

Research Article

Open Access

Chenguang Li* and Yongan Zhang

Percolation on the institute-enterprise R&D collaboration networks

Abstract: Realistic network-like systems are usually composed of multiple networks with interacting relations such as school-enterprise research and development (R&D) collaboration networks. Here, we study the percolation properties of a special class of R&D collaboration network, namely institute-enterprise R&D collaboration networks (IERDCNs). We introduce two actual IERDCNs to show their structural properties, and we present a mathematical framework based on generating functions for analyzing an interacting network with any connection probability. Then, we illustrate the percolation threshold and structural parameter arithmetic in the sub-critical and super-critical regimes. We compare the predictions of our mathematical framework and arithmetic to data for two real R&D collaboration networks and a number of simulations. We find that our predictions are in remarkable agreement with the data. We show applications of the framework to electronics R&D collaboration networks.

Keywords: networks; percolation; generating function

PACS: 64.60.aq, 64.60.ah, 64.60.-i, 89.75.Fb

DOI 10.1515/phys-2015-0008

Received November 21, 2013; accepted August 23, 2014

1 Introduction

In the past decade, complex networks have been studied intensively and widely applied in many real natural, physical and social systems. The structure and function of a single network component have already achieved great development due to numerous modeling and analysis works [1–7]. However, in fact, as one component in a larger complex of multiple systems, a single network does not live in isolation

because it always interacts with and interdepends on other networks [8]. Therefore, much attention has recently been focused on the topic of multiple networks with complex interplay and distinct topology.

Some studies on multiple networks, including interacting networks and interdependent networks, are starting to demonstrate excellent value. It is worth mentioning that several attractive models, which can be traced back to Buldyrev [9], focus on the properties of interdependent networks based on coupling between systems. The purpose of these write-ups is to elucidate distinct network nodes which depend on each other and determine the robustness of networks in general [10–13]. In providing proper functionality, mutually coupled and trigger processes have emphasized that when a failure has occurred in nodes from one network, it causes nodes in another network to fail. Furthermore, the failure of some initial nodes may trigger cascading failures from one network to another through a communication channel between a pair of nodes that can even destroy both networks [14, 15]. Beyond that, mathematical frameworks on interacting networks are another ingenious objective of the study of multiple networks. For instance, email, electronic commerce, electric grid, communications and socio-technical systems have been characterized by networks of networks, and the overall connectivity in these systems could be enhanced by calculating the properties of the components [8, 16]. Leicht and D'Souza [8] developed a framework based on generating functions for analyzing undirected interacting networks given the node connectivity within and between networks. Moreover, they derived exact expressions for the percolation threshold describing the onset of large-scale networks as well as each network individually. Aside from that, Fu et al. [16] proposed a mathematical framework based on generating functions for analyzing directed interacting networks, derived the necessary and sufficient condition for the absence of the system-wide giant in- and out-component and proposed arithmetic to calculate the corresponding structural measures in the sub-critical and super-critical regimes. Both these efforts extend the application of generating functions on percolation transition in multiple coupled networks.

***Corresponding Author: Chenguang Li:** School of Economics and Management, North China University of Technology, 100144 Beijing, P. R. China, E-mail: lichenguang@ncut.edu.cn

Yongan Zhang: Economics and Management School, Beijing University of Technology, 100124 Beijing, P. R. China, E-mail: bjutzhya@bjut.edu.cn

 © 2015 C. Li, Y. Zhang, licensee De Gruyter Open.

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 License.

It is generally known that regardless of individual enterprise, enterprise groups, or institutes in regional innovation systems, master cutting-edge knowledge and techniques are crucial. The R&D collaboration between institutes and enterprises is a vital form of new knowledge and technique creation, and the key elements in forming R&D collaboration networks is shared knowledge and technique creation. Thus, we pay special attention to percolation in the IERDCNs because it is helpful for further study of the transmission of knowledge and techniques in networks. In particular, percolation can be used to measure the number of enterprises that obtain the knowledge and techniques, while the giant component decides the transfer scope of the knowledge and technology.

This study is mainly focused on the percolation properties of IERDCNs which is a special type of school-enterprise R&D collaboration networks. Furthermore, we define R&D agents as nodes and collaborations as edges. There are two types of nodes which form different networks in IERDCNs, one is technology enterprises (hereinafter referred to as enterprises) and the other is research institutes (hereinafter referred to as institutes, which include colleges and private research institutes). Nodes always show their independence in their respective networks because of the intense competition between enterprises in the same industry. If one partner breaks down, others will still work. So, IERDCNs are interacting networks containing connectivity links only, and it is appropriate to choose the mathematical framework created by Leicht and D'Souza [8] as the base model. Something interesting and distinct from prior studies is that in this study, we tried to take connection probability into consideration. Moreover, the conditions for these components to become the giant ones are worthy of discussion. Hence, further investigations are needed to model the mechanism underlying discontinuous percolation processes. Our work can supplement and enrich existing studies on multiple networks.

The rest of the paper is organized as follows. From degree distribution, density, assortativity, and etc., Section 2 introduces the unique network topology and structural properties of IERDCNs. Section 3 puts forward two mathematical frameworks for IERDCNs, a general one and a special one with connection probability, which are useful for deriving percolation conditions and calculating the average sizes of components. We evaluate our arithmetic using a set of simulation instances and discuss the practical application of IERDCNs in Section 4. Finally, in Section 5, we discuss the possible implications and extensions of our work.

2 Structural properties

As mentioned previously, enterprises and institutes compose the IERDCNs which can be found mostly in technology innovation networks [17]. Resource-based theory emphasizes that there is a heterogeneous resource in enterprises and institutes, in which percolation would be able to create a synergistic effect of knowledge and technique flow in the overall R&D collaboration system. Meanwhile, any enterprise may not satisfy the need for innovation by utilizing its limited internal resources. Instead, it needs R&D collaboration to obtain more resources, knowledge and techniques. Moreover, enterprises in the same industry tend to select institutes as partners because of intense inter-firm competition caused by the homogeneity of the firms' commodity or services. Institutes with abundant intellectual resources have a team of professional researchers and technicians. Through external collaboration, institutes may transmit their accumulated knowledge and generate intellectual property rights for new technology. Certainly, they can obtain economic benefits. For instance, ENEA is an Italian government-sponsored R&D center. If an enterprise has innovative ideas and wants to invent a technique, it will entrust one or two institutes of ENEA with the development of cutting-edge knowledge and techniques, and after all research activities are complete, the enterprise and the institutes will apply for a patent and share ownership of the knowledge or technique. Thus, enterprises can greatly enhance their current workflow, productivity and quality with the new techniques, while institutes can increase their efforts on promoting new technology when collaborating with enterprises.

Some interesting things that we found are that most enterprises tend to select two institutes for R&D collaboration, which may be a necessary safeguard to prevent the failure of a single connection, as well as that there are few enterprise-with-enterprise collaborations due to the inter-firm competition we mentioned earlier. Furthermore, there are also few internal connections between institutes. The cause for this may be that each institute has abundant intellectual resources.

As already stated in our introduction, an IERDCN is comprised of an enterprise R&D network and an institute R&D network. In the enterprise R&D network, different enterprises have their partner selection preferences, but enterprises in the same industry have special uniform characteristics on collaborative R&D. Hence, we take the electronics enterprises in this study and investigate their combinations with the institute R&D network.

Table 1: Characteristics of IERDCNs 1 and 2.

IERDCN	1	2
Density	0.0397	0.0952
Ave degree	2.4194	2.0000
Ave closeness	0.0012	0.0083
Ave distance	3.8631	5.7619
Ave betweenness	63.7742	50.0000
Ave clustering coefficient	0.0000	0.0000

Two cases of IERDCNs in China have been investigated in detail (see Fig. 1 for their typologies and Table 1 for their characteristics). All enterprises and institutes are located in Beijing Zhongguancun Science Park. There are 50 electronics enterprises which compose the enterprise R&D network and 12 institutes which compose the institute R&D network in IERDCN 1 (Fig. 1(a)). Similarly, there are 11 electronics enterprises and 11 institutes in IERDCN 2 (Fig. 1(b)). Some large conglomerates such as Lenovo, Digital China, and BOE are very famous for independently conducting R&D, while some like Founder, Tsinghua Tongfang, Datang, Potevio, and etc. have formed strategic R&D partnerships with other enterprises that they have a long-term business connection with, meaning that there are enterprises that do not collaborate with any institute in the institute R&D network. It should be stressed that the links come from patent application, which is an important result of R&D collaboration. If a couple of partners have applied for patent protection for a new technique, R&D collaboration between them has become inevitable. All patent information can be found on the website of the Chinese State Intellectual Property Office.

