	**Max**	**1**	**2**	**3**	**4**	**5**	**6**
1. PatentCount	5.01	8.48	1	113						
2. InventLife	3.07	2.84	1	10	0.47					
3. ClusterCoefficient	0.54	0.48	0	1	0.00	-0.04				
4. StructuralHole	0.12	0.23	-0.13	0.73	0.22	0.07	0.19			
5. BetweenCentrality	2.72e-7	1.70e-6	0	2.19e-5	0.20	0.05	-0.03	0.33		
6. CloseCentrality	0.0012	0.0007	0.0006	0.0036	0.17	0.03	0.53	0.85	0.33	
7. DegreeCentrality	2.33	1.71	1	9	0.11	-0.01	0.61	0.74	0.14	0.89

[2] Here the "i's neighbor" denotes the vertices with a direct connection with i.

TABLE 2. IMPACT OF NETWORK STABILITY ON PATENT OUTPUT WITH NEGATIVE BINOMIAL MODEL

Sample	Inventors in the Largest Component				Inventors not in the Largest Component			
Model	Negative Binomial Model				Weighted Negative Binomial Model			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
$\widehat{\text{LengthofPeriod}}$	0.21***	0.21***	0.21***	0.21***	0.23***	0.23***	0.24***	0.23***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
ClusterCoefficient	0.21***	0.20***	0.19***	0.21***	0.00	0.01	-0.05	0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.09)	(0.09)
ClusterCoefficient ^{2}	-0.48**	-0.47**	-0.47**	-0.48**	0.01	0.01	0.01	0.01
	(0.22)	(0.21)	(0.22)	(0.21)	(0.00)	(0.00)	(0.00)	(0.00)
StructuralHole	1.24***	1.13***	1.06***	1.29***	1.02***	0.93***	0.92***	1.20***
	(0.07)	(0.07)	(0.07)	(0.09)	(0.12)	(0.13)	(0.26)	(0.20)
BetweenCentrality		20.36***				3.15*		
		(4.14)				(16.94)		
CloseCentrality			37.31***				6.49	
			(4.26)				(4.93)	
DegreeCentrality				0.02				-0.00
				(0.01)				(0.03)
Firm Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.41***	3.17***	-4.49***	3.46***	1.50***	1.41***	1.37***	1.70***
	(0.10)	(0.11)	(0.91)	(0.12)	(0.12)	(0.13)	(0.31)	(0.22)
Log Likelihood	-53,336	-53,164	-52,840	-53,335	-15,755	-15,705	-15,753	-15,755
LR Chi2	14,838	15,181	15,830	14,839	4,524	4,624	4,529	4,524
Prob>Chi2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
No. Obs.	17,808	17,808	17,808	17,808	4,616	4,616	4,616	4,616

Dependent Variable: PatentCount.

The clustering demonstrates a weakly inverted U relationship with inventor stability in the largest component, i.e., the estimates of ClusterCoefficient and its square term are only significant at 10% level. It exhibits similar inverted U effect on patent output in Table 2 in the largest component. This suggests that the coexistence of knowledge diffuse that improves innovation and common or even negative information that hampers creativity functions on both innovation performance and network stability. The structural hole also demonstrates positive impact on inventor stability and patent output in the largest component.

We may find from Table 2 that most network indicators demonstrate significant effect on the inventor stability and innovation performance in the largest component, while not significant or the significance level is reduced in the other mutually disconnected components, e.g., clustering coefficient, degree centrality, closeness centrality, degree centrality, which suggests that the disconnected component attenuated the network effect. However, the structural hole is an exception, which exhibits positive effect in both largest and other components.

IV. CONCLUSIONS

Our understanding of the impact of network connectivity remains incomplete. This research makes several theoretical and empirical contributions to our understanding of the moderate role of the network connectivity. Using the patent co-inventing data of top 9 ICT firms that filed the largest number of patents in China, this study establishes the co-inventing network and examines the moderate role of the network connectivity in the reciprocal effect between network stability and innovation output, as well as in the network effects, e.g., clustering, structural hole richness, centrality. The connectivity exhibits positive effect on both patent output and network stability. We further confirm that the clustering and centrality demonstrate significant effect in only the largest connected component, while not significant in other isolated components. This proves the key moderate role of network connectivity, which forms the basis for information transmission and knowledge spillovers. However, the effect structural holes richness demonstrate strong effects, which is not attenuated by network isolation.

Our study has important policy implications: As the largest component plays a major role in the innovation production process in the whole network, it is necessary to maximize the network connectivity. However, Figure 2 shows a declining trend of the size of the largest component, which should be noted by the firm managers as this will hamper knowledge spillovers and may be harmful to innovation; The positive interaction between network stability and innovation output suggests that a stable network structure is beneficial. How to refrain the employees, particularly the high-performers, from flowing out may always be one of the main focuses of firm managers; Additionally, firm managers should enhance the efficiency of the network by reducing redundant links and communications, which may lead to a network structure filled with more structural holes.

ACKNOWLEDGMENTS

This work was supported by National Natural Science Foundation of China, NSFC [71402175][71202116], National Social Science Foundation [11&ZD143], Beijing Natural Science Foundation [9144036], Ministry of Education, Humanities and Social Sciences Youth Fund Project

[13YJC630219].

