

## Research Article

# **Research on the Statistical Properties and Stability of Complex Interindustrial Networks**

## Xinyu Cheng

University of International Business and Economics, Beijing, China

Correspondence should be addressed to Xinyu Cheng; 201900173001@uibe.edu.cn

Received 30 October 2023; Revised 24 January 2024; Accepted 17 February 2024; Published 26 February 2024

Academic Editor: Hiroki Sayama

Copyright © 2024 Xinyu Cheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study consolidates input-output data from 42 sectors across 31 provinces and regions in China into a unified dataset for 42 industrial sectors within eight major economic zones. Leveraging the maximum entropy method, we identify significant interindustrial relationships, subsequently forming a directed, weighted, complex network of these ties. Building upon this intricate network, we analyze its foundational statistical attributes. The stability of the network's structure is further assessed through simulations of varied network attacks. Our findings demonstrate that the maximum entropy method is adept at extracting notable relationships between industrial sectors, facilitating the creation of a cogent complex interindustrial network. Although this established network exhibits high stability, it calls for targeted policy interventions and risk management, especially for industries with pronounced degree centrality and betweenness centrality. These pivotal industry nodes play a decisive role in the overall stability of the network. The insights derived from our examination of complex interindustrial networks illuminate the structure and function of industrial networks, bearing profound implications for policymaking and propelling sustainable, balanced economic progress.

## 1. Introduction

With the advent of globalization and the digital age, the interplay between industries has evolved into unprecedented complexity. This intricate nature challenges conventional analytical methods, rendering them inadequate for in-depth analyses of complex industrial networks within the economic domain. By the close of the 20th century, the rise of complex network theory not only garnered extensive academic attention but also paved the way for a series of groundbreaking research endeavors. Seminal theories such as Watts and Strogatz's small-world theory [1], as well as the scale-free network theory posited by Barabási et al. [2], continue to exert profound influence in contemporary discourse [3, 4].

The growing body of literature on network disruption and resilience, such as Iyer et al.'s work on attack robustness and centrality [5], Casali and Heinimann's study on the road network robustness [6], and Ficara et al.'s investigation into strategies for disrupting criminal networks [7, 8], further enriches our understanding of the complex interplays and resilience mechanisms inherent within various network systems.

Industrial interactions have transcended mere singular linkages, gradually morphing into a sophisticated network system. This structure elucidates the realities of industries under the twin forces of globalization and technological advancement, offering a renewed analytical lens for both theoretical exploration and practical application. Notably, within the realm of economics, the introduction of complex network theory has rejuvenated input-output analysis. Leontief's input-output table stands as a pivotal tool to unveil resource flows and dependencies between varying industries [9]. Methods grounded in this paradigm have been employed by scholars like Serrano et al. [10], offering insights into the collaborative actions, competitive relations, and stability inherent within industries, particularly in their analyses of world trade networks and China's energy flow networks.

Yet, the evolution of network science marches on. Many scholars are shifting their focus toward the dynamic behaviors within networks, especially concerning stability issues. This concern spans from natural ecosystems to manmade realms such as transportation and supply chain networks, addressing challenges of holistic stability and potential collapses triggered by localized node failures or attacks. As a result, network resilience and defense against attacks have gradually ascended to the forefront of complex network research. The stability and resilience of vital infrastructures, like finance and energy, have garnered widespread attention [11, 12]. Recent studies, through simulations and experiments, have unveiled the latent ramifications of network attacks while probing novel strategies to bolster network stability [13–15].

In summary, despite significant advancements in the study of complex industrial networks, numerous areas remain untouched and present challenges. It is especially salient to highlight that the stability analysis of input-output correlation within these networks has been somewhat overshadowed. Predominant studies tend to zoom in on specific industries or regions, often neglecting a holistic global viewpoint and missing out on interdisciplinary syntheses. Given the escalating intricacy and diversification of industrial interactions, there is a pressing need for more in-depth exploration and research in this domain.

The primary objective of this study is to delve deeply into the interactions between industries and their network structures using complex network theory, aiming to elucidate the network's stability and resistance to attacks. The research unfolds in distinct phases. Initially, leveraging the input-output table data, we employ the maximum entropy method to shape a complex network. Subsequently, we probe into the foundational statistical attributes—degree, node weight, clustering coefficient, shortest path, and network efficiency—of the crafted interindustrial complex network to discern the network's architecture. In the concluding phase, we assess the network's four centralities and gauge the stability and resilience of the overarching network through simulations of various attack paradigms.

### 2. Literature Review

2.1. Network Construction and Analysis Methods. The construction and analysis of networks from empirical data have become pivotal in understanding complex systems across various disciplines. The work of Donner et al. [16] highlights the quantitative assessment of structural properties in systems composed of interacting entities. This is particularly relevant for our study as it underscores the importance of understanding dynamic higher-order structures in complex networks, which can be applied to analyzing interindustrial relationships. Christensen et al. [17] provide insights into the universal nature of complex networks. Their research into the topological similarities among different systems can inform our understanding of the interconnected nature of industrial sectors, further enriching our network analysis. Furthermore, Cheng and Scherpen [18] discuss the challenges and solutions to dealing with high-dimensional dynamics and complex interconnections in network systems. This perspective is crucial for simplifying and effectively

analyzing the intricate network of industrial sectors in our study. Emmert-Streib et al. [19] emphasize the interconnectedness of economic and financial entities. Their approach to network science in economics and finance offers valuable parallels to our method of extracting network structures from real-world data, particularly in the context of economic networks. Lastly, Polishchuk's [20] analysis introduces concepts like flow adjacency matrices and dynamic characteristics of system elements. These concepts are instrumental in understanding the behavior of complex network systems and can be applied to our study to explore the flow core and dynamic interactions within the industrial network. Similar to our method, these studies often extract network structures from real-world data, underscoring the importance of leveraging empirical data to form directed, weighted complex networks. The process of forming these networks from input-output tables is akin to the approach taken by economic and ecological network analyses, which extract relational data to understand systemic dependencies and dynamics. The insights from these papers not only reinforce our methodology but also provide a broader context for understanding the complexities and interdependencies in industrial networks.