IERDCN 2, extracted from IERDCN 1, is a special case because it is not very easy to find an identical network. Actually, it is a classic IERDCN, where there is no intra-network R&D collaboration and each entity has only two partners which are from the other network. This occurs when all enterprises in the same industry have equal scale and market position, as well as there always being unsatisfied demand for future R&D. Furthermore, enterprises prefer to collaborate with two institutes to ensure a higher probability of success and lower costs. Every research institute, in the meantime, has the same powerful technological strength and is ready for collaborative R&D. It is important that equal amounts of collaborations ensure that the whole network works successfully.

Table 2 shows the average degree of intra-network and inter-network connections for IERDCNs 1 and 2. In addition to the illustrations in Fig. 1, we can see that rarely are there intra-network collaborations. In contrast, more

Table 2: Average degree of intra- and inter-network connections for IERDCNs 1 and 2.

IERDCN		1	2
Enterprise R&D network	$\bar{d}_{intra-}$	0.1600	0.0000
	$\bar{d}_{inter-}$	1.4000	2.0000
Institute R&D network	$\bar{d}_{intra-}$	0.1667	0.0000
	$\bar{d}_{inter-}$	5.8333	2.0000

inter-network collaborations exist in IERDCN 1 where the number of collaborations between institutes and enterprises is relatively large. It is obvious in IERDCN 2 that the inter-network collaborations in the enterprise R&D network and institute R&D network are equal.

We have to explain these structural properties based on the management attributes of enterprises and institutes. Firstly, collaboration and competition coexist in each network. In addition to similar or identical products and services, competition comes from information, skill and knowledge barriers. Collaboration relies on complementarities in creating common benefits for alliance partners, and it can help to diffuse advanced knowledge and techniques. However, due to the lack of profit-driven and informational communication, there are only a few collaborations and these occur when enterprises or institutes have the same stockholders. Furthermore, we find that enterprises enjoy collaborating with more than one institute as the more collaboration there is, the better the chances of generating new techniques and knowledge. Moreover, some institutes have large quantities of partners because of the institute's higher R&D capability and reputation. Even so, if an institute has too many partners, this may overburden the institute and even lead to failure. To avoid excessive competition, more regular collaboration has emerged after several years of competition and collaboration.

We must emphasize that despite there being a link between the generation and transmission of knowledge and techniques, it does not mean that knowledge and techniques might be transmitted for certain. The reason for this may be related to the diversity between different organizations and differences in absorptency. Especially in practice, the collaborations between enterprises and institutes or between themselves may fail, or enterprises and institutes may accomplish the R&D tasks of selected partners rather than complete all tasks. A collaboration between an enterprise and an institute may not result in the generation of new knowledge or techniques. We will introduce connection probability to the modeling process for studying this issue. In doing so, we implicitly assume that each collabo-

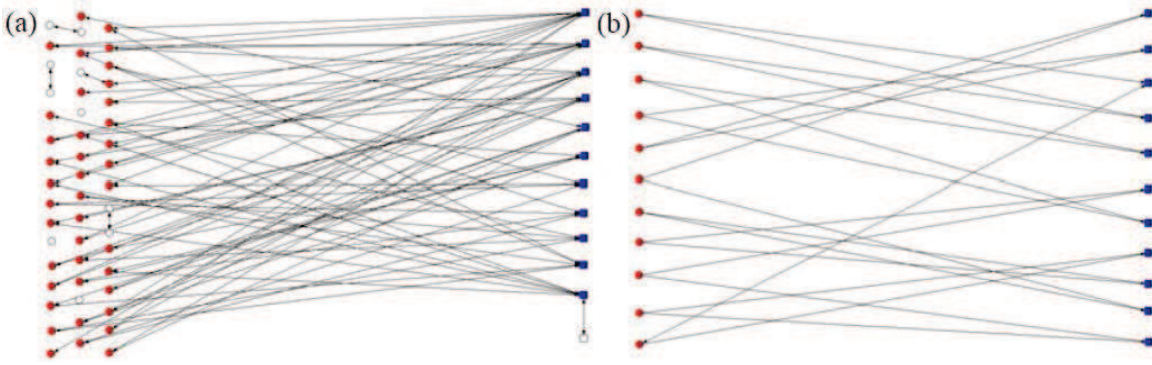


Figure 1: An IERDCN (a) and a special case (b). The red and white circles represent enterprises with and without, respectively, connections to institutes, while the blue and white squares represent institutes with and without, respectively, connections to enterprises.

ration has a risk of failing to generate knowledge and techniques, which is a departure from prior studies that use the assumption that collaborations are 100 percent successful. Nevertheless, studying the R&D collaborations in an eight-year period has allowed us to confirm our assumption and include it in this study based on the fact that it is reasonable in practice. Additionally, the actual probability of intra-network and inter-network connections can be determined by investigation and research. There may be some structural properties which reflect the economic, management or social attributes of networks that should be taken into a mathematical model's construction process. The second arithmetic equation in the next section will present the corresponding work.

3 A mathematical framework based on generating functions

3.1 The generating functions of classic IERDCNs

Here we give a brief description of generating functions, which have been used in many network connectivity studies [8–10, 16, 18–21] that all found that the functions are quite accurate when the structure of networks is approximately tree-like. Now, consider classic IERDCNs formed by two interacting networks, 1 and 2, whose characteristic is, the very rarely observed, nodes in the same network with inner links. Network 1 is composed of enterprises with the similar products and services while network 2 is composed of institutes with identical R&D capabilities. In practice, networks 1 and 2 represents a scenario full of intense competition and no collaboration between enterprises or institutes in their respective networks. Each indi-

vidual network μ can be characterized by a multi-degree distribution, $\{p_{k_1 k_2}^\mu\}$, where $p_{k_1 k_2}^\mu$ is the probability that a network μ 's node has k_1 edges to nodes in network 1 and 2 edges to nodes in network 2. The multi-degree distribution for network μ can be written in the form of a generating function:

$$G_\mu(x_1, x_2) = \sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^\mu x_1^{k_1} x_2^{k_2}. \quad (1)$$

We also assume that the distribution $p_{k_1 k_2}^\mu$ is correctly normalized, so that:

$$G_\mu(1, 1) = 1. \quad (2)$$

Firstly, consider selecting uniformly at random a v - μ edge which is used to connect a couple of nodes in network μ and network v . Relative to a single network, the remaining local connectivity to nodes in other networks is also accounted for by excess degrees [19]. We use $p_{k_1 k_2}^{\mu v}$ to denote the probability that a randomly chosen v - μ edge to a node with excess v degrees has total v -degrees of $k_v + 1$. Then, the generating function for the distribution $\{p_{k_1 k_2}^{\mu v}\}$ is:

$$\begin{aligned} G_{\mu v}(x) &= \sum_{k_1, \dots, k_l=0}^{\infty} p_{k_1 \dots k_l}^{\mu v} x_1^{k_1} \dots x_l^{k_l} \\ &= \sum_{k_1=0, k_2=0}^{\infty} \frac{p_{k_1 (k_v+1) k_2}^\mu (k_v+1)}{\sum_{j_1=0, j_2=0}^{\infty} p_{j_1 (j_v+1) j_2}^\mu (j_v+1)} x_1^{k_1} x_2^{k_2}. \end{aligned} \quad (3)$$

According to the structural properties of networks 1 and 2 in IERDCN 2, we have $p_{k_1 k_2}^{11} = p_{k_1 k_2}^{22} = 0$, which means that $G_{11}(x_1, x_2) = G_{22}(x_1, x_2) = 0$. Thus, we only consider the excess degree of a node in networks 1 and 2, and the generating functions can then be written as:

$$G_{12}(x_1, x_2) = \sum_{k_1=0, k_2=0}^{\infty} \frac{p_{k_1 (k_2+1)}^1 (k_2+1)}{\sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^1 k_2} x_1^{k_1} x_2^{k_2} \quad (4)$$

and

$$G_{21}(x_1, x_2) = \sum_{k_1=0, k_2=0}^{\infty} \frac{p_{(k_1+1)k_2}^2 (k_1+1)}{\sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^2} x_1^{k_1} x_2^{k_2}. \quad (5)$$

Secondly, consider the component sizes of IERDCN 2. All the component sizes are finite in the beginning, but after the emergence of a giant connected component, they become larger and larger. When the component sizes are too large to tolerate and to ignore the closed loop of edges, the generating functions can be adjusted to calculate the average component size [8, 16, 18]. They can also be adjusted to calculate the probability that a randomly chosen node belongs to the giant component in a supercritical regime [8, 16, 18]. We have to scrutinize components of a randomly chosen undirected edge. Let $H_{\nu\mu}(x_1, x_2)$ be the generating function for the distribution of the sizes of components reached by following randomly chosen edges connecting nodes in network ν with nodes in network μ .

