Dynamics of Research Team Formation in Complex Networks

Caihong Sun[1,2], Yuzi Wan[1], and Yu Chen[1]

[1] School of Information, Renmin university of China, 100872, Beijing, China
[2] Key Laboratory of Data Engineering and Knowledge Engineering, MOE,
Renmin University of China, Beijing 100872, China
chsun@ruc.edu.cn, wanyuzi0@yahoo.com.cn, yuchen318@gmail.com

Abstract. Most organizations encourage the formation of teams to accomplish complicated tasks, and vice verse, effective teams could bring lots benefits and profits for organizations. Network structure plays an important role in forming teams. In this paper, we specifically study the dynamics of team formation in large research communities in which knowledge of individuals plays an important role on team performance and individual utility. An agent-based model is proposed, in which heterogeneous agents from research communities are described and empirically tested. Each agent has a knowledge endowment and a preference for both income and leisure. Agents provide a variable input ('effort') and their knowledge endowments to production. They could learn from others in their team and those who are not in their team but have private connections in community to adjust their own knowledge endowment. They are allowed to join other teams or work alone when it is welfare maximizing to do so. Various simulation experiments are conducted to examine the impacts of network topology, knowledge diffusion among community network, and team output sharing mechanisms on the dynamics of team formation.

Keywords: Group formation, complex networks, agent-based modeling, knowledge diffusion.

1 Introduction

The concept of team formation or joint action [1,2], is crucial to a wide variety of research topics, including organizational design, computational organization theory, computer-supported collaborative work, game theory, multi-agent systems and artificial social systems. Many researchers focus on studying how to foster the efficient formation of the team to accomplish complex tasks in both real and artificial societies [1]. In the multi-agent systems community, there has been a significant amount of research on team formation and self-organization. Much of the work on team formation focuses on mental states of the agents and their willingness to form teams and collaborate [2, 3]. These studies have driven the development and implementation of frameworks in which teams coordinate closely to develop and execute distributed plans [4, 5]. Recent research on real world networks has revealed

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 2004 – 2015, 2009.

that social systems have a rich structure [6,7]. Gaston et. al. suggested that network structure among individuals can have a significant effect on the dynamics of team formation. Axtell talked about the conditions under which the rational agents form coalitions and how they are engaged in collective action [8]. In [9], Axtell not only built a computer model to simulate the emergence of firms, but studied the empirical firm size distributions produced by the model. Furthermore he compared the results with the empirical data on U.S. firms, such as firm sizes, growth rates, and related aggregate regularities.

The above models on team formations do not discuss knowledge diffusion effects in a research institutes. A member of a research team actually could learn something (knowledge) from other members in-team and out-of-team. However, if knowledge is diffused, models of its diffusion must take explicit account of the structure of connections between members too. Cowan [13,14] proposed a knowledge diffusion model by treating knowledge as a vector of knowledge type, and each agent's knowledge evolved over time through a process of barter exchange. Cowan et. al. also proposed a knowledge improvement in a network industry. They studied the relations between network structure and knowledge diffusion. But in their models, the barter exchange knowledge transmission mechanism could not reflect the knowledge diffusion in cooperative research. In [10], Li and Sun proposed a model different from Cowan knowledge diffusion model; they treated the knowledge gain as the result of cooperation production of knowledge. Cowan et. al. also introduced the concept of knowledge pooling and production to model innovation as knowledge creation [16].

In this paper, we focus on studying the dynamics of team formation in a research community (or organization), since in such a community, knowledge plays important role in team formation. We consider the process of team formation as a self-organizing process in which knowledge diffusion and network structure affect people's strategy of choosing teams. We build a multi-agent model in which heterogeneous agents are self-interested when joining a team. Agents have knowledge endowment and have preference on both income and leisure. They provide a variable input ('effort') and their whole knowledge endowment to team production. There are increasing returns to cooperation, and self-interested agents are self-organized into productive teams. Each agent periodically adjusts its effort level to maximize its welfare non-cooperatively; they could learn from others inside their team and those who are not in their team but has private connections in community. As a result of learning, an agent's knowledge endowment is increased via cooperation. Agents are allowed to join other teams or work alone when it is welfare maximizing to do so. Our main contribution is to introduce knowledge diffusion and sharing mechanisms into the process of team formation.

