INTERNATIONAL JOURNAL OF SOCIAL SCIENCE, INNOVATION AND EDUCATIONAL TECHNOLOGIES (ONLINE) - ISSN: 2717-7130

Vol:1, Issue: 3 pp: 229-236

JEL Codes: C8, L9

WANG C., SHEN A., GUO J, (2020). "**Research on Cascading Failure in Complex Networks Based on Generative Functions**", Vol: 1 Issue: 3 pp: 229-236

Keywords: Complex Networks, Cascading Failure, Robustness

Article Type Research Article

Research on Cascading Failure in Complex Networks Based on Generative Functions

Arrived Date	Accepted Date	Published Date
06.07.2020	16.07.2020	31.07.2020

Chao Wang*, Anwei Shen†, Jilian Guo‡

ABSTRACT

Interdependent networks have become a hot research direction in the field of complex networks in recent years. In order to study the robustness of interdependent networks, firstly, a generation function was introduced as a tool for complex network calculations. A single complex network generation function was analyzed, combined with the interdependent network modeling mechanism. The theory of robustness of interdependent networks based on generative functions is presented. The results show that as the removal ratio p increases, the remaining giant component P_∞ is a function of the removal ratio p and presents a first-order phase transition; as the dependence node ratio q decreases, The remaining giant component P_∞ appears as a second-order phase change with respect to the removal ratio p.

INTRODUCTION

After decades of development, the theoretical system of complex network research has been gradually established and improved, but these theories still have certain limitations. Because most previous researches on complex networks are based on isolated networks, in the real world each network has more or less various relationships with other networks, such as physical dependence, strict logic, isolated networks, information exchange. An isolated network in the strict sense does not exist. Once a node in a network fails, the connections among the networks will propagate and amplify the impact of the failure and may affect the functions of other networks. In the end, a small failure may have catastrophic consequences for the entire system.

Robert et al. (2008) provided a certain operational means for the management and control of interdependent networks. Reed et al. (2009) provided a quantitative method for the vulnerability of interdependent networks through empirical research. In the same year, Bagheri et al. (2010) built a reference model of the interdependent infrastructure network on this basis. Buldyrev (2011) and others first proposed a fully interdependent dependent network model, that is, there is a one-to-one interdependent relationship between nodes of two networks. In this case, when a node in one network fails, nodes that depend on it in another network will also fail, which opens a new direction for research on the problem of cascading faults in interdependent networks. Since then, Johansson et al. (2010) and Gao et al. (2011) a large number of research results have emerged in the study of interdependence networks.

* Dr., Air Force Engeeering University, cc_970110@163.com, Xi'an/CHINA

† Shen A., Air Force Engeeering University, Xi'an /CHINA

‡ Guo J., Air Force Engeeering University, Xi'an /CHINA

ANALYSIS OF CASCADING FAILURES OF SINGLE NETWORK

First, the process of single complex network cascading failures is introduced. The robustness of a single complex network refers to the ability of the entire network to maintain its structural and functional integrity after the node is destroyed.

Therefore, the basic idea of complex network cascading failures theory and simulation analysis is to change the number or proportion of nodes destroyed, and then study the ability of the network to maintain robustness at different scales.

After some nodes in the general network are destroyed, they will show two forms. One is that there is a giant component, then the network structure is considered to be functional, as shown in Figure 1(a); the other is that there is no giant component. The network is considered to be completely paralyzed, as shown in Figure 1(b).