Fig. 2 shows all the types of connections possible for the μ node as the receiver of a randomly chosen edge in its component. Because $G_{11}(x_1, x_2) = G_{22}(x_1, x_2) = 0$, when there is no giant component, we have:

$$\begin{cases} H_{12}(x_1, x_2) = x_1 \cdot G_{12}[H_{11}(x_1, x_2), H_{21}(x_1, x_2)] \\ H_{21}(x_1, x_2) = x_2 \cdot G_{21}[H_{12}(x_1, x_2), H_{22}(x_1, x_2)]. \end{cases} \quad (6)$$

By taking the partial derivative of both sides in each sub-equation of Eq. (6) with respect to x_μ , the average component size for any $H_{\mu\nu}(x_1, x_2)$ can be calculated. Then, let $x_1 = 1$ and $x_2 = 1$ and put them into the calculation process. For the sake of brevity, we would like to use $H_\mu^\lambda(1, 1)$, $H_{\mu\nu}^\lambda(1, 1)$, $G_\mu^\lambda(1, 1)$ and $G_{\mu\nu}^\lambda(1, 1)$ instead of $\frac{\partial H_\mu(x_1, x_2)}{\partial x_\lambda} \Big|_{x_1=1, x_2=1}$, $\frac{\partial H_{\mu\nu}(x_1, x_2)}{\partial x_\lambda} \Big|_{x_1=1, x_2=1}$, $\frac{\partial G_\mu(x_1, x_2)}{\partial x_\lambda} \Big|_{x_1=1, x_2=1}$ and $\frac{\partial G_{\mu\nu}(x_1, x_2)}{\partial x_\lambda} \Big|_{x_1=1, x_2=1}$ in the following parts of this paper when there is no ambiguity. Thus:

$$\begin{cases} H_{12}^\lambda(1, 1) = 1 + G_{12}^\lambda(1, 1) \cdot H_{11}^\lambda(1, 1) \\ \quad + G_{12}^{\lambda'}(1, 1) \cdot H_{21}^\lambda(1, 1) \\ H_{21}^\lambda(1, 1) = G_{21}^\lambda(1, 1) \cdot H_{12}^\lambda(1, 1) \\ \quad + G_{21}^{\lambda'}(1, 1) \cdot H_{22}^\lambda(1, 1) \\ H_{12}^{\lambda'}(1, 1) = G_{12}^{\lambda'}(1, 1) \cdot H_{11}^\lambda(1, 1) \\ \quad + G_{12}^{\lambda''}(1, 1) \cdot H_{21}^\lambda(1, 1) \\ H_{21}^{\lambda'}(1, 1) = 1 + G_{21}^{\lambda'}(1, 1) \cdot H_{12}^\lambda(1, 1) \\ \quad + G_{21}^{\lambda''}(1, 1) \cdot H_{22}^\lambda(1, 1). \end{cases} \quad (7)$$

In Eq. (7), all the $G_{\mu\nu}^\lambda(1, 1)$ terms can be calculated based on Eqs. (4) and (5) if we know the multi-degree distribution of each network. For instance, $G_{12}^\lambda(1, 1) = \frac{k_{12} \cdot k_{11}}{k_{12}}$, where k_{11} and k_{12} have the same meanings as k_1 and k_2 in Eq. (1) and Eq. (3). Meanwhile, all the $H_{\mu\nu}^\lambda(1, 1)$ terms can

be solved in Eq. (7), which has four sub-equations and four unknowns.

The generating function for the probability distribution of a component's size of a randomly selected μ node when there is no giant component can be written as $H_\mu(x_1, x_2) = x_\mu[H_{1\mu}(x_1, x_2), H_{\mu 2}(x_1, x_2)]$, where we have:

$$\begin{cases} H_1^{\lambda'}(1, 1) = G_1^{\lambda'}(1, 1) \cdot H_{11}^{\lambda'}(1, 1) + G_1^{\lambda''}(1, 1) \cdot H_{21}^{\lambda'}(1, 1) \\ H_2^{\lambda'}(1, 1) = G_2^{\lambda'}(1, 1) \cdot H_{12}^{\lambda'}(1, 1) + G_2^{\lambda''}(1, 1) \cdot H_{22}^{\lambda'}(1, 1). \end{cases} \quad (8)$$

Because the $H_\mu^\lambda(1, 1)$ terms can be calculated by putting the results of Eq. (7) into Eq. (8), the average component size will be solved.

Thirdly, we would like to discuss the percolation threshold. Because of the absence of inner links in both networks, $H_{11}^{\lambda'}(1, 1)$, $H_{22}^{\lambda'}(1, 1)$, $H_{11}^{\lambda''}(1, 1)$ and $H_{22}^{\lambda''}(1, 1)$ should be canceled. For that reason, we can conclude that $G_{12}^{\lambda'}(1, 1)$ and $G_{21}^{\lambda'}(1, 1)$ equal 0 (because $k_{11} = k_{22} = 0$). Thus, $H_{12}^{\lambda'}(1, 1)$ and $H_{21}^{\lambda'}(1, 1)$ equal 0. Plugging $G_{11}^{\lambda'}(1, 1) = G_{11}^{\lambda''}(1, 1) = G_{22}^{\lambda'}(1, 1) = G_{22}^{\lambda''}(1, 1) = 0$ into Eq. (7), it can now be written as:

$$\begin{cases} H_{12}^{\lambda'}(1, 1) = 1 + G_{12}^{\lambda''}(1, 1) \cdot H_{21}^{\lambda'}(1, 1) \\ H_{21}^{\lambda'}(1, 1) = G_{21}^{\lambda'}(1, 1) \cdot H_{12}^{\lambda'}(1, 1). \end{cases} \quad (9)$$

Solving Eq. (9), we get:

$$\begin{cases} H_{12}^{\lambda'}(1, 1) = \frac{1}{1 - G_{21}^{\lambda'}(1, 1) \cdot G_{12}^{\lambda''}(1, 1)} \\ H_{21}^{\lambda'}(1, 1) = \frac{G_{21}^{\lambda'}(1, 1)}{1 - G_{21}^{\lambda'}(1, 1) \cdot G_{12}^{\lambda''}(1, 1)}. \end{cases} \quad (10)$$

We believe that the necessary and sufficient condition for the absence of the giant component of IERDCN 2 is:

$$G_{12}^{\lambda''}(1, 1) \cdot G_{21}^{\lambda'}(1, 1) < 1. \quad (11)$$

Finally, let us calculate the probability that a randomly chosen node belongs to the giant component. Here, $u_{\mu\nu}$ represents the probability that a randomly chosen edge pointing at a μ node leaving from a network ν node is not part of the giant component. We get:

$$\begin{cases} u_{12} = G_{12}(u_{11}, u_{21}) \\ u_{21} = G_{21}(u_{12}, u_{22}). \end{cases} \quad (12)$$

Taking G_{12} and G_{21} as known, we can solve for u_{12} and u_{21} , and then, we can calculate the probability that a randomly chosen μ node belongs to the giant component S_μ as follows:

$$\begin{cases} S_1 = 1 - G_1(u_{11}, u_{21}) \\ S_2 = 1 - G_2(u_{12}, u_{22}). \end{cases} \quad (13)$$

Above all, we discuss generating functions of the classic IERDCNs. In the following, we will show that our arithmetic is based on connection probability.

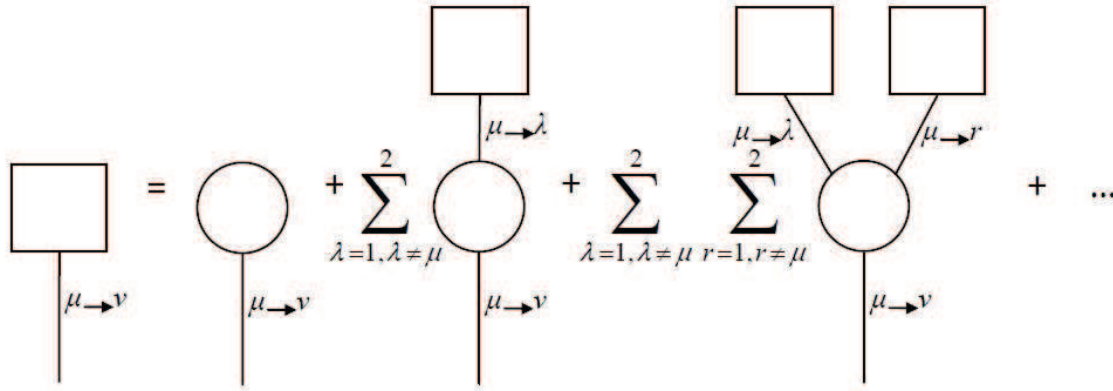


Figure 2: A diagrammatical representation of the topological constraints placed on the generating function $H_{v\mu}(x_1, x_2)$ for the distribution of sizes of components reachable by follow a randomly chosen v - μ edge. When $G_{11}(x_1, x_2) = 0$ and $G_{22}(x_1, x_2) = 0$, the functions may be expressed as this.