There are five main assumptions in our model:

1. Agents are self-interested, and bounded rational, they interact via their network connections in the community.
2. The network connections in the community are static.
3. An agent can not belong to two or more teams at the same time.
4. Research team is a cooperative coalition to produce output; their production environment is characterized by increasing returns and various rules for dividing team output.

5. Knowledge endowment of each agent is increased by learning knowledge from other agents inside the team and outside the team.

In our model, both knowledge diffusion and interactions among agents are related to the network structure of communities. Our main focuses are to observe:

- How network structures affect team formations?
- How team output allocation mechanisms impact team formation and the whole knowledge level of the community.

In the next section (section 2), we present an agent based computational model of team formation in details. Then, in section 3, the experimental designs are described and their results are analyzed. Section 4 summarizes the main findings and draws conclusions.

2 Model of Research Team Formation

As we discussed in above section, we treat a research team as a kind of firm to produce some output not to fulfill a complicated task. We treat each member in a research community as an agent, and there is a social network among members in the community.

Suppose there is a finite, fixed set of agents, A, each of whom has an initial knowledge endowment $k_{i\in A}$ and works with some effort level $e_{i\in A} \in [0,1]$. Consider a representative team composed of N agents. The team knowledge level and team effort level are simply defined by Equation (1) and (2).

$$K = \sum_{i=1}^{N} k_i \tag{1}$$

$$E = \sum_{i=1}^{N} e_i \tag{2}$$

2.1 Team Production Functions

The team produces output, O, as a function of K and E,

$$O(K,E) = \rho K E^2 . \tag{3}$$

Equation (3) represents the team's production function. ρ ($\rho > 0$) is an adjusting coefficient. It is a cobb douglas production function with increasing returns to cooperation.

2.2 Mechanisms of Output Sharing

How the agents in a team share total output? Here, we proposed four basic mechanisms to share output.

Sharing mechanism 1: All agents share output equally. For agent i, its share of team output is

$$O_i = O(K, E) / N . \quad (4)$$

Where K, E is team knowledge level, team effort level of the team that agent i belongs to respectively, and N is the number of agents in the team.

Sharing mechanism 2: Each agent's sharing output is proportional to its knowledge endowment. For agent i, its share of team output is

$$O_i = \frac{O(K, E) * k_i}{\sum_{j=1}^{N} k_j} \quad (5)$$

Sharing mechanism 3: The sharing output of an agent is proportional to its effort. For agent i, its share of team output is

$$O_i = \frac{O(K, E) * e_i}{\sum_{j=1}^{N} e_j} \quad (6)$$

Sharing mechanism 4: Consider the combinations of mechanism 1, 2 and 3.

2.3 Community Network

Generally, there are connections among people in a community. These connections could represent the relationship such as friendship, acquaintance or kinship etc. In our model, we consider agents in the research community have such private connections between each other. We treat each agent as a node, if two nodes have a private connection (friendship), we add a link between them, then we get a community network among agents. As discussed in Section 1, network structure of community has impacts on team formation and knowledge diffusion among agents. In the past, the network structure is assumed to be either completely regular or completely random. But many empirical studies [12, 17] indicate that many social networks exhibit small world characteristics which are somewhere between these two extremes. In this paper, we consider a WS model proposed by Watts and Strogatz [12] to simulate different community network structures. WS model describes the process from regular network to random network. The steps of WS model are as follows [12]:

(1) Starting from a ring lattice with n vertices and k edge per vertex (for a vertex, it connects to its k/2 nearest neighbor in both left and right sides). To obtain a sparse but connected network, let $n > k > \ln n > 1$.

(2) Rewiring each edge at random with probability p. When p=0, the network is a regular network, when p=1, it is a random network, when p is between 0 and 1, the network exhibits small world characteristics.

2.4 Knowledge Diffusion Model

When people work in a team, knowledge diffusion happens. When people interact, knowledge transmission may occur. Based on this assumption, we bring forward a knowledge diffusion model adapted from a model proposed by the authors in [10].