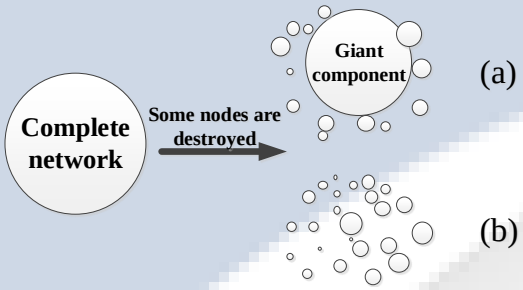

Figure 1. Two ways in which the network may fail

The parameter that measures the robustness of a complex network is generally the scale of the giant component remaining after the cascading failure. As shown in Figure 1(a), there are 16 nodes in the network. Now remove 3 red nodes. The remaining giant components in the network are the yellow shaded parts in Figure 2. If the proportion of nodes removed in the network is defined as p, then the proportion of reserved nodes is $1-p$, and the proportion of the remaining giant components to the total number of nodes in the network after the cascading failures occurs is P_∞. For a complex network with an infinite number of nodes, when changing the proportion p of nodes removed in the network, the change of the proportion of nodes remaining in the complex network is shown in Figure 1(b). It can be seen that only when the proportion $p \le p_c$ of the nodes is removed, the network giant components are retained, and the system presents a second-order phase transition. This key point p_c in the network giant component is called the seepage threshold and is another important parameter to measure the robustness of the network. When the value of the seepage threshold is larger, the robustness of the network is better, and it can resist a large number of node deletions; when the value of the seepage threshold is small, the robustness of the network is worse.

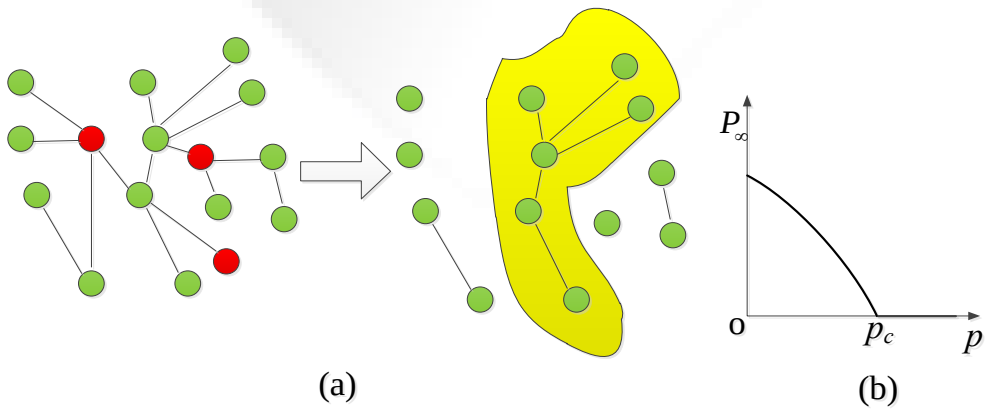

Figure 2. Schematic diagram of robust network robustness parameters

Basic concepts of generating functions

For any set of sequences $a_n, n=0,1,2\cdots$, $\psi(\mathrm{z})=\sum_{n=0} a_n z^n$ can be called a generating function of the sequence a_n. Obviously, the sequence a_n is a discrete sequence, but $\psi(\mathrm{z})$ is a continuous function, and the two are connected through a z-transform. If we already have a set of generator function expressions for a series, we can Taylor expand it and display it in the form of a series to know the value of each item in the series.

$$a_n=\frac{1}{n!}\frac{d^n\psi(\mathrm{z})}{dz^n}\bigg|_{z=0}=\frac{1}{2\pi i}\oint\frac{d\psi(\mathrm{z})}{z^{n+1}} \tag{1}$$

In the above formula, when $z=1$, it is obvious that the sum $S=\psi(1)$ of this series is obtained. And can also get some meaningful conclusions:

$$\sum_{n=1} a_{n-1}z^n = z\psi(\mathrm{z}) \tag{2}$$

$$\sum_{n=0} a_{n+1}z^n = \frac{1}{z}\left[\psi(\mathrm{z})-a_0\right] \tag{3}$$

$$\sum_{n=1} a_n n z^{n-1} = \psi'(\mathrm{z}) \tag{4}$$

$$\sum_{n=0} a_n n z^n = z\psi'(\mathrm{z}) \tag{5}$$

Further, if the sequence a_n is the probability distribution of the random variable A, or the probability of the occurrence of n, then $0\le a_n\le 1$, and we know that $\psi(1)=1$. Therefore, we can name $\psi(z)$ the probability generation function. At this point, you can use the above-mentioned generation function to solve for any k order moment of the random variable A.

$$\left\langle n^k\right\rangle=\sum_{n=0} a_n n^k=\left[\left(z\frac{d}{dz}\right)^k\psi(\mathrm{z})\right]_{z=1} \tag{6}$$

At this time, if we let $n=1$, we can find the average value of random variables A.