3.2 The generating functions of IERDCNs with connection probability

Now we are in the position to introduce the case with connection probability ω , where collaborations achieve partial success in IERDCNs. Before discussing, we use ω_μ , ω_v and $\omega_{\mu v}$ to denote, separately, the threshold value of connecting a couple of intra-network and inter-network nodes. In practice, the collaboration will not be maintained if it does not generate knowledge or techniques which can flow through networks. Nevertheless, many researchers believe that the collaboration is successful (links exist) when enterprises try to contact institutes even before collaborative agreements have been signed. To better describe this complex and dynamic phenomenon, we introduce connection probability which represents the probability of success in a collaboration. According to different situations, the connection probability of enterprises and institutes may be different. However, here they have equal collaboration choices as inner links rarely exist, so we let ω_μ and ω_v equal ω_1 , and $\omega_{\mu v}$ equal ω_2 . Based on the generating functions mentioned above, we also use G_μ to calculate the multi-degree distribution.

Firstly, connection probability was used to endow different weights for every v - μ edge, meaning that it may break some of the connections between nodes, selected randomly, from network μ and network v . Furthermore, the optimal target range of the connection probability is from 0 to ω_i , as if the connection probability is beyond ω_i , the connection will break. So, by using the probability of a randomly chosen v - μ edge to a node with excess v degrees, the generating functions of IERDCNs with connection probabilities based on Eq. (3) can be written as:

$$\left\{ \begin{array}{l} G_{11}(\omega_1 \cdot x_1, \omega_2 \cdot x_2) = \sum_{k_1=0, k_2=0}^{\infty} \frac{p_{(k_1+1)k_2}^1(k_1+1)}{\sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^1 k_1} \cdot (\omega_1 \cdot x_1)^{k_1} \cdot (\omega_2 \cdot x_2)^{k_2} \\ G_{12}(\omega_1 \cdot x_1, \omega_2 \cdot x_2) = \sum_{k_1=0, k_2=0}^{\infty} \frac{p_{k_1(k_2+1)}^1(k_2+1)}{\sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^1 k_2} \cdot (\omega_1 \cdot x_1)^{k_1} \cdot (\omega_2 \cdot x_2)^{k_2} \\ G_{21}(\omega_1 \cdot x_1, \omega_2 \cdot x_2) = \sum_{k_1=0, k_2=0}^{\infty} \frac{p_{(k_1+1)k_2}^2(k_1+1)}{\sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^2 k_1} \cdot (\omega_1 \cdot x_1)^{k_1} \cdot (\omega_2 \cdot x_2)^{k_2} \\ G_{22}(\omega_1 \cdot x_1, \omega_2 \cdot x_2) = \sum_{k_1=0, k_2=0}^{\infty} \frac{p_{k_1(k_2+1)}^2(k_2+1)}{\sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^2 k_2} \cdot (\omega_1 \cdot x_1)^{k_1} \cdot (\omega_2 \cdot x_2)^{k_2}. \end{array} \right. \quad (14)$$

The multi-degree distribution for each network may be written in the form of a generating function:

$$\left\{ \begin{array}{l} G_1(\omega_1 \cdot x_1, \omega_2 \cdot x_2) = \sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^1 \cdot (\omega_1 \cdot x_1)^{k_1} \cdot (\omega_2 \cdot x_2)^{k_2} \\ G_2(\omega_1 \cdot x_1, \omega_2 \cdot x_2) = \sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^2 \cdot (\omega_1 \cdot x_1)^{k_1} \cdot (\omega_2 \cdot x_2)^{k_2}. \end{array} \right. \quad (15)$$

Secondly, we discuss component sizes. We also have to scrutinize components of a randomly chosen undirected edge with connection probability. Let $H_{v\mu}(x_1, x_2)$ be the new generating function for the distribution of the sizes of components reached by following randomly chosen v - μ edges in the optimal target interval. A key point that needs to be emphasized is that the links between the v - μ nodes exist in 100 percent of cases.

Fig. 3 shows all the types of connections possible for the μ node as the receiver of a randomly chosen edge with some connection probability in its component. Meanwhile, let $\omega_{\lambda\mu} \cdot H_{\lambda\mu}(x_1, x_2)^{k_1} + (1 - \omega_{\lambda\mu}) \cdot H_{\lambda\mu}(x_1, x_2)^0$ denote the generating function for the distribution of the sizes of components with total probability. We explicitly exclude the giant component from $H_{\lambda\mu}(x_1, x_2)$ and get:

$$H_{\mu\nu}(x_1, x_2) = x_\mu \cdot \sum_{k_1=0, k_2=0}^{\infty} p_{k_1 k_2}^{\mu\nu} \cdot [\omega_{1\mu} \cdot H_{1\mu}(x_1, x_2)^{k_1} + 1 - \omega_{1\mu}, \omega_{\lambda\mu} \cdot H_{1\mu}(x_1, x_2)^{k_\lambda} + 1 - \omega_{\lambda\mu}] \quad (16)$$

where if $\mu = \lambda$, then $\omega_{\mu\nu} = \omega_1$, else $\omega_{\mu\nu} = \omega_2$. We could change the form of Eq. (16) using Eq. (14) and Eq. (15). Thus:

$$H_{\mu\nu}(x_1, x_2) = x_\mu \cdot G_{\mu\nu}[\omega_{1\mu} \cdot H_{1\mu}(x_1, x_2) + 1 - \omega_{1\mu}, \omega_{\lambda\mu} \cdot H_{\lambda\mu}(x_1, x_2) + 1 - \omega_{\lambda\mu}]. \quad (17)$$

If we consider starting from a randomly chosen μ node instead of a ν - μ edge, we can obtain a topology that describes the end of each edge incident to the μ node, such as the one from Fig. 3. The generating function that represents the probability distribution of component sizes can be written as:

$$H_\mu(x_1, x_2) = x_\mu \cdot G_\mu[\omega_{1\mu} \cdot H_{1\mu}(x_1, x_2) + 1 - \omega_{1\mu}, \omega_{\lambda\mu} \cdot H_{\lambda\mu}(x_1, x_2) + 1 - \omega_{\lambda\mu}]. \quad (18)$$

In IERDCNs, Eq. (17) becomes:

$$\begin{cases} H_{11}(x_1, x_2) = x_1 \cdot G_{11}[\omega_1 \cdot H_{11}(x_1, x_2) + 1 - \omega_1, \omega_2 \cdot H_{21}(x_1, x_2) + 1 - \omega_2] \\ H_{12}(x_1, x_2) = x_1 \cdot G_{12}[\omega_1 \cdot H_{11}(x_1, x_2) + 1 - \omega_1, \omega_2 \cdot H_{21}(x_1, x_2) + 1 - \omega_2] \\ H_{21}(x_1, x_2) = x_2 \cdot G_{21}[\omega_2 \cdot H_{12}(x_1, x_2) + 1 - \omega_2, \omega_1 \cdot H_{22}(x_1, x_2) + 1 - \omega_1] \\ H_{22}(x_1, x_2) = x_2 \cdot G_{22}[\omega_2 \cdot H_{12}(x_1, x_2) + 1 - \omega_2, \omega_1 \cdot H_{22}(x_1, x_2) + 1 - \omega_1] \end{cases} \quad (19)$$

and Eq. (18) becomes:

$$\begin{cases} H_1(1, 1) = x_1 \cdot G_1[\omega_1 \cdot H_{11}(1, 1) + 1 - \omega_1, \omega_2 \cdot H_{21}(1, 1) + 1 - \omega_2] \\ H_2(1, 1) = x_2 \cdot G_2[\omega_2 \cdot H_{12}(1, 1) + 1 - \omega_2, \omega_1 \cdot H_{22}(1, 1) + 1 - \omega_1]. \end{cases} \quad (20)$$