Here, we consider knowledge improvement incurred by knowledge diffusion as a result of cooperation production of knowledge, and use cobb-douglas production function to model the process of knowledge improvement. In a community, the knowledge improvement of an individual depends on its own knowledge and the knowledge difference from others. In our model, knowledge is treated as a simple scalar for simplicity, and people can only learn knowledge from the ones who have higher knowledge level. Suppose at time t, agent i is the one who imparts the knowledge, agent j is the receiver, when agent i and agent j interact; the knowledge increment level is defined as

$$\Delta k_{j,i,t+1} = \begin{cases} 0, & k_{i,t} \le k_{j,t} \\ \eta k_{j,t}{}^{\alpha} (k_{i,t} - k_{j,t})^{\beta}, & k_{i,t} > k_{j,t} \end{cases}. \tag{7}$$

Where $0 < \eta, \alpha, \beta < 1$, η is a learning rate, $\Delta k_{j,i,t+1}$ denotes in time $t+1$, the amount of knowledge agent j learn from agent i, $k_{i,t}$ denotes the knowledge level of agent i in time t and $\Delta k_{j,t+1}$ denotes the total learning from all others and it is defined as

$$\Delta k_{j,t+1} = \max_{i, who\ has\ \mathrm{int}eraction\ with\ j} (\Delta k_{j,i,t+1}). \tag{8}$$

Hence,

$$k_{j,t+1} = k_{j,t} + \Delta k_{j,t+1} \tag{9}$$

But when an agent interacts with others? In our model, we suppose all agents in the same team interact with each other. An agent could interact with those agents outside its team, only when there are private connections between them.

To measure the knowledge performance of a community, we define two measurements, one is knowledge level and the other is knowledge dispersion.

The average knowledge level of the community in time t is defined as

$$\mu_t = \frac{1}{|A|} \sum_{i \in A} k_{i,t} \tag{10}$$

The standard error which reflects the degree of even distribution is defined as

$$\sigma_t = \sqrt{\sum_{i \in A} k_{i,t}{}^2 / |A| - \mu_t{}^2}. \tag{11}$$

2.5 Utility Function of Agents

In this part, we adopt the concept of utility in Axtell's papers in studying firm emergence. Each agent has Cobb-Douglas preferences for income and leisure [8,9]. He supposed that all time not spent working is spent in leisure, thus agent i's utility can be written as a function of its effort level, e_i, as

$$U^i(e_i;\theta_i, E_{\sim i}, K_i, N) = (O_i(e_i; E_{-i}, K_i))^{\theta_i}(1-e_i)^{1-\theta_i} \tag{12}$$

Where K_i is the knowledge level of the team which agent i belongs to, $E_{\sim i}$ is the amount of effort put in by the other agents in the team, θ_i is agent i's preference on income and leisure. The utility function of agents is a function of knowledge level, effort level, and its preference. Individual effort is not observable, each agent, i, selects the effort level, e_i*, that maximizes its utility; formally,

$$e_i* = \arg\max_{e_i}[U^i(e_i;\theta_i, K_i, E_{\sim i})]. \tag{13}$$

2.6 Team Formation Dynamics

Initially, we choose M agents who have the first M largest knowledge endowment as M team leaders, then let others randomly join the M teams. Each agent has its initial knowledge endowment and effort endowment. What make an agent decide to stay in its team or join another team? In our model, an agent will try to get its optimal effort level according to Equation (13), and then it uses this new effort as input effort level to compute its sharing output and its utility. At the same time, it seeks to know the output sharing of its neighbors (through the community network). If one of its neighbors has higher utility, it would like to join its neighbor's team; otherwise it stays in its team. Furthermore, if an agent can gain a bigger utility by working alone, it works alone as a singleton team. In addition, knowledge diffusion happens in the dynamic process of team formation according to Equation (7), (8) and (9).

3 Numerical Simulations and Analysis

The team formation model just developed a complex dynamic process, which we study numerically. We examine how sharing mechanisms, initial settings and network structures impact the results of team formation and knowledge distribution in community level. We use multi-agent simulation tool Repast (Recursive Porous Agent Simulation Toolkit) [11] to implement our model.

3.1 Set-Up of Numerical Simulations

Suppose a research community consists of 100 agents, and agents are grouped into 20 teams as described in Section 2.6. Each agent has its initial knowledge endowment and effort endowment. We define four initial settings to examine how agents' initial knowledge endowment and effort endowment affect the knowledge distribution in community level.

Initial setting 1: Individual knowledge endowment k_i are randomly drawn from a uniform distribution over [0, 1], and individual effort level e_i are randomly drawn from a uniform distribution over [0, 1] too.