2.2. Simulation and Attack Modeling in Network Stability Analysis. Simulation methods, including those for targeted network attacks, are extensively utilized to evaluate network stability and robustness. Researchers simulate the removal of key nodes to determine a network's vulnerability to specific disruptions. For example, Iyer et al. analyzed complex network centrality to assess network attack robustness, bridging theoretical models with practical applications. Furthering this field, Dshalalow and White [21] employ stochastic processes to model network attacks, enabling predictions about the timing and scale of network failures. This method introduces a probabilistic aspect to network robustness assessment, shedding light on the unpredictability and impact of network attacks. Fabris and Zelazo [22] investigate the resilience of multi-agent consensus networks against attacks that manipulate edge weights. Their research broadens the scope of network robustness understanding by examining the effects of such attacks on network convergence performance, underscoring the need for structured defense strategies. Sarraute et al. [23] present a prototype for simulating large-scale network attacks, focusing on realism from the attacker's perspective. This study highlights the importance of comprehensive simulation environments that accurately reflect the complexities of real-world networks. P. Y. Chen and S. M. Chen [24] explore the effectiveness of sequential defense strategies against both random and intentional attacks, emphasizing the importance of adaptive defense mechanisms for maintaining network integrity. Bel et al. [25] introduce a cosimulation framework for generating and monitoring network attacks, especially in power grids with integrated distributed energy resources. This study points out the changing nature of network attack surfaces and the necessity for advanced simulation tools for evaluating and improving network resilience.

Collectively, these studies enhance the understanding of network stability and robustness, highlighting the critical role of advanced simulation methods in assessing network vulnerabilities and devising defense strategies against various potential attacks.

2.3. Assessment of Network Stability. Evaluating the stability of network structures is crucial for understanding the resilience of complex systems to various perturbations. Our study extends the literature by simulating varied network attacks to assess the stability of the interindustrial network. This methodology aligns with the approaches seen in Casali and Heinimann, who evaluated the robustness of the Zurich road network under different disruption processes. By employing simulations of targeted attacks, we can identify critical nodes within the network, akin to the analysis of network centrality in determining the pivotal roles certain nodes play in maintaining network integrity and stability. Building upon this foundation, Platig et al. [26] further enhance our understanding of network stability. Their exploration of the robustness of network measures, including centrality, in the presence of link inaccuracies is particularly relevant. It underscores the importance of considering the reliability of network connections in assessing overall network stability. Ufimtsev et al. [27] reveal that the impact of structural noise on centrality ranks is examined. This study complements our approach by highlighting how minor structural changes can significantly influence the stability and centrality of nodes, thereby affecting the network's resilience to disruptions. Saxena et al. [28] argue for the importance of stability in centrality measures under information loss or noise. This perspective is crucial for our study as it emphasizes the need for robust centrality measures that can withstand variations in network data, ensuring accurate identification of critical nodes. Oldham et al. [29] provide insights into the roles of different nodes through various centrality analyses. Understanding the consistency and uniqueness of these measures across different network types aids in accurately assessing the roles of nodes in maintaining network stability. Gupta et al. [30] discuss the significance of centrality measures in large networks with community structures. This research is pertinent to our study as it offers methods to identify influential nodes in complex networks, which is key to understanding and enhancing the resilience of the interindustrial network. Together, these studies enrich our methodology by providing a comprehensive view of how centrality measures and network robustness assessments can be effectively utilized to understand and improve the stability of complex networks, such as the interindustrial network in our study.

## 3. Establishment of the Complex Input-Output Network

3.1. Data Sources and Processing. This study utilizes data derived from the work of Zheng et al. [31], specifically the CEADS 2017 interregional input-output table (The China City-level MRIO Table data used to support the findings of

this study may be released upon application to the Carbon Emission Accounts & Datasets, who can be contacted at shanyuli@outlook.com or guandabo@hotmail.com or zhuliu@tsinghua.edu.cn.) for 31 provinces, autonomous regions, and municipalities in mainland China covering 42 industries. Constructing a complex network based on this table would generate 1,302 nodes, considering the 31 areas and 42 industries, resulting in a substantial number of node interconnections. For research clarity, the 31 regions were consolidated based on recommendations from the "Strategies and Policies for Coordinated Regional Development" report published by the Development Research Center of the State Council. This report delineated mainland China into four primary sectors during the "Eleventh Five-Year Plan" period: eastern, central, western, and northeastern areas. These were further subdivided into eight comprehensive economic regions, with divisions primarily informed by each region's economic traits and geographical position. Table 1 provides a detailed breakdown of these divisions. Such categorization offers a more macroscopic perspective on China's economic framework and interindustrial dynamics.

In the course of data processing, input-output data from the 42 industries across each province, region, and city were aggregated. This aggregation was undertaken to compute the consolidated input-output figures for the 42 industries within the eight principal economic regions, yielding a comprehensive input-output table. For analytical clarity, the eight key economic regions are denoted by letters A through H, while the 42 industries are numerically represented from 1 to 42. Refer to Tables 1 and 2 for a detailed breakdown.

3.2. Modeling Method. In this study, the construction of the complex network pivots on the 0-1 adjacency matrix. Within this matrix, an element denoted as "1" signifies that the nodes corresponding to that particular row and column are interconnected, whereas an element marked "0" suggests the absence of such a connection. To achieve this configuration, the matrix undergoes a binarization process to manifest as a 0-1 adjacency matrix. This study employs the maximum entropy method for binarization, a statistical approach grounded in information theory tailored to estimate probability distributions contingent on certain predefined constraints. This approach is in line with Loaiza-Ganem et al. [32], who highlight the adaptability and efficacy of maximum entropy in statistical models, especially for optimizing network density. In addition, Metzig and Colijn [33] demonstrate the use of Gibbs-Shannon entropy in network size and degree distributions, providing a theoretical foundation for predictive analysis, and Zenil et al. [34] offer insights into employing maximum entropy for prior probability distributions, essential for understanding the probabilistic aspects of industrial networks.

In information theory, entropy quantifies the uncertainty inherent in random variables and simultaneously reflects the informational richness of the data. The information entropy of an event is, essentially, the expected value of the logarithm of the event's probability. When entropy peaks, it signals the extraction of optimal

Code	Economic region
A	Northeast comprehensive economic region (Liaoning, Jilin, Heilongjiang)
В	Northern coastal comprehensive economic region (Beijing, Tianjin, Hebei, Shandong)
C D	Eastern coastal comprehensive economic region (Shanghai, Jiangsu, Zhejiang) Southern coastal economic region (Fujian, Guangdong, Hainan)
Е	Middle yellow river comprehensive economic region (Shaanxi, Shanxi, Henan, Inner Mongolia)
F	Comprehensive economic region in the middle reaches of the Yangtze River (Hubei, Hunan, Jiangxi, Anhui)
G	Southwest comprehensive economic region (Yunnan, Guizhou, Sichuan, Chongqing, Guangxi)
Н	Northwest comprehensive economic region (Gansu, Qinghai, Ningxia, Tibet, Xinjiang)

TABLE 1: Division and codes of the eight major economic regions.

TABLE 2: 42 industries and corresponding codes.