Take the partial derivatives of both sides in each sub-equation of Eqs. (14)-(20) with respect to x_1 and x_2 . Then, let $x_1 = 1$ and $x_2 = 1$ and put them into the calculation process. We have:

$$\begin{cases} H'_{11}(1, 1) = 1 + \omega_1 \cdot G'_{11}(1, 1) \cdot H'_{11}(1, 1) + \omega_2 \cdot G'_{11}(1, 1) \cdot H'_{21}(1, 1) \\ H'_{12}(1, 1) = 1 + \omega_1 \cdot G'_{12}(1, 1) \cdot H'_{11}(1, 1) + \omega_2 \cdot G'_{12}(1, 1) \cdot H'_{21}(1, 1) \\ H'_{21}(1, 1) = \omega_2 \cdot G'_{21}(1, 1) \cdot H'_{12}(1, 1) + \omega_1 \cdot G'_{21}(1, 1) \cdot H'_{22}(1, 1) \\ H'_{22}(1, 1) = \omega_2 \cdot G'_{22}(1, 1) \cdot H'_{12}(1, 1) + \omega_1 \cdot G'_{22}(1, 1) \cdot H'_{22}(1, 1) \\ H'_{11}(1, 1) = \omega_1 \cdot G'_{11}(1, 1) \cdot H'_{11}(1, 1) + \omega_2 \cdot G'_{11}(1, 1) \cdot H'_{21}(1, 1) \\ H'_{12}(1, 1) = \omega_1 \cdot G'_{12}(1, 1) \cdot H'_{11}(1, 1) + \omega_2 \cdot G'_{12}(1, 1) \cdot H'_{21}(1, 1) \\ H'_{21}(1, 1) = 1 + \omega_2 \cdot G'_{21}(1, 1) \cdot H'_{12}(1, 1) + \omega_1 \cdot G'_{21}(1, 1) \cdot H'_{22}(1, 1) \\ H'_{22}(1, 1) = 1 + \omega_2 \cdot G'_{22}(1, 1) \cdot H'_{12}(1, 1) + \omega_1 \cdot G'_{22}(1, 1) \cdot H'_{22}(1, 1) \end{cases} \quad (21)$$

and

$$\begin{cases} H'_1(1, 1) = 1 + G'_1(1, 1) \cdot \omega_1 \cdot H'_{11}(1, 1) + G'_1(1, 1) \cdot \omega_2 \cdot H'_{21}(1, 1) \\ H'_2(1, 1) = G'_2(1, 1) \cdot \omega_1 \cdot H'_{12}(1, 1) + G'_2(1, 1) \cdot \omega_2 \cdot H'_{22}(1, 1) \\ H'_1(1, 1) = G'_1(1, 1) \cdot \omega_2 \cdot H'_{12}(1, 1) + G'_1(1, 1) \cdot \omega_1 \cdot H'_{22}(1, 1) \\ H'_2(1, 1) = 1 + G'_2(1, 1) \cdot \omega_2 \cdot H'_{12}(1, 1) + G'_2(1, 1) \cdot \omega_1 \cdot H'_{22}(1, 1). \end{cases} \quad (22)$$

Then, $H'_\mu(1, 1)$ can be calculated by putting the results of Eq. (21) into Eq. (22). Now, the average component size will be solved.

Thirdly, we would like to discuss the percolation threshold. By solving Eq. (21), we can get:

$$\begin{cases} H'_{11}(1, 1) = \frac{1 - \omega_1 \cdot G'_{22}(1, 1) + \omega_2 \cdot (G'_{11}(1, 1) - G'_{12}(1, 1)) \cdot [G'_{21}(1, 1) \cdot (1 - \omega_1 \cdot G'_{22}(1, 1)) + \omega_1 \cdot G'_{21}(1, 1) \cdot G'_{22}(1, 1)]}{\theta} \\ H'_{12}(1, 1) = \frac{(1 - \omega_1 \cdot G'_{22}(1, 1)) \cdot (1 - \omega_1 \cdot G'_{11}(1, 1) + \omega_1 \cdot G'_{12}(1, 1))}{\theta} \\ H'_{21}(1, 1) = \frac{(1 - \omega_1 \cdot G'_{11}(1, 1) + \omega_1 \cdot G'_{12}(1, 1)) \cdot \omega_2 \cdot [G'_{21}(1, 1) \cdot (1 - \omega_1 \cdot G'_{22}(1, 1)) + \omega_1 \cdot G'_{21}(1, 1) \cdot G'_{22}(1, 1)]}{\theta} \\ H'_{22}(1, 1) = \frac{G'_{22}(1, 1) \cdot \omega_2 \cdot (1 - \omega_1 \cdot G'_{11}(1, 1) + \omega_1 \cdot G'_{12}(1, 1))}{\theta} \\ H'_{11}(1, 1) = \frac{G'_{11}(1, 1) \cdot \omega_2 \cdot (1 - \omega_1 \cdot G'_{22}(1, 1) + \omega_1 \cdot G'_{21}(1, 1))}{\sigma} \\ H'_{12}(1, 1) = \frac{\omega_2 \cdot (1 - \omega_1 \cdot G'_{22}(1, 1) + \omega_1 \cdot G'_{21}(1, 1)) \cdot [G'_{11}(1, 1) \cdot (1 - \omega_1 \cdot G'_{11}(1, 1) + \omega_1 \cdot G'_{12}(1, 1)) + G'_{12}(1, 1)]}{\sigma} \\ H'_{21}(1, 1) = \frac{(1 - \omega_1 \cdot G'_{11}(1, 1)) \cdot (1 - \omega_1 \cdot G'_{22}(1, 1) + \omega_1 \cdot G'_{21}(1, 1))}{\sigma} \\ H'_{22}(1, 1) = \frac{1 - \omega_1 \cdot G'_{11}(1, 1) + \omega_2 \cdot (G'_{21}(1, 1) - G'_{22}(1, 1)) \cdot [G'_{11}(1, 1) \cdot (\omega_1 \cdot G'_{11}(1, 1) - 1) - \omega_1 \cdot G'_{11}(1, 1) \cdot G'_{12}(1, 1)]}{\sigma} \end{cases} \quad (23)$$

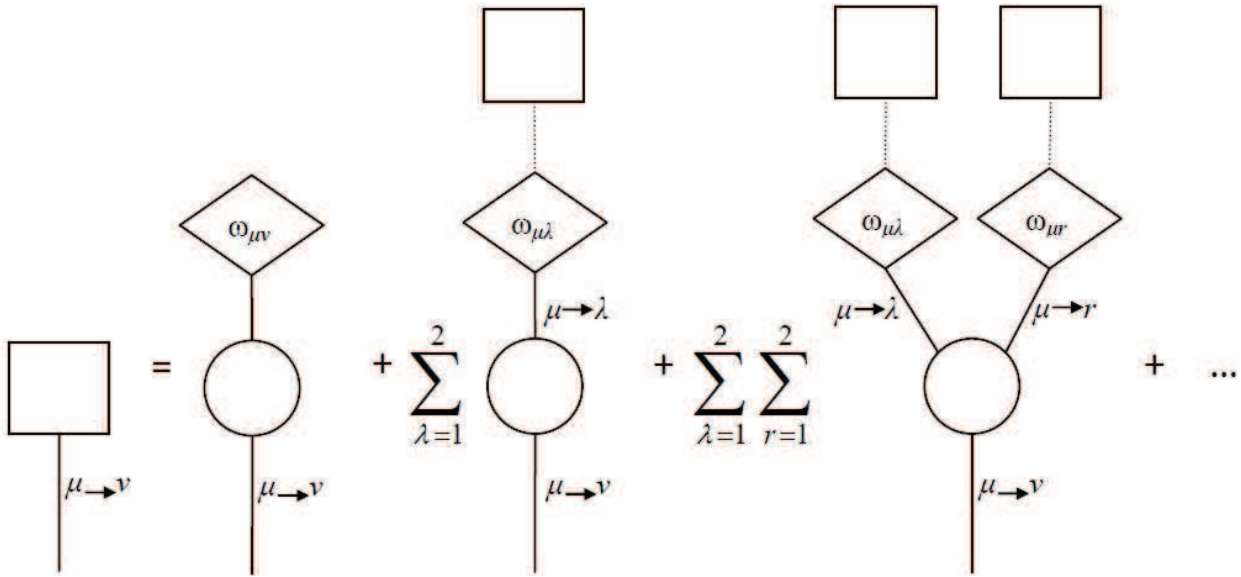


Figure 3: A diagrammatical representation of the topological constraints placed on the generating function $H_{\mu\nu}(\omega_1 \cdot x_1, \omega_2 \cdot x_2)$ for the distribution of the sizes of components reachable by following a randomly chosen v - μ edge, where connection probability decides all the links in IERDCNs.