Initial setting 2: Individual knowledge endowment k_i are randomly drawn from a uniform distribution over [0, 1], and individual effort level e_i are set to a homogenous value 0.5.

Initial setting 3: Individual knowledge endowment k_i are randomly drawn from a Gaussian distribution over [0.5, 0.3], and individual effort level e_i are randomly drawn from a uniform distribution over [0, 1] too.

Initial setting 4: Individual knowledge endowment k_i are randomly drawn from a Gaussian distribution over [0.5, 0.3], and individual effort level e_i are set to a homogenous value 0.5.

The model set-up is almost identical to section 2. Preferences are homogeneous in the agent population, $\theta = 0.5$. The community network is generated by a WS model with every agent having 6 connections with a probability p (p=0.001, 0.1 and 0.9).

3.2 Experimental Design

3.2.1 The Impacts of Sharing Mechanisms and Initial Settings

In this experiment, we let the rewiring parameter p=0.1 in WS-model simulate the community network with small world characteristics. We assume that the learning rate in-team and out-of-team are different, and we set the learning rates are 0.1 and 0.05 for in-team and out-of-team respectively. We apply four sharing mechanisms (denote as SM1, SM2, SM3 and SM4) proposed in Section 2.2 to examine the performance in community level. Note, here SM4=SM2*0.5+SM3*0.5. In addition, we simulate four different initial settings of knowledge endowment and effort level corresponding to 1, 2, 3 and 4.

Fig. 1 shows the performance of community. When team evolution is stable (i.e., no agents want to change team), the performance of community could be measured by knowledge level, knowledge dispersion, and effort level of the community. In Fig 1, in x axis, 1, 2, 3 and 4 stand for initial settings 1, 2, 3 and 4 respectively.

As we can see from the top left panel of Fig. 1, both initial knowledge endowments and initial effort endowments have impacts on the knowledge level of community (as in Equation (10)), no matter which sharing mechanism is used. From the top right panel of Fig. 1, initial settings also have effects on knowledge dispersion. If knowledge endowments follow the same distribution, setting all effort level to the same value could lead to higher knowledge level and more even knowledge distribution.

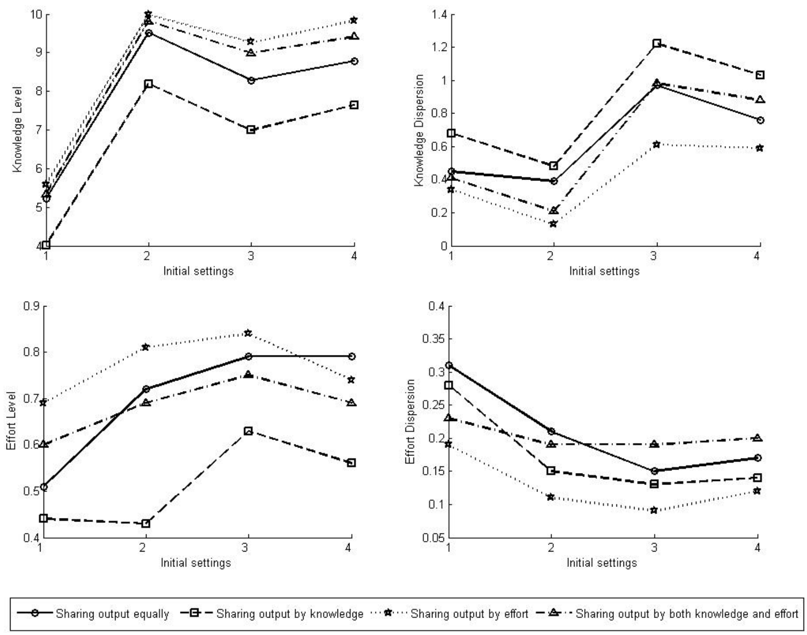

Fig. 1. Effects of initial settings and sharing mechanisms on performance in community level

From the left bottom panel and right bottom panel of Fig.1, we can see that the knowledge endowment with Gaussian distribution performs well, i.e., the average effort level of community is high and effort dispersion of community is low. It might imply that agents have more even knowledge level initially, and putting more effort would be the main source to increase their utilities.