Code	Industry
1	Agriculture, forestry, animal husbandry, and fishery products and services
2	Coal mining and processing products
3	Oil and gas extraction products
4	Metal ore mining and processing products
5	Nonmetallic minerals and other mineral processing products
6	Food and Tobacco
7	Textiles
8	Textiles, clothing, shoes, hats, leather, down, and their products
9	Wood processed products and furniture
10	Papermaking, printing and cultural, educational, and sporting goods
11	Petroleum, coking products, and nuclear fuel processed products
12	Chemical products
13	Nonmetallic mineral products
14	Metal smelting and rolling products
15	Metal products
16	General equipment
17	Special equipment
18	Transportation equipment
19	Electrical machinery and equipment
20	Communication equipment, computers, and other electronic equipment
21	Instrumentation
22	Other manufactured products
23	Scrap
24	Metal products, machinery and equipment repair services
25	Production and supply of electricity and heat
26	Gas production and supply
27	Water production and supply
28	Architecture
29	Wholesale and retail
30	Transportation, warehousing, and postal services
31	Accommodation and meals
32	Information transmission, software and information technology services
33	Finance
34	Real estate
35	Leasing and business services
36	Scientific research and technical services
37	Water conservancy, environment and public facilities management
38	Residential services, repairs, and other services
39	Education
40	Health and social work
41	Culture, Sports, and Entertainment
42	Public administration, social security, and social organizations

information from the event. The quantitative representation of an event's information is expressed as follows:

$$I = -\log p(x), \tag{1}$$

where p(x) represents the probability of the event's occurrence. Given a set with *n* events, the information entropy for this event set is defined as follows:

$$H = -\sum_{i=1}^{n} p(x_i) \log p(x_i),$$
 (2)

where  $p(x_i)$  denotes the probability of the *i*<sup>th</sup> event occurring. The crux of the maximum entropy method lies in its principle: Among all probability distributions that align with given constraints, the distribution with the maximum entropy is the one capturing the optimal amount of information.

In the construction of a complex network, the maximum entropy method plays a pivotal role in determining the network's adjacency matrix. Specifically, an appropriate threshold is chosen to classify the connections between nodes into two categories: present or absent. When the sum of the average entropies for these two categories is maximized, we attain the greatest amount of information. This, in turn, helps ascertain the optimal threshold, which dictates whether a connection between nodes exists. The primary steps for the maximum entropy algorithm in identifying the best threshold are as follows:

- (1) Initially, acquire the original matrix indicating the correlation strength between nodes and derive the probability distribution of the correlation strengths. If the total count of distinct correlation strengths is N, the probability of the  $i^{\text{th}}$  correlation strength appearing is represented as  $p_i$ .
- (2) Set an initial threshold Th, which equates to the  $j^{\text{th}}$  correlation strength,  $1 \le j < N$ . Partition the correlation strength matrix into two categories: those less than or equal to Th, and those greater than Th.
- (3) Calculate the average relative entropy for both categories:

$$E_{1} = -\sum_{i=1}^{j} \left(\frac{p_{i}}{p_{j}}\right) \ln\left(\frac{p_{i}}{p_{j}}\right), p_{j} = \sum_{i=1}^{j} p_{i},$$

$$E_{2} = -\sum_{i=j+1}^{N} \left(\frac{p_{i}}{(1-p_{j})}\right) \ln\left(\frac{p_{i}}{(1-p_{j})}\right).$$
(3)

(4) When the value of  $E_1 + E_2$  peaks, the corresponding Th is identified as the optimal threshold. By binarizing the correlation strength matrix using this optimal threshold, we obtain the adjacency matrix.

3.3. Building Complex Networks. To construct an adjacency matrix based on the input-output tables from 42 sectors (industries) within the eight major economic zones, it is first essential to establish a matrix denoting the linkage intensity between these industries. In this study, the directly

calculated consumption coefficient matrix is employed as this linkage intensity matrix. This direct consumption coefficient matrix illustrates the coefficient of direct inputs from various sectors into a specific output, shedding light on the direct dependencies between industries. Let's denote this coefficient matrix as A, where  $A_{ii}$  represents the direct input required from industry *i* to produce 1 unit of the product from industry *j*. Applying the maximum entropy method described in the previous section, we determined the optimal threshold for the direct consumption coefficient matrix. This threshold serves to extract significant interindustry relationships from the original direct consumption coefficient matrix. Specifically, elements in matrix A that are less than or equal to this threshold are assigned a value of 0, while those exceeding the threshold are designated a value of 1. Consequently, we obtain the adjacency matrix M, which delineates the salient relationships between industries. This procedure is referred to as the binarization process.

Based on the adjacency matrix M, we construct a complex network. If  $M_{ii} = 1$ , an oriented edge is drawn from industry *i* to industry *j*. Conversely, if  $M_{ij} = 0$ , there is no edge from industry i to industry j. Given the inherent asymmetry in inputs (or consumption) between two industries, typically  $M_{ij} \neq M_{ji}$ . Moreover, some industries have inputs (or consumption) relating to themselves,  $M_{ii} \neq 0$ , indicating the presence of loops within the network. Keeping in mind the practical significance of the input-output association in complex networks, it is essential to account for the differences in the magnitude of inputs (or consumption) between industries. As such, edges within the network carry weights, denoted as  $W_{ij}$ , with the set of all weights represented as W. Here,  $W_{ij}$  corresponds to the direct con-sumption coefficient  $A_{ij}$  of industry *i* to industry *j*. The resulting complex network is illustrated in Figure 1. In the figure, blue dots represent network nodes, red labels denote node codes, black lines signify edges, and arrows indicate the direction of the edges. Nodes within the network symbolize industry sectors, with the eight major economic zones contributing 336 nodes in total. Within each economic zone, the 42 industry nodes are distributed in a rectangular uniform layout, while the grouping of these nodes reflects the rough geographical placement of the respective economic zones. Edges in the network represent relationships between industries, amounting to 30,864 directed and weighted edges in total.

In summary, starting from the direct consumption coefficient matrix, and using the maximum entropy method to go through a series of calculations and processing, a directed weighted complex network with self-loop is finally established. This network provides the basis for subsequent network statistical properties and stability analysis.