where

$$\begin{aligned} \theta = & (1 - \omega_1 \cdot G'_{11}(1, 1)) \cdot (1 - \omega_1 \cdot G'_{22}(1, 1)) \cdot (1 - \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1)) \\ & - \omega_1 \cdot G'_{11}(1, 1) \cdot \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1) \cdot (1 - \omega_1 \cdot G'_{22}(1, 1)) \\ & - \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1) \cdot \omega_1 \cdot G'_{22}(1, 1) \cdot (1 - \omega_1 \cdot G'_{11}(1, 1)) \\ & - \omega_1 \cdot G'_{11}(1, 1) \cdot \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1) \cdot \omega_1 \cdot G'_{22}(1, 1) \end{aligned} \quad (24)$$

and

$$\begin{aligned} \sigma = & (1 - \omega_1 \cdot G'_{11}(1, 1)) \cdot (1 - \omega_1 \cdot G'_{22}(1, 1)) \cdot (1 - \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1)) \\ & - \omega_1 \cdot G'_{11}(1, 1) \cdot \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1) \cdot (1 - \omega_1 \cdot G'_{22}(1, 1)) \\ & - \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1) \cdot \omega_1 \cdot G'_{22}(1, 1) \cdot (1 - \omega_1 \cdot G'_{11}(1, 1)) \\ & - \omega_1 \cdot G'_{11}(1, 1) \cdot \omega_2 \cdot G'_{12}(1, 1) \cdot \omega_2 \cdot G'_{21}(1, 1) \cdot \omega_1 \cdot G'_{22}(1, 1). \end{aligned} \quad (25)$$

Eqs. (23)-(25) show that the necessary and sufficient condition for the absence of a system-wide giant component is:

$$\begin{cases} \theta > 0 \wedge \omega_1 \cdot G'_{11}(1, 1) < 1 \wedge \omega_1 \cdot G'_{22}(1, 1) < 1 \wedge \omega_2^2 \cdot G'_{12}(1, 1) \cdot G'_{21}(1, 1) < 1 \\ \sigma > 0 \wedge \omega_1 \cdot G'_{11}(1, 1) < 1 \wedge \omega_1 \cdot G'_{22}(1, 1) < 1 \wedge \omega_2^2 \cdot G'_{12}(1, 1) \cdot G'_{21}(1, 1) < 1. \end{cases} \quad (26)$$

To prove sufficiency, we apply these conditions to Eq. (23) and find that according Eq. (14) and Eq. (21), $H'_{11}(1, 1) = \frac{1 + \omega_2 \cdot G'_{11}(1, 1) \cdot H'_{21}(1, 1)}{1 - \omega_1 \cdot G'_{11}(1, 1)}$, $G'_{\mu\nu}(1, 1) \equiv \bar{k}_{\mu\nu} \geq 0$ and $G'_{11}(1, 1)$ almost equal to 0 (because of the rarity of inner links), so $H'_{11}(1, 1)$ converges to a value equal to or higher than 1. Furthermore, $H'_{12}(1, 1)$ converges to a value equal to or higher than 1, and $H'_{21}(1, 1)$ and $H'_{22}(1, 1)$ converge to values equal to or higher than 0. In the same way, $H'_{11}(1, 1)$ and $H'_{12}(1, 1)$ converge to values equal to or higher than 0. Both $H'_{21}(1, 1)$ and $H'_{22}(1, 1)$ converge to values equal to or higher than 1. For all parameters, when Eq. (26) is satisfied, sufficiency is proved. When $\omega_1 \cdot G'_{11}(1, 1) \geq 1 \vee \omega_1 \cdot G'_{22}(1, 1) \geq 1 \vee \omega_2 \cdot G'_{12}(1, 1) \geq 1 \vee \omega_2 \cdot G'_{21}(1, 1) \geq 1$, a part of the whole network has already undergone the transition phase. Otherwise, a giant component appears. Considering the sufficiency of Eq. (25), which has been proven, we have that it is the necessary and sufficient condition for the absence of a system-wide giant component.

Finally, the probability that randomly chosen nodes and the fraction of network nodes that belong to the giant component will be calculated. Once a giant component appears, we can calculate the properties of the components not belonging to it. As Eq. (12) shows, $u_{\mu\nu}$ represents the probability that a randomly chosen edge pointing at a μ node leaving from a network ν node is not part of the giant component. As we mentioned, whether or not each edge exists depends on the connection probability, and our work, therefore, must ensure that all randomly chosen ν - μ edges do not belong to the giant component. Now, letting $G_{\mu\nu}$ remain the same, we divide the parameters of $G_{\mu\nu}$ into two parts. As a probability, both parts are used for calculating the distribution of outgoing edges excluding the giant component. When the connection probability equals to $1 - \omega_{\mu\nu}$, neither an outgoing edge nor a giant component will appear in all networks. On the contrary, when the connection probability equals to $\omega_{\mu\nu}$, which means that an outgoing edge exists, and according to Eq. (15), the probability that outgoing edges do not belonging to the giant component is $\omega_{\mu\nu} \cdot u_{\mu\nu}$. Thus, the probability that a randomly chosen node does not belong to the giant component is $1 - \omega_{\mu\nu} + \omega_{\mu\nu} \cdot u_{\mu\nu}$ and we get:

$$\begin{cases} u_{11} = G_{11}(1 - \omega_1 + \omega_1 \cdot u_{11}, 1 - \omega_2 + \omega_2 \cdot u_{21}) \\ u_{12} = G_{12}(1 - \omega_1 + \omega_1 \cdot u_{11}, 1 - \omega_2 + \omega_2 \cdot u_{21}) \\ u_{21} = G_{21}(1 - \omega_2 + \omega_2 \cdot u_{12}, 1 - \omega_1 + \omega_1 \cdot u_{22}) \\ u_{22} = G_{22}(1 - \omega_2 + \omega_2 \cdot u_{12}, 1 - \omega_1 + \omega_1 \cdot u_{22}) \end{cases} \quad (27)$$

and

$$\begin{cases} u_1 = G_1(1 - \omega_1 + \omega_1 \cdot u_{11}, 1 - \omega_2 + \omega_2 \cdot u_{21}) \\ u_2 = G_2(1 - \omega_2 + \omega_2 \cdot u_{12}, 1 - \omega_1 + \omega_1 \cdot u_{22}). \end{cases} \quad (28)$$

Let S_μ be the fraction of μ -nodes belonging to the giant component, and it may be written as:

$$\begin{cases} S_1 = 1 - u_1 \\ S_2 = 1 - u_2. \end{cases} \quad (29)$$

Our discussion for this case is general and applicable to randomly connecting networks.

4 Applications

4.1 Comparison with actual networks

In this section, we apply our mathematical framework to two actual IERDCNs with 100 percent connection probabilities mentioned in Section 2. As Table 3 shows, there are

some discrepancies in the first one, but the discrepancies between the theoretical calculations and the actual IERDCNs is proved to be acceptable. Meanwhile, in the second one, the analytical results completely match the empirical data. We will discuss this in more detail and see how precise our arithmetic is. For IERDCN 1, we get $G'_{11}(1, 1) = 0$, $G'_{11}(1, 1) = 0.125$, $G'_{12}(1, 1) = 0.0142857143$, $G'_{12}(1, 1) = 0.8571428571$, $G'_{21}(1, 1) = 6.2$, $G'_{21}(1, 1) = 0.1$, $G'_{22}(1, 1) = 3.5$, and $G'_{22}(1, 1) = 0$. Obviously, given that $G'_{12}(1, 1) \cdot G'_{21}(1, 1) = 5.3142857140 > 1$, a giant connected component exists according to Eq. (26) when $\omega_2 = 1$. The two values from the analytical results and empirical data are extremely close to each other. For IERDCN 2, we get $G'_{12}(1, 1) = 1$, $G'_{21}(1, 1) = 1$, and $G'_{12}(1, 1) \cdot G'_{21}(1, 1) = 1$ which disagree with Eq. (11) and Eq. (26) when $\omega_1 = 0$ and $\omega_2 = 1$. Thus, IERDCN 2 also has a giant connected component.