Now we proceed to examine the effects of sharing mechanism. From the left top panel and right top panel of Fig.1, we can see that sharing output by effort performs well, i.e., the community with high average knowledge level, and small knowledge dispersion. This is probably caused by our initial knowledge settings. Initially the difference of knowledge among agent is small; there are no elites in the community who possess a large amount of knowledge. We would like to do more experiments to examine this aspect in future. From the bottom panels of Fig.1, we can see that sharing output by effort performs well, and sharing output by knowledge performs badly. If agents share output by knowledge, they do not have more incentives to work hard. In this context, sharing output by knowledge performs worse than sharing output equally does. This situation probably is caused by our utility function in which knowledge improvement has no or little impacts. We are interested in seeing what will happen about the mechanism if the utility function is changed.

Table 1 shows results of different sharing mechanisms and initial settings in team levels.

Table 1. Impacts of sharing mechanisms on team formation

Initial Settings		SM1	SM2	SM3	SM4
1	Time steps of convergence	186	169	179	175
	No. of teams	14	17	16	15
	Max. size of teams	17	19	16	19
	Avg. size of teams	7	6	6	7
	Min. size of teams	1	2	1	2
2	Time steps of Convergence	164	132	171	147
	No. of teams	15	19	16	18
	Max. size of teams	13	17	15	14
	Avg. size of teams	7	5	6	6
	Min. size of teams	1	1	1	1
3	Time steps of Convergence	221	236	223	221
	No. of teams	10	13	11	12
	Max. size of teams	19	24	21	27
	Avg. size of teams	10	8	8	8
	Min. size of teams	2	3	2	3
4	Time steps of Convergence	204	182	201	189
	No. of teams	12	16	12	15
	Max. size of teams	18	21	20	18
	Avg. size of teams	8	6	8	7
	Min. size of teams	1	2	1	2

Note: The data describe the features of teams when team evolution process converges to a stable state.

From Table 1, we can see, four sharing mechanisms have little difference on team level. Initial settings with uniform knowledge endowment converge a little faster. Further examination inside teams is expected in our future work.

To sum up, initial settings and sharing mechanisms have impacts on knowledge and effort distribution in community, and have little impacts on the number of teams, and the size of teams, and converging time. Our simulation results show that sharing output by effort is a better choice under our current parameter settings.

3.2.2 The Impacts of Network Structure on Team Formation

To study the effects of network structure on team formation, we set the rewiring parameter p discussed in section 2.3 as 0.001 to simulate regular network, 0.1 to simulate small world network, and 0.9 to simulate the nearly random network. Table 2 gives out features of team formation. As shown in Table 2, if the community network is small world, both the average knowledge level and average effort level are high in community. The result conforms to other research work which indicated that average knowledge level tends to high in small world network [10, 13, 15] although we change the knowledge diffusion model and introduce the team formation process. Regular network and random network have little difference between each other in average knowledge level and average effort level, but in the regular network, team formation process converges very fast.

Table 2. Effects of network structures on team formation

Network	K. Mean	K. Stderr	E.Mean	E.Stderr	Tim steps
Regular	3.87	0.17	0.44	0.13	65
Small World	9.41	0.88	0.69	0.20	189
Random	3.81	0.21	0.47	0.15	246

*Here, we adopt initial setting 4 and sharing mechanism 4. Learning rate inside the team is 0.1, and learning rate outside the team is 0.05.

3.2.3 The Effects of Learning In-team and Out-of-Team

To study the impacts of learning effects in-team and out-of-team, we consider different learning rates η_1, η_2 which denote learning rate inside and outside teams respectively). The experimental results are shown in Table 3.

Table 3. Impacts of knowledge diffusion on team formation

Learning rate	K. Mean	K. Stderr	E.Mean	E.Stderr	Tim steps
$\eta_1 = 0.1, \eta_2 = 0.05$	9.41	0.88	0.69	0.20	189
$\eta_1 = 0, \eta_2 = 0.1$	9.45	0.83	0.64	0.19	215
$\eta_1 = 0.1, \eta_2 = 0$	9.38	0.79	0.58	0.17	183

*Here, we adopt initial setting 4 and sharing mechanism 4. And p=0.1(small world)

From Table 3, we can see, if learning exists, learning from inside team and outside team does not matter much for average knowledge level. We boldly guess that this result is caused by the knowledge learning rule in Equation (8), in which agents only learn a maximal knowledge increment which decreases the effects of learning both inside team and outside team. Further experiments are needed to be done to support the guess. But agents learning from both inside the team and outside the team could encourage them to put more efforts which is shown in Table 3. To sum up, learning from network has little effects on knowledge level, but encourages agents to put more efforts in team production.