## 4. Analysis of Complex Network Statistical Properties

4.1. Degrees and Node Weight. Within the realm of graph theory, complex networks offer a robust representation of systems with interconnected entities, known as nodes, connected by links, termed edges. This perspective is

FIGURE 1: The complex network of 42 industries across the eight major economic regions.

foundational in understanding the organization and dynamics of diverse complex systems, from biological networks to social structures and technological infrastructures. The application and significance of graph theory in analyzing such systems are extensively supported by a body of research. Christensen and Albert [17] underscore the universal applicability of graph concepts across various fields, highlighting the shared topological features of different complex systems. Jalving et al. [35] propose graph-based modeling abstractions that articulate the dependencies and interactions within complex systems. Their work is particularly pertinent to studies focused on interconnected entities, such as industrial networks, demonstrating the utility of graphbased models in capturing the intricate relationships that define complex systems. Torres et al. [36] delve into the intricacies of representing complex systems, emphasizing the necessity for effective representation strategies across various domains. Their discussion on the why, how, and when of complex system representations sheds light on the methodological challenges and considerations in employing graph theory to model complex interactions, providing a critical lens through which to view our research endeavors.

The number of connections or edges linked to a node, or its degree, provides insights into its role and significance within the network. In complex networks, the degree of a node is indicative of its connectivity. In directed networks, the concept further bifurcates into in-degree (incoming connections) and out-degree (outgoing connections). From an economic perspective, a higher degree, whether it be indegree or out-degree, generally signifies an industry's influential role and interconnectedness in the marketplace. A node's degree in economic networks can be seen as a measure of its interdependence with other industries or sectors.

In economic networks, the in-degree can be thought of as the diversity of resources or inputs an industry requires, whereas the out-degree can indicate the variety of outputs or services it provides to other sectors. An industry with a high in-degree, for instance, might be crucial for multiple sectors due to its essential products or services. Conversely, a high out-degree may suggest the industry relies on diverse inputs from various sectors, denoting its intricate integration into the overall economic fabric.

For the established directed weighted network, we computed the degree, in-degree, and out-degree and derived a degree distribution histogram as shown in Figure 2. Statistical analysis reveals that the degree distribution does not follow a power-law distribution, with 85% of industry sectors having a degree ranging between 70 and 220. This indicates a tight topological linkage among industries. The industry with the lowest degree is Petroleum and Natural Gas Extraction Products (East coast) with an in-degree of 35 and an out-degree of 4, suggesting that the East Coast economic zone's petroleum and natural gas extraction heavily relies on input from other industries. The industries with the highest degrees include Wholesale and Retail (East Coast), Transportation, Warehousing, and Postal Services (East Coast), Nonmetallic Minerals and Other Mining Products (Mid-Yellow River), and Transportation, Warehousing, and Postal Services (North coast) with degrees ranging from 328 to 325. These industries maintain connections with almost all other sectors, denoting their significant role in the economic network.

Figure 2 also reveals that the in-degree distribution roughly follows a bell-shaped curve, indicative of a Poisson distribution, with 91% of industries having an in-degree ranging between 50 and 130. The highest in-degree is 163 for construction (Mid-Yellow River), followed by metal products machinery and equipment repair services (Mid-Yellow River) with an in-degree of 156, and then various industries from the Mid-Yellow River and the Northwest. Notably, construction, public facilities, and the service sector are high-consumption industries, particularly those in the Mid-Yellow River and Northwest economic zones, as they require resources and products from nearly half of the industries.

The out-degree distribution in Figure 2 shows a higher proportion of industries with lower out-degrees. The industry with the highest out-degree of 326 is Transportation, Warehousing, and Postal Services (Mid-Yellow River), followed closely by Wholesale and Retail (East Coast) with 325. It is evident that transportation, warehousing, and postal services, as well as wholesale and retail, play a central role, as nearly all industries rely on consuming their products and resources, underscoring their critical position in the national economy.

Analyzing our directed weighted network, we found nuances in the connectivity and centrality of various sectors, reflected in their degree distributions. A significant observation from Figure 2 is that the degree distribution does not align with the typical power-law seen in many real-world networks, implying that our economic network deviates from scale-free characteristics. Instead, the observed Poisson distribution suggests a more homogenous network structure where most nodes have a degree close to the average.

Some sectors, as evident from our analysis, act as central hubs. Their high connectivity, both in terms of inputs and outputs, indicates their pivotal role in the economic





FIGURE 2: Degree distribution histogram.

network. The East Coast's Petroleum and Natural Gas Extraction Products sector, for example, showcases how regional specializations and dependencies can emerge, underlining the nuances of economic geography.

Further, the central roles of sectors like wholesale and retail, and transportation, warehousing, and postal services, indicate their foundational importance in facilitating and sustaining the operations of other industries. In the context of network theory, these can be seen as hub nodes, exerting disproportionate influence on the network's overall connectivity and flow.

Given the directed, weighted nature of our established complex network, our initial degree distribution analysis did not account for the significance of edge weights. Consequently, we extended our investigation to ascertain the node strength, in-strength, and out-strength for each industry node within the network.

In weighted network analyses, the node strength represents the sum of the weights of all edges connected to a specific node. In the context of a directed weighted network, the in-strength embodies the cumulative weight of all incoming edges (edges directed towards the node), while the out-strength conveys the sum of the weights of all outgoing edges (edges originating from the node).

From an economic perspective, these network metrics take on heightened significance. The in-strength can be interpreted as the total volume of inputs an industry receives from other sectors, indicating its dependency or reliance on external factors for operation. Conversely, the out-strength offers insights into the volume of inputs or resources an industry provides to others, reflecting its contribution and potential influence over other sectors in the economic landscape. By examining these weighted network measures, we can derive a more nuanced understanding of each industry's role and importance within the broader economic network, leading to deeper insights into intersectoral dependencies and influences.

The direct consumption coefficient between industries serves as an edge weight in this analysis. The higher its value, the more significant the input or consumption volume between two industries, indicating a closer interrelation. Therefore, this edge weight operates as a similarity weight. Table 3 presents industries with the highest and lowest node weight, in-weight, and out-weight.

A striking observation is that the industries ranking in the top four for node weight also dominate the top four for out-weight. These industries, in sequence, are as follows: chemical products (Eastern Coastal), chemical products (Northern Coastal), leasing and business services (Eastern Coastal), and chemical products (Middle Yangtze River). Predominantly, the node weight of these sectors is dictated by their out-weight. Chemical products, as an export-driven sector, provide a significant volume of resources or products to other industries.

Regionally, the Jiangsu-Zhejiang-Shanghai area is underscored by the robust standing of its wholesale and retail sector, which ranks fifth in out-weight, reinforcing the prosperous nature of this sector and its substantial contributions to other industries.

On the flip side, the industries ranking high in in-weight include textiles (Northern Coastal), other manufacturing products (Northwest), other manufacturing products (Northeast), textile, clothing, footwear, leather, down and its products (Middle Yellow River), and textile, clothing, footwear, leather, down and its products (Northern Coastal).