Furthermore, if there are two independent random variables A and B, and their joint distribution is defined as $A+B$, then the probability distribution function of the joint distribution can be characterized by:

$$\begin{aligned}\psi_{A+B}(\mathrm{z})&=\sum_{n=0}\sum_{i=0}\binom{n}{i}a_i b_j z^n\\&=\sum_{i,j} a_i b_j z^{i+j}\\&=\sum_i a_i z^i\sum_j b_j z^j\\&=\psi_A(z)\psi_B(z)\end{aligned} \tag{7}$$

If it is generalized to a more general case, for the joint probability distribution of m independent random variables X, the density generation function can be characterized as $\left[\psi_X(\mathrm{z})\right]^m$. For any

random variable Y among the m independent random variables, the probability density generation function can be characterized as:

$$\psi_Y(z) = \sum_m y_m \left[\psi_X(z)\right]^m = \psi_y\left(\psi_X(z)\right) \tag{8}$$

In the formula, $\psi_y\left(x\right) = \sum_m y_m x^m$. So far, the generating function formulas (7) and (8) derived from the z-transform will be applied to the complex network cascading failure in the next section.

Generating functions for a single complex network

Assume that there is a single complex network with N nodes, and all nodes in the network are randomly connected, so that its degree distribution is $p(k)$. The degree distribution $p(k)$ refers to the probability k that the degree of nodes in the network is $p(k)$, where $1 \le k < \infty$. The generation function that defines the degree distribution is: $G(x) = \sum_{k=0}^{\infty} p(k)x^k$.

Among them, x generally represents a probability value, the value of which can be expressed as the connection status of edges in a complex network. Therefore, it can be known that the average degree in the network can be expressed by the following formula:

$$\langle k \rangle = \sum_{k=0}^{\infty} kP(k) = \left.\frac{\partial G}{\partial x}\right|_{x \to 1} = G^{'}(1) \tag{9}$$

For an infinitely large network $N \to \infty$, the growth process can follow the following rules: starting from a network with N_0 nodes, adding a new node each time, and connecting to each node with probability $f = kP(k)/\langle k \rangle$. k is the degree of each node that already exists, $P(k)$ is the degree distribution probability of each node that already exists, and $\langle k \rangle$ is the average degree of the entire network. Therefore, the generation function of the network addition process can be expressed by the following formula::

$$H(k) = \frac{\sum_{k=0}^{\infty} P(k)kx^{k-1}}{\langle k \rangle} = \frac{G^{'}(x)}{G^{'}(1)} \tag{10}$$

When the number of nodes in the network is randomly removed, its generating function expression must be recalculated. When the cascade failure is over, the size of the remaining giant components in the network can be characterized by the following formula:

$$P_\infty = (1-p)g(1-p) \tag{11}$$

Let the physical meaning of $r = 1-p$ and r be the number of nodes remaining in the network when the proportion p of nodes in the network is removed for the first time, then the above formula becomes $P_\infty = rg(r)$.

among them,

$$g(r) = 1 - \mathrm{G}[rf(r) + 1 - r] \tag{12}$$

And $f(r)$ satisfies

$$f(r) = H[rf(r) + 1 - r] \tag{13}$$

When the removal ratio p is within a certain range, the above formula can find the effective solution $f(r)<1$; however, as the removal ratio p increases, the result will tend to the invalid solution $f(r)=1$.

Therefore, with the increase of the removal ratio p, when the removal ratio $p=p_c$ is the key node, the size of the remaining giant component P_∞ in the network will eventually approach 0 in the form of a second-order phase transition. Therefore, we can find

$$p_c = 1 - r_c = 1 - \frac{1}{H'(1)} \tag{14}$$

Due to the different network models of complex networks, the degree distribution forms are different, and the process of cascading failure will also change greatly. Commonly accepted network models currently include ER random networks, small-world network models, and scale-free network models. For example, for ER random networks, the degree distribution is Poisson distribution, that is,

$$p_k = \frac{e^{-\langle k\rangle}\langle k\rangle^k}{k!} \tag{15}$$

According to the above derivation process, the generating function is:

$$G(x) = H(x) = \exp\left[\langle k\rangle(x-1)\right] \tag{16}$$

$$g(r) = 1 - f(r) \tag{17}$$

$$f(r) = \exp\left\{p\langle k\rangle\left[f(r)-1\right]\right\} \tag{18}$$

Therefore, we can get the relationship of the remaining giant components after the cascade failure of the ER random network:

$$P_\infty = r\left[1-\exp(-\langle k\rangle P_\infty)\right] \tag{19}$$

Obviously, the equation is constant over a point $(r, P_\infty) = (0,0)$. At this time, the equation does not have a solution for the closed form of P_∞, but the phase transition point where the cascade failure occurs can be accurately obtained:

$$p_c = 1 - r_c = 1 - \frac{1}{\langle k\rangle} \tag{20}$$

In order to solve the giant component P_∞ in formula (19), we discuss it further, assuming the function

$$w(P_\infty) = P_\infty - r\left[1-\exp(-\langle k\rangle P_\infty)\right] \tag{21}$$

There is $P_\infty \to 0$ near the phase transition point p_c of the function, and Taylor expansion of the above formula to $P_\infty = \varepsilon << 1$ gives

$$w(P_\infty) = w(0) + w'(0)\varepsilon + \frac{1}{2}w''(0)\varepsilon^2 + \frac{1}{3}w'''(0)\varepsilon^3 + O(\varepsilon^4) \tag{22}$$

among them , $w(0)=0$, $w'(0) = 1 - p\langle k\rangle$, $w''(0) = p\langle k\rangle^2$, $w'''(0) = -p\langle k\rangle^3$。

Obviously, for the second-order phase transition, it satisfies $w^{'}(0)=0, w^{''}(0)>0$. Obviously, the phase transition point formula (19) can be obtained. If we consider the higher-order term $O(\varepsilon^4)$ and let equation $w(P_\infty)=0$, we get

$$\varepsilon = \frac{-2w^{'}(0)}{w^{''}(0)} \tag{23}$$

From this, $P_\infty \propto (r-r_c)^\beta$ can be obtained. Among them, β is the critical index, and for an infinite network, its critical index $\beta=1$.

THEORETICAL STUDY OF ROBUSTNESS OF INTERDEPENDENT NETWORKS BASED ON GENERATIVE FUNCTIONS

Assume that an existing interdependent network consists of two sub-networks S and C. Each sub-network can be represented mathematically as a graph, $G_s=(V_s,E_s)$ and $G_c=(V_c,E_c)$, respectively. Where V_s and V_c respectively represent the point set of the two sub-networks, and E_s and E_c represent the edge sets connected internally of the two sub-networks, respectively. Define the coupling interdependent edge set $E_{sc}=\left\{(u,v)\middle|u\in V_{s,}v\in V_c\right\}$ of the sub-network. Therefore, we call E_s and E_c the internal connecting edges of the network, and E_{sc} is the external interdependent edge of the network. The external interdependent edges between the network S and the network C adopt a random coupling method, and it is guaranteed that there is at most one external interdependent edge connected to the interdependent network of any node in the network S or the network C.

The number of nodes in network S is N_S, the number of nodes in network C is N_C, and their degree distributions are $P_S(k)$ and $P_C(k)$, respectively. And the node with the proportion q_S in the network S and the node with the proportion q_C in the network C have an interdependent relationship. Next, this article will theoretically analyze the cascading failure of interdependent networks step by step. To trigger a cascading failure, we first remove a certain percentage p of nodes in network S.

Step 1 of the cascading failure process: The number of remaining nodes in the sub-network is $\psi_1^{'}=N_S(1-p)$, but the part of the remaining nodes that does not belong to the huge component must also be removed. Therefore, the number of remaining huge component nodes in the sub-network is $\psi_1=\psi_1^{'}g_S(\psi_1^{'})$. Among them, $g_S(\psi_1^{'})$ is consistent with the definition of formula (11) and formula (12) of the single network cascading failure part.

Step 2 of the cascading failure process: There are a number of nodes q_C in network C and interdependent connection edges in network S. Therefore, with the failure of the nodes in network S, there will be a number of nodes in network C that fail. At this time, the number of remaining nodes in the network C is $\phi_1^{'}=1-q_C(1-\psi_1^{'}g_S(\psi_1^{'}))$. Furthermore, the nodes in the network C may also be divided into several parts, where the node size of the huge component is $\phi_1=\phi_1^{'}g_B(\phi_1^{'})$.