However, according to our arithmetic, we can give two cases with a smaller connection probability, and by letting $\omega_1 = 0.1$, $\omega_2 = 0.1$ and $\omega_1 = 0.1$, $\omega_2 = 0.2$

$$\text{for IERDCN 1, we can get: } \begin{cases} G'_{11}(1, 1) = 0 \\ G'_{11}(1, 1) = 0.125 \\ G'_{12}(1, 1) = 0.0143 \\ G'_{12}(1, 1) = 0.8571 \\ G'_{21}(1, 1) = 6.2 \\ G'_{21}(1, 1) = 0.1 \\ G'_{22}(1, 1) = 3.5 \\ G'_{22}(1, 1) = 0 \end{cases} \quad \text{and}$$

$$\begin{cases} H'_{11}(1, 1) = 1.0083 \\ H'_{11}(1, 1) = 0 \\ H'_{12}(1, 1) = 1.0580 \\ H'_{12}(1, 1) = 0.0028 \\ H'_{21}(1, 1) = 0.6597 \\ H'_{21}(1, 1) = 1.9539 \\ H'_{22}(1, 1) = 0.3703 \\ H'_{22}(1, 1) = 1.5385 \end{cases}, \begin{cases} G'_{11}(1, 1) = 0 \\ G'_{11}(1, 1) = 0.125 \\ G'_{12}(1, 1) = 0.0143 \\ G'_{12}(1, 1) = 0.8571 \\ G'_{21}(1, 1) = 6.2 \\ G'_{21}(1, 1) = 0.1 \\ G'_{22}(1, 1) = 3.5 \\ G'_{22}(1, 1) = 0 \end{cases} \quad \text{and}$$

$$\begin{cases} H'_{11}(1, 1) = 1.0397 \\ H'_{11}(1, 1) = 0 \\ H'_{12}(1, 1) = 1.2738 \\ H'_{12}(1, 1) = 0.0084 \\ H'_{21}(1, 1) = 1.5841 \\ H'_{21}(1, 1) = 1.9541 \\ H'_{22}(1, 1) = 0.8916 \\ H'_{22}(1, 1) = 1.5385. \end{cases}$$

And at this point, we have:

$$\theta = 1 > 0 \wedge G'_{11}(1, 1) = 0 < 1 \wedge G'_{22}(1, 1) = 0 < 1 \wedge 0.1 \cdot G'_{12}(1, 1) \cdot 0.1 \cdot G'_{21}(1, 1) = 0.0531 < 1$$

and

Table 3: Analytical results and empirical data of IERDCN 1 and 2.

	IERDCN 1		IERDCN 2	
	S_1	S_2	S_1	S_2
Analytical results	0.8174	0.9582	1.0000	1.0000
Empirical data	0.8200	1.0000	1.0000	1.0000

$$\theta = 1 > 0 \wedge G'_{11}(1, 1) = 0 < 1 \wedge G'_{22}(1, 1) = 0 < 1 \wedge 0.2 \cdot G'_{12}(1, 1) \cdot 0.2 \cdot G'_{21}(1, 1) = 0.2126 < 1$$

Because both of them satisfy Eq. (26), there is no giant component. Now, we can calculate the average component size of a randomly selected μ network. By taking into account all possible situations, we work out the average true values of component size. Table 4 shows the analytical values and true values of component size. We find that all pairs have a tolerated discrepancy (approximately 0.67%-1.4%) which is lower than the tolerance of 1 node. These discrepancies may come from the inherent inaccuracy of the generating function. However, our arithmetic is precise enough to estimate whether or not the networks have undergone the transition phase and to calculate important structural property measures.

4.2 Application in actual networks

All the work we do possesses theoretical meaning, and also has a very strong meaning in practice.

First of all, the component of nodes utilizes the resources of the networks most efficiently by competing inside and collaborating outside the network, which provides a path for every node in the component to learn some technique or knowledge from others without directly working together. As Table 3 shows, in IERDCN 2, every enterprise and institute has a path to the giant component. Original technological innovation ideas may come from every enterprise and institute and must be done in partnership. As for the links within the R&D network of enterprises or institutes, it can be found that even if we remove all of them, the network still has a giant component. This means that the necessary knowledge of technological R&D can flow freely throughout the whole network and that there is a high rate of success for patent output.

Secondly, we introduce connection probability to extend the application of generating functions. Currently, links in many networks are unstable in the sense that they can be changed with time or human will. As far as social networks go, links between one person and another person will be weak and can even disappear due to mistrust,

alienation, and sabotage. Likewise, in IERDCNs, a connection probability exists in all collaborations. In the collaboration process, everyone likes a partner with powerful research capabilities, and never works together when either side of the R&D collaboration fails. Additionally, enterprises have some interest in collaborating with institutes for technological R&D. As Table 4 and all parameters show in the previous part of this section, this means there is no giant component and connection probability that can make the component size smaller. Furthermore, self-governing choices of enterprises and institutes in practice reduce the chance that a lot of enterprises collaborate with the same institute, which is relatively more conducive to the technological R&D innovation of the whole network. So, by using connection probability, we can explain the process of autonomous choice and better describe the complicated network environment in the real world.

In addition, connection probability could explain the disposable link phenomenon where if one link is used, it cannot be used again. Such as in IERDCNs, any side of the network can establish a partnership with each other, but the collaboration is not successful without the outputting of a patent or new technique. Some scholars such as Fu et al. [16] use directed interacting networks for discussing this issue. However, the unique contribution of our mathematical framework and arithmetic is that we provide another method to solve the situation where there is an outgoing edge with no return or opposite. Connection probability may be seen as a precondition for whether or not a R&D collaboration can generate a new technique, and its value could refer to the average success rate of the R&D collaboration with knowledge diffusion. Furthermore, it could also give some basis for reference for government to work out innovation policies. To percolate and promote R&D collaboration to output as many new technologies and patents as possible, the government could moderately increase or reduce the connection probability. It is obvious that collaboration between enterprises and institutes is important to technological R&D because of their complementarily advantages. So, we always have $\omega_2 > \omega_1$ in IERDCNs as previously mentioned. After the results listed in Table 4, we give different analytical and simulation results (50 times) under four pairs of connection probabilities (see Table 5 and Table 6). Neither Table 4 nor Table 5 has a giant component. However, Table 6 shows that the network has already undergone the transition phase. Comparing Table 6 with Table 5, we find that the relative growth rate of all average component sizes under ω_2 is larger than under ω_1 . Furthermore, obviously, $H'_1(1, 1)$ and $H'_2(1, 1)$ independently increase from three nodes to 11 nodes when

Table 4: Analytical results and empirical data of IERDCN 1.

Enterprise R&D network	When $\omega_1 = 0.1$ and $\omega_2 = 0.1$		When $\omega_1 = 0.1$ and $\omega_2 = 0.2$	
	$H_1^{'1}(1, 1)$	$H_1^{'2}(1, 1)$	$H_1^{'1}(1, 1)$	$H_1^{'2}(1, 1)$
Analytical results	1.0396	0.1172	1.1906	0.2345
Empirical data	1.0250	0.1200	1.1735	0.2285
Institute R&D network	When $\omega_1 = 0.1$ and $\omega_2 = 0.1$		When $\omega_1 = 0.1$ and $\omega_2 = 0.2$	
	$H_2^{'1}(1, 1)$	$H_2^{'2}(1, 1)$	$H_2^{'1}(1, 1)$	$H_2^{'2}(1, 1)$
Analytical results	0.5202	1.0014	1.2526	1.0072
Empirical data	0.5000	1.0000	1.2417	1.0000

changing ω_2 from 0.3 to 0.4, while remaining the same when changing ω_1 from 0.1 to 0.2. Similarly, the discrepancy between the analytical results and simulation results enlarges with the increase in connection probability and the precision of the results change from 0.26% to 0.5% and 0.09% to 0.48%, respectively. All results are given to show the effectiveness and applicability of the proposed arithmetic.

Finally, using our arithmetic, it is easy to calculate every percolation threshold under some fixed connection probability. As Fig. 4 shows, we give two groups of values to discuss how the connection probability of different networks affects the percolation of the whole network. The threshold value (0.43) of IERDCN 1 is easy to calculate according to Eq. (26) which we have discussed. Fig. 4(a) indicates that the probability that a randomly chosen node belongs to the giant component is hardly influenced by ω_1 . Instead, Fig. 4(b) indicates that ω_2 is the main influencer. When we let ω_2 equal some fixed value and ω_1 runs from 0 to 1, S_1 and S_2 measured a tiny change. However, if we let ω_1 equal some fixed value and ω_2 runs from 0.4301 to 1, we can see double inward curves and the curvature of S_2 changes greatly. For this reason, we believe that the inter-firm competition and desire for new techniques makes collaborations between institutes and enterprises very important to IERDCNs and IERDCN 2 is considered as an ideal state for both sides. In other words, every enterprise can find one or two institutes to partner with, while institutes have equal opportunities to collaborate with enterprises. There two situations that may occur such that they are depicted as IERDCNs 1 and 2. In the first one, both sides are strongly inclined to collaborate with a famous and technologically strong partner, which is very much like the celebrity effect. In another, technological monopoly advantages are difficult to maintain with the flow of knowledge and techniques in the industry's network, especially because institutes accumulate a wealth of R&D experience and master the advanced technology in the industry's field after a period of collabora-

tion. Additionally, some enterprises and institutes have an established long-term R&D collaborative relationship, which helps the network achieve the ideal state. We find that the lower the interdependence between enterprises or institutes, the higher the percolation in IERDCNs. That's why networks have a giant component with a lesser inner connection probability and a higher interactive one. Additionally, it is common in many high-tech industries, such as in IERDCN 1, that only 30 links between enterprises and institutes is enough to maintain the supercritical regime.