Node weight	Node	In-weight	Node	Out-weight	Node
3.67007402	C12	0.86712420	B7	3.03943861	C12
3.38395668	B12	0.84029975	H22	2.62444130	B12
2.89679630	C35	0.84007252	A22	2.24181543	C35
2.89500907	F12	0.83725058	E8	2.15906770	F12
2.79223454	B14	0.83029668	B8	2.12924387	C29
0.25412811	B23	0.14281446	B23	0.01646612	B40
0.23549762	H23	0.13928177	C4	0.01474259	A37
0.18969335	A23	0.13047039	C23	0.01018285	G24
0.15067719	C3	0.13043612	C3	0.00602547	H40
0.14479969	C4	0.12417176	D23	0.00551791	C4

This suggests some regional dynamics at play: other manufacturing product industries in the Northeast and Northwest regions appear to be underdeveloped, demanding significant inputs from other sectors. The textile sector in the Northern Coastal area stands out in its reliance on other industries for resources.

From a regional economic perspective, the spatial distribution and development of industries often reflect a region's historical, geographical, and socio-economic conditions. The thriving chemical and leasing businesses along the coastlines indicate the coastal regions' advantages in trade, port logistics, and market accessibility, leading to economies of scale and agglomeration benefits.

Comparing the industries ranking in the bottom five for both node weight and out-weight, there's a discernible pattern: sectors such as oil and natural gas extraction products and metallic mineral mining products from the eastern coastal region have relatively low in-weights, signaling minimal direct consumption from other industries. Furthermore, waste material sectors in economic zones A, B, C, D, and H are similarly characterized by their low consumption from other sectors. This could hint at either the self-sufficient nature of these industries or perhaps a need for more integration and collaboration for sustainable regional economic development.

4.2. Clustering Coefficient. The clustering characteristic is a critical feature in complex networks, typically used to describe the aggregation tendency among nodes within the network. The clustering coefficient is a commonly used metric to quantify this characteristic. Although its computation varies slightly across different network types (like undirected and directed networks), its core idea revolves around describing the interconnection situation amongst the neighbors of a node. For a given node *i* in an undirected network, the local clustering coefficient is defined as follows:

$$C_i = \frac{2E_i}{K_i(K_i - 1)},\tag{4}$$

where

- (1)  $C_i$  represents the local clustering coefficient of node *i*.
- (2)  $E_i$  denotes the number of links between the neighbors of node.

- Complexity
- (3) K<sub>i</sub> signifies the degree of node *i*, i.e., the number of links connected to node *i*.

The local clustering coefficient describes the ratio between the number of actual connections formed among a node's neighbors and the maximum possible number of such connections.  $C_i = 1$  indicates that all neighbors of node *i* are interconnected, while  $C_i = 0$  signifies that there are no connections among the neighbors of node *i*. The formula essentially captures the ratio of the number of actual links between the neighbors of a node to the maximum possible number of such links. A higher clustering coefficient for a node indicates that its neighbors are more densely interconnected. The global clustering coefficient *C* is the average of the local clustering coefficients for all nodes in the network. Mathematically, it can be expressed as follows:

$$C = \frac{1}{n} \sum_{i=1}^{n} C_i, \tag{5}$$

where n is the total number of nodes in the network. The global clustering coefficient is also commonly referred to as the average clustering coefficient of the network.

In the realm of directed networks, computing the clustering coefficient presents intricate challenges, largely due to the inherent directionality of the edges. A key complication arises from the fact that one triangle formation in an undirected network can manifest in seven possible configurations in a directed network context. To circumvent this complexity, it is standard practice to consider directed edges in the network as bidirectional, undirected edges, thus allowing the application of clustering coefficient calculation methods originally developed for undirected networks. In weighted networks, the calculus extends beyond mere node-to-node connection states to incorporate the strength of these connections as well. As a result, this study adopts a specific definition of the clustering coefficient that is tailored for weighted networks:

$$C_{i} = \frac{1}{K_{i}(K_{i}-1)} \sum_{j,k} \left( \hat{W}_{ij} \hat{W}_{ik} \hat{W}_{jk} \right)^{1/3}, \hat{W}_{ij} = \frac{W_{ij}}{\max_{k,l} W_{kl}}, \quad (6)$$

where  $W_{ij} \in [0, 1]$  is the normalized weight.

In this section, our study lays the groundwork by constructing an undirected weighted network, a subset of the complex network discussed earlier. In a regional economic context, this exercise serves as a crucial step in understanding spatial interdependencies and resource allocations among various industrial nodes. The weight of the edge between two nodes is determined as the sum of the weights of any existing directed edges between them, which is particularly relevant for capturing the flow of goods, services, or information between industries. These weights are calculated using the previously defined direct cost coefficients, which function as similarity weights.

Following this, we apply (4) to compute the clustering coefficient for the unweighted, undirected network and (6) for the weighted clustering coefficient of the undirected weighted network. Table 4 presents the top ten industries ranked by both their clustering and weighted clustering coefficients. In the realm of regional economics, a higher clustering coefficient signifies strong local synergies and interindustrial cooperation. This can often be seen in regional clusters where industries benefit from shared resources, expertise, and markets. Similarly, industries with elevated weighted clustering coefficients are indicative of substantial capital flows, both in terms of investments and consumption, among the industries constituting the vertexassociated triangles. The weighted measures give a nuanced understanding of the economic robustness and the depth of interindustry relationships. According to (5), the network's average clustering coefficient is ascertained to be 0.63737820, while the average weighted clustering coefficient is 0.00252026. These metrics can serve as valuable indicators for policymakers and stakeholders in identifying regional economic strengths and potential areas for fostering industrial collaboration.

4.3. Shortest Path and Network Efficiency. The shortest path in a graph or network refers to the path connecting two nodes that minimizes the sum of lengths or weights along that path. In an unweighted network, the length of the shortest path is typically the number of edges it contains. In a weighted network; however, the length is determined by the sum of the edge weights along the path. When the edge weight is used to represent the distance between nodes, a longer path between two nodes implies a greater distance and thus a more distant or weaker relationship. Consequently, in this context, edge weights should serve as dissimilarity measures.