REFERENCES

[1] Ahuja G., Soda G., and Zaheer A., "The genesis and dynamics of organizational networks," *Organ. Sci.*, vol. 23, pp. 434–448, 2012.

[2] Ahuja, G., "Collaboration networks, structural holes, and innovation: A longitudinal study," *Admin. Sci. Quart*., vol. 45, pp. 425–455, 2000.

[3] Ahuja, G., Soda, G., and Zaheer, A, "The genesis and dynamics of organizational networks," *Organ. Sci.*, vol. 23, pp. 434–448, 2012.

[4] Aime, F., Johnson, S., Ridge, J. W., and Hill, A. D., "The routine may be stable but the advantage is not: competitive implications of key individual mobility," *Strateg. Manag. J*, vol. 31, pp. 75–87, 2010.

[5] Casper, S., "How do technology clusters emerge and become sustainable? Social network formation and inter–firm mobility within the San Diego biotechnology cluster," *Res. Policy*, vol. 36, pp. 438–455, 2007.

[6] Cattani, G., and S. Ferriani., "A core/periphery perspective on individual creative performance: Social networks and cinematic achievements in the Hollywood film industry," *Organ. Sci.*, vol. 19, pp. 824–844, 2008.

[7] Chen, Z. F., and Guan, J. C., "The impact of small world on innovation: An empirical study of 16 countries," *J. of Informetrics*, vol. 4, pp. 97–106, 2010.

[8] Fleming, L. C., and Marx, M, "Managing creativity in small worlds," *California Manag. Rev*., vol. 48, pp. 6–27, 2006.

[9] Fleming, L. C., King, III., and Juda, A. I. "Small worlds and regional innovation," *Organ. Sci.*, vol. 18, pp. 938–954, 2007.

[10] Gardner, T. M, "Interfirm competition for human resources: Evidence from the software industry," *Acad Manag. J.* vol. 48, pp. 237–256, 2005.

[11] Glebbeek, A. C., and Bax, E. H., "Is high employee turnover really harmful? An empirical test using company records," *Acad. Manag. J.*, vol. 47, pp. 277–286, 2004.

[12] Griliches, Z., "Patent statistics as economic indicators: A survey," J. Econ. Lit., vol. 28, pp. 1661–1707, 1990.

[13] Harris, D, and Helfat, C. E., "Specificity of CEO human capital and compensation," *Strateg. Manag. J.*, vol. 18, pp. 895–920, 2002.

[14] Kolaczyk, E. D., "Statistical analysis of network data: Methods and models," New York, NY: Springer, 2009.

[15] Nerkar, A. A., and S. Paruchuri., "Evolution of R&D capabilities: The role of knowledge networks within a firm," *Manag. Sci.*, vol. 51, pp. 771–785, 2005.

[16] Nooy W., Mrvar A., and Batagelj V., "Exploratory social network analysis with Pajek," Cambridge University Press, 2011.

[17] Paruchuri, S., "Intraorganizational networks, interorganizational networks, and the impact of central inventors: A longitudinal study of pharmaceutical firms," *Organ. Sci.*, vol. 21, pp. 63–80, 2010.

[18] Phillips, D. J., "A genealogical approach to organizational life chances: the parent–progeny transfer among Silicon Valley law firms, 1946–1996," *Admin. Sci. Quart.*, vol. 47, pp. 474–506, 2002.

[19] Rao, H, and Drazin R, "Overcoming resource constraints on product innovation by recruiting talent from rivals: a study of the mutual fund industry," *Acad Manag. J*., vol. 45, pp. 491–507, 2002.

[20] Schilling, M. A., "A 'small–world' network model of cognitive insight," *Creat. Res. J.*, vol. 17, pp. 131–154, 2005.

[21] Schilling, M. A., and Phelps, C. C., "Interfirm collaboration networks: The impact of large–scale network structure on firm innovation," *Manag. Sci.*, vol. 53, pp. 1113–1126, 2007.

[22] Shaw, J. D., Duffy, M. K., Johnson, J. L., and Lockhart, D. E., "Turnover, social capital losses, and performance," *Acad. Manag. J.*, vol. 48, pp. 594–606, 2005.

[23] Shaw, J. D., Gupta, N., and Delery, J. E., "Voluntary turnover and organizational performance," *Acad. Manag. J*, vol. 48, pp. 50–68, 2005.

[24] Song, J, Almeida, P, and Wu, G., "Learning–by–hiring: when is mobility more likely to facilitate interfirm knowledge transfer?", *Manag. Sci*., vol. 49, pp. 351–365, 2003.

[25] Sturman, M. C., Walsh, K., and Cheramie, R. A. 2008. The value of human capital specificity versus transferability. *J. Manag.*, vol. 34: pp. 290–316, 2008.

[26] Wezel, F. C., Cattani, G., and Pennings J.M., "Competitive implications of interfirm mobility," *Organ. Sci*., vol. 17, pp. 691–709, 2006.

[27] Zhang, G. P., Guan, J. C., amd Liu, X. L., "The impact of small world on patent productivity in China," *Scientometrics*, vol. 98, pp. 945–960, 2014.

[28] Zhang, G. P., Lv, X. F., and Duan, H. B., "How do prolific inventors impact firm innovation in ICT: Implications from patent co–inventing network," *Tech. Anal. Strateg. Manag.*, vol. 26, pp. 1091–1110, 2014.