5 Conclusion

Connection probability is able to consider pairs of nodes with a temporary edge but that cannot effectively connect to each other. The connection probability is so important to a dynamically evolving network of R&D collaboration, as the R&D collaboration network is a knowledge dissemination network. From the perspective of the success or failure of a collaboration, connection probability illustrates that knowledge and techniques cannot be disseminated with invalid collaboration.

Our mathematical framework and the arithmetic discussed in this paper are a near-perfect depiction of the real state of IERDCNs and are accurate, as the results mentioned in Section 4 show. In this case, despite the sizes of IERDCNs being not very large and the existence of discrepancies, it does not greatly damage the validity of the arithmetic. So, we may revise generating larger networks using simulations for a while. However, we have simulated 50 networks randomly to study, and sometimes the discrepancies are artificially enlarged. We believe that discrepancies will likely be quite minimal if all samples are considered.

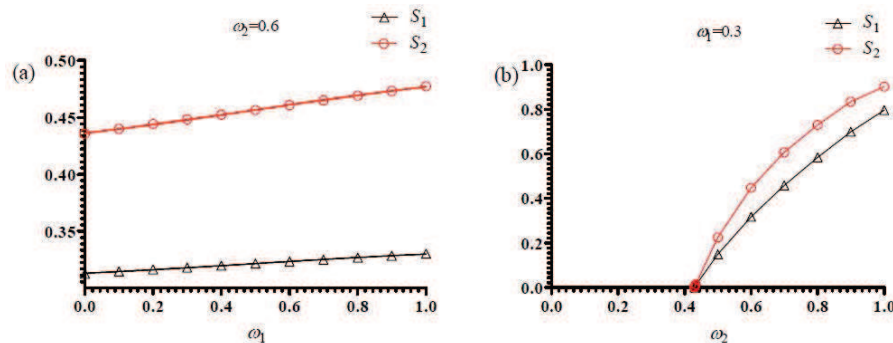
In this study, we have investigated two types of mathematical frameworks and arithmetic based on generating functions for analyzing IERDCNs. Furthermore, we have

Table 5: Analytical and simulation results of IERDCN 1 with no giant component.

	$\omega_1 = 0.1, \omega_2 = 0.3$				$\omega_1 = 0.1, \omega_2 = 0.4$			
	$H_1^1(1, 1)$	$H_1^2(1, 1)$	$H_2^1(1, 1)$	$H_2^2(1, 1)$	$H_1^1(1, 1)$	$H_1^2(1, 1)$	$H_2^1(1, 1)$	$H_2^2(1, 1)$
Analytical results (S_1)	1.6498	0.3517	2.8465	1.0124	5.1414	0.4690	13.6074	1.0220
Simulation results (S_1')	1.6405	0.3492	2.8371	1.0000	5.1365	0.4635	13.5233	1.0000
	$\omega_1 = 0.2, \omega_2 = 0.3$				$\omega_1 = 0.2, \omega_2 = 0.4$			
	$H_1^1(1, 1)$	$H_1^2(1, 1)$	$H_2^1(1, 1)$	$H_2^2(1, 1)$	$H_1^1(1, 1)$	$H_1^2(1, 1)$	$H_2^1(1, 1)$	$H_2^2(1, 1)$
Analytical results (S_1)	1.6579	0.9241	2.8661	1.0325	5.3193	1.2323	14.1124	1.0577
Simulation results (S_1')	1.6463	0.9200	2.8401	1.0167	5.3115	1.2200	14.0067	1.0167

Table 6: Analytical and simulation results of IERDCN 1 with a giant component.

	$\omega_1 = 0.4, \omega_2 = 0.5$		$\omega_1 = 0.5, \omega_2 = 0.6$		$\omega_1 = 0.6, \omega_2 = 0.7$	
	S_1	S_2	S_1	S_2	S_1	S_2
Analytical results (S_1)	0.1522	0.2302	0.3214	0.4565	0.4633	0.6193
Simulation results (S_1')	0.1518	0.2300	0.3224	0.4650	0.4656	0.6350
	$\omega_1 = 0.7, \omega_2 = 0.8$		$\omega_1 = 0.8, \omega_2 = 0.9$		$\omega_1 = 0.9, \omega_2 = 1.0$	
	S_1	S_2	S_1	S_2	S_1	S_2
Analytical results (S_1)	0.5905	0.7471	0.7074	0.8554	0.8156	0.9541
Simulation results (S_1')	0.5940	0.7800	0.7112	0.8950	0.8198	1.0000

**Figure 4:** Analytical results and simulation results of S versus different ω_1 s and ω_2 s, where S denotes the probability that a randomly chosen node is a part of the giant component. In this figure (a) shows the probability when $\omega_2 = 0.6$ and ω_1 changes from 0 to 1. And (b) shows the probability when $\omega_1 = 0.3$ and ω_2 changes from 0.4301 to 1.

given the necessary and sufficient conditions for the absence of system-wide giant components. Using our arithmetic, we can calculate the corresponding parameters in the sub-critical and supercritical regimes. Using two actual IERDCNs, we discussed the application and validity of our mathematical framework and arithmetic. It is quite clear that the interactive connection probability of networks is a determining factor of percolation, while the inner connection probability has less influence. Some reasonable and helpful advice is given to promote regional technological R&D and innovation. By adjusting the probability values, we also found that the supercritical regime

of the whole network is maintained mainly by collaboration between enterprises and institutes.

Acknowledgement: This work was partly supported by Foundation of China education Ministry for Youth under Grant No.14YJC630035.

References

- [1] R. Albert, H. Jeong, A. L. Barabasi, Nature 401, 130 (1999)
- [2] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts, Phys. Rev. Lett. 85, 5468 (2000)
- [3] M. E. J. Newman, Proc. Natl. Acad. Sci. 98, 404 (2001)

- [4] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, E. D. Gilles, *Nature* 420, 190 (2002)
- [5] C. M. Song, S. Havlin, H. A. Makse, *Nature* 433, 392 (2005)
- [6] J. Nagler, A. Levina, M. Timme, *Nat. Phys.* 7, 265 (2011)
- [7] Y. Y. Liu, E. Csoka, H. J. Zhou, M. Posfai, *Phys. Rev. Lett.* 109, 205703 (2012)
- [8] E. A. Leicht, R. M. D'Souza, *arXiv:0907.0894 [cond-mat.dis-nn]*
- [9] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, S. Havlin, *Nature* 464, 1025 (2010)
- [10] R. Parshani, S. V. Buldyrev, S. Havlin, *Phys. Rev. Lett.* 105, 048701 (2010)
- [11] S. V. Buldyrev, N. W. Shere, G. A. Cwlich, *Phys. Rev. E* 83, 016112 (2011)
- [12] J. Shao, S. V. Buldyrev, S. Havlin, H. E. Stanley, *Phys. Rev. E* 83, 036116 (2011)
- [13] J. X. Gao, S. V. Buldyrev, H. E. Stanley, S. Havlin, *Nat. Phys.* 8, 40 (2012)
- [14] A. Bashan, R. Parshani, S. Havlin, *Phys. Rev. E* 83, 051127 (2011)
- [15] X. Q. Huang, J. X. Gao, S. V. Buldyrev, S. Havlin, H. E. Stanley, *Phys. Rev. E* 83, 065101 (2011)
- [16] T. Fu, Y. N. Chen, Z. Qin, L. P. Guo, *Physica A* 392, 2807 (2013)
- [17] S. Valverde, R. V. Sole, M. A. Bedau, N. Packard, *Phys. Rev. E* 76, 056118 (2007)
- [18] M. E. J. Newman, S. H. Strogatz, D. J. Watts, *Phys. Rev. E* 64, 026118 (2001)
- [19] M. E. J. Newman, *Phys. Rev. E* 66, 016128 (2002)
- [20] T. Kalisky, R. Cohen, *Phys. Rev. E* 73, 035101 (2006)
- [21] J. X. Gao, S. V. Buldyrev, S. Havlin, H. E. Stanley, *Phys. Rev. Lett.* 107, 195701 (2011)