In this section, the study recalibrates the edge weights in the directed weighted network by taking their reciprocal values, which are then used as dissimilarity measures. This is consistent with using the direct cost coefficient as an inverse measure. Specifically, the greater the direct cost coefficient between two industries, the closer and more tightly-knit their relationship is expected to be. The concept of the shortest path takes on a critical role. It acts as a proxy for transaction costs between industries, and a shorter average path length could indicate a more efficient, agile, and wellintegrated regional economy. Furthermore, understanding the network efficiency and the average shortest path length provides actionable insights for policymakers aiming to optimize resource allocation and improve the economic interconnectivity of industrial clusters. After determining the shortest path and its length between every pair of nodes, the network's overall efficiency can be evaluated by computing the average shortest path length among all pairs of nodes using the following equation:

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d_{ij},$$
 (7)

where *n* represents the total number of nodes in the network, while  $d_{ij}$  denotes the shortest path length between node *i* and node *j*. By taking the reciprocal of the shortest path lengths, one can obtain a measure of efficiency between nodes. The

TABLE 4: The top ten industrial nodes in clustering coefficient and weighted clustering coefficient.

Clustering coefficient	Node	Weighted clustering coefficient	Node
0.83686275	C4	0.00554410	H29
0.82589984	H40	0.00544787	C34
0.80869821	G11	0.00507715	G11
0.79953244	B40	0.00463909	D40
0.78962844	B21	0.00459691	B1
0.77793361	H4	0.00453500	F25
0.77466106	G40	0.00444192	D33
0.77345018	H8	0.00437377	F1_
0.77191358	B9	0.00431369	D38
0.77124888	H18	0.00414265	H30

average of these reciprocal values across all pairs of nodes in the network is termed the global network efficiency.

In this section, we employ the Dijkstra algorithm to calculate the shortest paths and their corresponding lengths within the directed weighted network under investigation. The fundamental idea behind the algorithm is to start with a source node and incrementally expand the set of nodes for which the shortest paths are known. This expansion occurs by exploring nodes that are adjacent to the current set but have not yet been visited. More specifically, the algorithm maintains two sets: one consisting of nodes with already-known shortest paths and another set comprising candidate nodes. The algorithm iteratively selects the next node with the shortest path from the candidate set until the shortest paths to all nodes have been identified. After calculating the shortest path lengths between all node pairs, the paths corresponding to the minimum and maximum values of these lengths are highlighted in the network, as shown in Figure 3. The shortest path length from node A1 to A6 is the smallest, at 2.286. This suggests that the input from the Agriculture, Forestry, Fishing, and Hunting sector in the Northeast to the Food and Tobacco sector in the same region is exceptionally direct, bypassing intermediary industries, and is also significant in volume. The shortest path length in Figure 3  $H8 \longrightarrow H7 \longrightarrow H30 \longrightarrow H5 \longrightarrow$  $H3 \longrightarrow C11 \longrightarrow C35 \longrightarrow A33 \longrightarrow A23$ . In the network diagram, the longest shortest path length is 1525.267, extending from the Textile, Clothing, Footwear, and Leather Goods sector in the Greater Northwest region to the Waste and Scrap sector in the Northeast. This path sequentially traverses seven different industrial sectors: textiles (Greater Northwest), transportation, warehousing, and postal services (Greater Northwest), nonmetallic minerals and other mining products (Greater Northwest), crude petroleum and natural gas (Greater Northwest), petroleum, coking products, and nuclear fuel processing goods (Eastern Coast), leasing and business services (Eastern Coast), and finance (Northeast). Importantly, this path crosses three major economic zones. The pivotal point is the input from crude petroleum and natural gas (Greater Northwest) to petroleum, coking products, and nuclear fuel processing goods (Eastern Coast), highlighting the transfer of abundant petroleum and natural gas resources from the Greater Northwest to the Eastern Coastal economic zone. This path sequentially traverses seven different industrial sectors,

#### 10

each of which could be a manifestation of regional economic specialization. For example, the Greater Northwest's focus on Crude Petroleum and Natural Gas might be a function of its resource endowments. Importantly, this path crosses three major economic zones, illustrating how regional capabilities contribute to forming complex interindustrial relationships. The pivotal point is the input from crude petroleum and natural gas (Greater Northwest) to petroleum, coking products, and nuclear fuel processing goods (Eastern Coast). This could signify a vital input-output linkage in the supply chain, reinforcing the structural interconnectedness of these industries across regions. Such relationships might be indicative of high transaction costs, as evidenced by the involvement of sectors like Transportation, Warehousing, and Postal Services in the longest shortest path.

The calculated average shortest path length for the entire network is 219.7, indicating that, on average, the shortest path lengths between industries are relatively long. This relatively long average path length suggests that there may be inefficiencies or bottlenecks in the system, either due to regulatory hurdles or inherent complexities in production processes. The global efficiency of the weighted network is 0.00728, signifying that the efficiency of resource or product propagation among the industries is low. This could reflect a high degree of market power or anticompetitive behavior among some sectors, impeding the efficient flow of goods and services.

#### 5. Network Structure Stability Analysis

This section explores the stability of the network structure, a critical aspect related to the network's resilience and antidisturbance capabilities. By simulating attacks on the complex network, we analyze its fault tolerance and resistance to targeted disruptions, thereby shedding light on the overall network stability.

5.1. Centrality Measures. A primary focus is the analysis of network centrality, an essential facet of network stability research. Centrality reveals key nodes and vulnerabilities within the network, contributing to our understanding of the structural characteristics and relative importance of individual nodes. This section evaluates four types of centrality in the directed, weighted network: degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality. This is echoed in the work of Ufimtsev et al. [27], who emphasize the impact of centrality measures on understanding the stability of networks, especially under conditions of noise and disturbance. This section evaluates four types of centrality in the directed, weighted network: degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality. The relevance of these measures in different network contexts, as discussed by Grando et al. [37], underscores their utility in identifying influential nodes and assessing network resilience. Furthermore, the study by Rajeh et al. [38] highlights the importance of considering community structures in centrality analysis, which can be particularly pertinent in complex

## Complexity



FIGURE 3: Schematic diagram of the shortest and longest paths in the shortest path.

industrial networks. The correlation analysis of centrality measures by Ficara et al. [39] provides a comprehensive understanding of how these different centrality types interact and influence each other, enriching the analysis of network stability in this research.

*5.1.1. Degree Centrality.* This measure reflects the number of connections a node has, thereby identifying the most popular and active nodes in the network. The failure of these nodes could severely impact network stability.

5.1.2. Closeness Centrality. Calculated as the average shortest path length from one node to all other nodes, this centrality measure gauges the node's importance. A higher closeness centrality implies better accessibility and faster information dissemination, indicating that the node may play a critical role in the network.

5.1.3. Betweenness Centrality. This measure is based on the frequency with which a node appears in all shortest paths across the network. It uncovers nodes that act as "bridges" in the network. These nodes appear in the shortest paths between many pairs of nodes; thus, their failure could substantially alter the network structure.