4 Conclusions and Future Work

In this paper, we build an agent-based model to study the dynamics of team formation in a research community (or organization), since in such a community, knowledge plays important role in team formation. We treat the process of team formation as a self-organizing process in which knowledge diffusion and network structure affect people's strategy of joining teams. In our model, heterogeneous agents are self-interested and have knowledge endowment and have preference on both income and leisure. Agents are self-organized into productive teams by periodically adjusting their effort levels to maximize their welfare non-cooperatively; they could learn from others inside their team and those who are not in their team but have private connections in community.

As demonstrated in our experimental results, we find that network structure, initial settings and sharing mechanisms affect the average knowledge level and effort level of the community a lot. But initial settings and sharing mechanisms have little impacts on the number of teams, and the size of teams, and converging time. Our simulation results show that sharing output by effort is a better choice under our current parameter settings. Average knowledge level of community tends to high in small world network although we change the knowledge diffusion model and introduce the team formation process. In addition, although learning from network has little effects on knowledge level in our model, but it encourages agents to put more efforts in team production.

The model and the simulation we have done in this paper is only the very first step of our research, our future work includes: (1) doing more experiments to further examine the initial settings and parameters which include enlarging network size, comparing different initial team generation methods, setting initial knowledge endowments, applying different knowledge diffusion models etc.; and (2) examining the knowledge distribution inside teams; (3) considering the co-evolution of network and knowledge; (4) extending the model by treating agents' knowledge endowment as a vector; (5) introducing different types of utility functions.

Acknowledgments. This paper is supported by Natural Science Foundation Project (No.70671102).

References

1. Gaston, M., desJardins, M.: Agent-organized Networks for Dynamic Team Formation. In: AAMAS 2005, pp. 230–237 (2005)
2. Cohen, R., Levesque, H.J., Smith, I.: On Team Formation. In: Contemporary Action Theory, pp. 87–114 (1997)
3. Wooldridge, M., Jennings, N.: The Cooperative Problem Solving Process. Journal of Logic and computation 9(4), 563–592 (1999)
4. Tambe, M.: Towards Flexible Teamwork. Journal of artificial intelligence Research 7, 83–124 (1997)
5. Durfee, D.H., Lesser, V.R.: Partial Global Planning: A Coordination Framework for Distributed Hypothesis Formation. IEEE Transactions on Systems, Man, and Cybernetics 21(5), 1167–1183 (1991)
6. Newman, M.E.: The Structure and Function of Complex Networks. SIAM Review 45, 167–256 (2003)
7. Strogatz, S.: Exploring Complex Networks. Nature 410, 268–276 (2002)
8. Axtell, R.: Non-Cooperative Dynamics of Multi-Agent System. In: Proceedings of the international conference on autonomous agents and multi-agent systems, Bologna, Italy (2002)
9. Axtell, R.: The Emergence of Firms in a Population of Agents: Local Increasing Returns. Unstable Nash Equilibria, and Power Law Size Distributions. Working paper (1999)
10. Li, J., Sun, D.: Knowledge Propagation Model in Complex Networks. Journal of South China University of Technology 34(6) (2006) (in Chinese)
11. (Last visited October 15, 2008), `http://repast.sourceforge.net/index.html`

12. Watts, D., Strogatz, S.: Collective Dynamics of “Small-World” Networks. Nature 393, 440–442 (1998)
13. Cowan, R., Jonard, N.: Network Structure and Diffusion of Knowledge. Journal of economic dynamics and control 28, 1557–1575 (2004)
14. Cowan, R., Jonard, N., Zimmermann, J.B.: Evolving Networks of Inventors. Journal of Evolutionary Economy (16), 155–174 (2006)
15. Gaston, M., desJardin, M.: The Effect of Network Structure on Dynamic Team Formation in Multi-Agent Systems. Computational Intelligence 24(2), 122–157 (2008)
16. Cowan, R., Jonard, N., Zimmermann, J.B.: Bilateral Collaboration and the Emergence of Innovation Networks. Management Science 53(7), 1051–1067 (2007)
17. Collins, J., Chow, C.: It’s a small world. Nature 393, 409 (1998)