Step of cascading failure process: The failure of nodes in network C will cause the failure of its interdependent nodes in network S. By analogy, the interdependent network will continue to cascade failures until it reaches a stable state. Therefore, its general formula is:

$$\begin{cases} \psi_1^{'} = N_S(1-p) \\ \phi_1^{'} = 1 - q_C(1-\psi_1^{'} g_S(\psi_1^{'})) \\ \ldots\ldots \\ \psi_t^{'} = \psi_1^{'}\left[1 - q_s(1-\mathrm{g}_c(\phi_{t-1}^{'}))\right] \\ \phi_1^{'} = 1 - q_C(1-\psi_1^{'} g_S(\psi_{t-1}^{'})) \end{cases} \tag{24}$$

In addition, the judgment condition for the failure of the interdependent network cascade to reach the end point is:

$$\begin{cases} \psi_{t+1}^{'} = \psi_t^{'} \\ \phi_{t+1}^{'} = \phi_t^{'} \end{cases} \tag{25}$$

If t is large enough and satisfies $\psi_{t+1}^{'} = \psi_t^{'}$ and $\phi_{t+1}^{'} = \phi_t^{'}$, let $x = \psi_t^{'}$, $y = \phi_t^{'}$ then the general formula of formula (24) can be used to obtain the final state of the dependent network after the cascade failure has ended:

$$\begin{cases} x = N_S(1-p)\left[1 - q_s(1-\mathrm{g}_c(y))\right] \\ y = 1 - q_C(1-g_S(\psi_{t-1}^{'})N_S(1-p)) \end{cases} \tag{26}$$

If the above equation is drawn in the (x, y) plane as two curves, for q_S and q_C that are large enough, the two curves will have two intersection points $(1 < x_0, 0 < y_0)$, $(x_0 < x_1 < 1, y_0 < y_1 < 1)$. Studies have shown that the second point of intersection (x_1, y_1) of the two curves has certain physical significance. As the removal ratio p increases, the distance between these two intersections will get closer and closer, but the intersections are both within the specified range $(0 < x < 1, 0 < y < 1)$. And at the key point $p = p_c$, the two intersection points of the two curves will meet, namely $0 < x_0 = x_1 = x_c < 1$, $0 < y_0 = y_1 = y_c < 1$. For sufficiently large q_S and q_C, $P_{\infty,S}$ and $P_{\infty,\mathrm{C}}$ can be functions with respect to the removal ratio p, and present a first-order phase change. As q_C decreases, the function of $P_{\infty,S}$ with respect to the removal ratio p appears as a second-order phase transition.

CONCLUSIONS AND RECOMMENDATIONS

In order to facilitate the analysis of the robustness of complex networks, a common method in complex network calculations—generating functions is introduced first. A single complex network generating function is analyzed. In combination with the interdependent network modeling mechanism, a interdependent network based on generating functions is performed.

Research shows that with the increase of the removal ratio p, for a sufficiently large proportion of dependent nodes, the remaining giant component P_∞ is a function of the removal ratio p, and presents a first-order phase transition;

As the dependence node ratio q decreases, The remaining giant component P_∞ appears as a second-order phase change with respect to the removal ratio p.

REFERENCES

Bagheri, E., Ghorbani, A. A.. (2010). Uml-ci: a reference model for profiling critical infrastructure systems. Information Systems Frontiers, 12(2), 115-139.

Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E., Havlin, S.. (2010). Catastrophic cascade of failures in interdependent networks. Nature, 463(7291), p.1025-1028.

Buldyrev, S. V., Shere, N. W., Cwilich, G. A.. (2011). Interdependent networks with identical degrees of mutually dependent nodes. Physical Review E, 83(1 Pt 2), 016112.

Gao, J., Buldyrev, S. V., Havlin, S., Stanley, H.E.. (2011). Robustness of a network of networks. Physical review letters, 107(19), p.195701.1-195701.5.

Johansson, J., Hassel, H.. (2010). An approach for modelling interdependent infrastructures in the context of vulnerability analysis. Reliability Engineering and System Safety, 95(12), 1335-1344.

Reed, D. A., Kapur, K. C., Christie, R. D. (2009). Methodology for assessing the resilience of networked infrastructure. IEEE Systems Journal, 3(2), 174-180.

Robert, B., Morabito, L. (2008). The operational tools for managing physical interdependencies among critical infrastructures. International Journal of Critical Infrastructures, 4(4), 353-367.