5.1.4. Eigenvector Centrality. This form of centrality is a function of the importance of the neighbors to which a node is connected. It reflects a node's social influence within the network. Nodes connected to influential nodes often have high eigenvector centrality, helping to identify nodes that may be crucial within the network.

This multi-dimensional approach to centrality provides a nuanced understanding of the elements that contribute to or jeopardize network stability. By identifying these critical nodes and potential vulnerabilities, we gain insights that can inform strategies for enhancing network resilience and efficiency.

In Table 5, the top ten industrial nodes are ranked according to the four different types of centrality measures. Industries leading in degree centrality are mainly those in the sectors of transportation, warehousing, and postal services, specifically located in the middle reaches of the Yellow River, the Eastern Coast, and the Northeast of China. This highlights the pivotal role these sectors in China's centraleastern and northeastern regions play in the overall national economy.

For closeness centrality, the values are generally low, and the industries that rank the highest are predominantly in the Greater Northwest, specifically in sectors such as transportation equipment, metal product machinery, electrical machinery, and timber processing and furniture. Despite their peripheral geographical locations, these sectors may have a strategic position that allows for efficient information dissemination.

High betweenness centrality is mainly observed in industries located in the Eastern Coastal economic zone, specifically in sectors such as petroleum coking products and nuclear fuel processing, leasing and business services. In addition, the petroleum and natural gas extraction industries in the Greater Northwest and Northeast, along with the production and supply of electric and thermal Power in the Northeast and Eastern Coast, play key "bridge" roles in the network. Industries involved in coal mining in the middle reaches of the Yellow River and chemical products on the Eastern coast also exhibit high betweenness centrality.

Lastly, the top ten industries in terms of eigenvector centrality are all concentrated in the Northern Coastal economic zone, predominantly in light industry, manufacturing, and chemical products sectors. Particularly noteworthy are the textile and apparel sectors in this region, whose eigenvector centrality is far higher than all other industries, indicating the critical importance of their neighboring industries.

Analysis of centrality measures can unveil the most influential nodes, the best nodes for dissemination, and the key "bridge" nodes within the network. Protecting and closely monitoring these nodes can significantly enhance the network's stability and resilience to disturbances. This analytical approach is particularly important for policymakers and stakeholders who aim to safeguard critical infrastructure and optimize resource allocation. In the realm of regional economics and industrial organization theory, identifying these central nodes can also guide regional development policies, investment decisions, and crisis management strategies. Therefore, understanding centrality measures is not just a theoretical exercise but also an essential practice for ensuring effective economic management and sustainable development.

5.2. Network Connectivity Metrics. The resilience of a network to attacks refers to its ability to maintain structural and functional integrity in the face of adversarial actions, such as the failure or removal of nodes or edges. In this study, changes in network connectivity serve as a measure for assessing the network's robustness against attacks. Specifically, two key metrics are employed: network efficiency and the size of the largest connected subgraph.

Network efficiency, as elaborated in Section 3.3, is an important gauge of global connectivity within the network. The mathematical expression for network efficiency is provided in the following equation:

$$E = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d_{ij}},$$
 (8)

where is  $d_{ij}$  the shortest path length between node *i* and node *j*. When there is no edge between node  $d_{ij} = +\infty i$  and node *j*, the local efficiency between the pair of nodes is 0. After the previous calculation, the global efficiency of the initial directed weighted network before the attack started was 0.00728.

The largest connected subgraph refers to the connected portion of the network containing the maximum number of nodes. In directed networks, one can further distinguish between weakly and strongly connected subgraphs. In a weakly connected subgraph, any pair of nodes is mutually reachable through a sequence of directed edges (ignoring edge direction), whereas in a strongly connected subgraph, directionality must be considered. The size and characteristics of the largest connected subgraph serve as important indicators of network connectivity and stability. In this study, the strongly connected subgraphs of the directed network are computed using Kosaraju's algorithm. The ratio of the number of nodes in the largest connected subgraph is defined in (9):

$$Z = \frac{m'}{m},\tag{9}$$

where m' is the number of nodes in the largest connected subgraph postattack and m is the number of nodes in the largest connected subgraph of the initial network. The ratio Z is used to reflect changes in network connectivity subsequent to an attack. After computational analysis, it was established that the initial directed weighted network constitutes a strongly connected subgraph. Therefore, in this specific case, m = n = 336.

5.3. Random vs. Targeted Attacks on Network Resilience. Network attacks can broadly be classified into two categories: random attacks and targeted (or deliberate) attacks. Random attacks refer to nonspecific assaults on nodes or edges in the network. In such scenarios, the elements attacked are randomly chosen without taking into account their unique roles or significance within the network structure. Contrastingly, targeted attacks are orchestrated to impact key nodes or edges within the network selectively. These attacks can be executed based on various criteria. In the present study, we employ four types of centrality metrics, previously discussed, as the criteria for simulating targeted attacks.

TABLE 5: Centrality of complex networks.

Degree centrality	Node	Closeness centrality	Node	Betweenness centrality	Node	Eigenvector centrality	Node
1.26865672	E30	0.00589318	H18	0.13004737	C11	0.57392775	B7
1.23582090	C12	0.00570312	H24	0.12980606	C35	0.41973487	B8_
1.20895522	A30	0.00561065	H19	0.10832067	H3_	0.19114992	B10_
1.20895522	B30	0.00556078	H9_	0.10095630	A25	0.18414660	B9_
1.19104478	C30	0.00551666	A24	0.09131290	A3_	0.13041251	B17_
1.18208955	C35	0.00547256	H10_	0.08829207	E2_	0.12438422	B12
1.16119403	E12	0.00539353	F11	0.08822951	C12	0.12312023	B15
1.13731343	C29	0.00531417	D24	0.06560014	D35	0.10514796	B22
1.11940299	D10	0.00530257	A10	0.06124765	C25	0.10499113	B19
1.11044776	D12	0.00529745	H16	0.06085441	E30	0.10390616	B21

This section outlines five distinct attack strategies: random attacks, attacks targeting nodes with the highest degree centrality, the highest closeness centrality, the highest betweenness centrality, and the highest eigenvector centrality. For instance, under the highest degree centrality attack, the node with the maximum degree centrality is removed along with its associated edges. Subsequently, the connectivity indices and degree centrality are recalculated for the newly formed network. This process is iteratively repeated until the network's overall connectivity is reduced to zero. The remaining three targeted attacks follow a similar methodology.

Figure 4 illustrates the changes in network efficiency and the ratio of the number of nodes in the largest connected subgraph under the five attack strategies. From the efficiency change curve, it is evident that attacks based on the highest closeness centrality are markedly less effective than random attacks. Intriguingly, after 300 instances of these two attack types, there is a noticeable uptick in global efficiency. Specifically, at 326 instances of highest closeness centrality attacks, the network efficiency peaks at 0.0081, surpassing the initial network's efficiency, followed by a linear decrease as attacks continue.

Furthermore, the effectiveness of the attack strategies, as observed from the graph, is sequentially best to worst as follows: highest betweenness centrality, highest degree centrality, and highest eigenvector centrality; all three being more effective than random attacks. The ratio of the number of nodes in the largest connected subgraph reveals a consistent trend across all five strategies when the number of attacks is below 200. Beyond this point, the performance of targeted attacks based on the highest closeness centrality and random attacks remains similar, but the other targeted attacks exhibit superior performance. Among them, attacks based on the highest betweenness centrality are the most effective, followed by those based on the highest degree centrality and the highest eigenvector centrality.

The analysis quantified the resilience of the network by computing the number of attacks needed to completely disrupt its connectivity. Specifically, for each of the five strategies—random attacks, highest degree centrality attacks, highest closeness centrality attacks, highest betweenness centrality attacks, and highest eigenvector centrality attacks—the required number of attacks was 332, 305, 331, 312, and 324, respectively. This data suggests that the network in question possesses a commendable level of resilience and structural stability. Among the examined attack strategies, the highest betweenness centrality, and the highest degree centrality attacks warrant special attention. Industries characterized by high values of these centrality metrics emerge as critical nodes in the network and substantially influence its overall stability. Thus, they are pivotal in safeguarding the economic system. Targeted policy interventions and risk management strategies aimed at these high-centrality industries can further fortify the stability of the intricate input-output network under study.

5.4. Further Interpretations and Suggestions. In China, the central, eastern, and northeastern regions serve as pivotal hubs for the transportation, warehousing, and postal sectors, which wield considerable influence over the national economy and logistics network. The elevated centrality of these industries is accentuated by their extensive interconnections with other industrial sectors. Concurrently, peripheral industries such as transportation equipment, metal product machinery, electrical machinery, wood processing, and furniture, primarily situated in the remote Great Northwest, manifest notable metrics of network cohesiveness. This suggests their capacity to act as early indicators for network perturbations, efficiently disseminating information and adaptations across the network.

It is noteworthy that certain sectors, including petroleum coking products and nuclear fuel processing, exhibit pronounced intermediary centrality, thereby serving as critical nodes or "bridges" in the economic network. These industries facilitate essential conduits for the transit of resources and information. Furthermore, industries such as light manufacturing and chemical production, concentrated in the northern coastal economic zone, not only maintain high eigenvector centrality but also indicate analogous levels of centrality in their proximate regions, underlining the network resilience within this geographical area.

In light of these findings, policy implications emerge. Targeted infrastructure investments should be allocated preferentially to regions characterized by both high industrial density and intermediary centrality, thereby catalyzing more expansive economic activities. Concurrently, risk mitigation strategies should be proactively formulated



FIGURE 4: Changes in network efficiency and maximum connected subgraph node ratio under different attack strategies.

for highly network-centric industries, even those located in economically peripheral areas, to preempt potential cascading effects stemming from local network disruptions. Given the indispensable "bridging" role of industries with high intermediary centrality, supply chain risk management warrants particular focus to assure the uninterrupted and stable flow of goods and services. To bolster the vigor of regions with elevated eigenvector centrality, the cultivation of innovation clusters is advised to generate industrial synergies. In addition, strategic tax incentives and subsidies can be deployed to sustain competitiveness and stimulate growth in pivotal industries. Lastly, a real-time monitoring mechanism is recommended for tracking critical performance metrics in these key sectors, enabling timely interventions should severe fluctuations in centrality indicators arise

This analysis, rooted in empirical data, offers actionable insights for policy-making, aiming to enhance network resilience through data-driven strategies.

## 6. Conclusion

6.1. Further Interpretations and Suggestions. This study employs the Maximum Entropy Method to construct a directed, weighted complex network comprising 42 industrial sectors across China's eight major economic regions. The research offers an in-depth statistical analysis of the network's attributes and stability. Key findings include

- The maximum entropy method effectively uncovers significant intersectoral relationships, maximizing the extraction of information from input-output tables.
- (2) Degree analysis reveals that 85% of industrial sectors have degrees ranging between 70 and 220, indicating tight topological connections between industries. Notably, sectors such as transportation, warehousing, and postal services, along with wholesale and retail, play pivotal roles in the economic network. The distribution of in-degrees and out-degrees reflects distinct regional industrial structures and interdependencies. For instance, construction, public utilities, and service sectors are highly consumed in the Yellow River Basin and the Northwest regions.
- (3) In the weighted network, point, in-strength, and outstrength metrics further highlight interindustry connections and dependencies. Coastal Eastern regions display strong export-oriented characteristics in chemical product industries, while the Northern

coastal regions excel in textiles. In contrast, the Northeast and Northwest regions exhibit importdependent traits in various manufacturing sectors. These patterns are indicative of regional industrial specialization and varying levels of economic development.

The observed topological structure and industry roles can be situated within broader industrial and regional theories. For example, the concept of agglomeration economies may explain why certain industries like transportation and warehousing are centrally positioned. They potentially serve as clusters that generate additional economic advantages for nearby sectors.

Simulated network attacks reveal that the most effective targeted strategies are highest-betweenness centrality attacks, followed by highest-degree centrality and highesteigenvector centrality attacks. These deliberate attack methods outperform random attacks. While the industrial network demonstrates high resilience, targeted policy interventions are warranted for industries with high degree and betweenness centrality. These sectors emerge as critical nodes, influencing the stability of the network structure.

In summary, the network exhibits strong resilience but requires nuanced policy and risk management strategies aimed at industries with high centrality metrics, as these sectors are pivotal in maintaining the network's structural stability.

6.2. Limitations and Future Directions. Our investigation into the intricate relationships among industrial sectors within China's economic zones contributes to the understanding of complex networks in an economic context. Despite the insights gained, our study acknowledges several limitations that pave the way for future research directions. Firstly, the reliance on static data for network modeling in existing literature does not adequately reflect the dynamic nature of economic activities, resulting in diminished predictive capabilities. This highlights the necessity of incorporating dynamic data and models that can more accurately mirror the fluctuations and trends within economic networks.

Second, the simulation of network disruptions often fails to account for real-world economic shocks, such as financial crises or sudden market changes, which limits the practical relevance of these models. Future studies should aim to integrate real economic shock scenarios to enhance the applicability and resilience of network models.

Third, there is a notable gap in cross-disciplinary research regarding the examination of network analysis methods tailored to specific economic situations. This oversight may overlook the complex realities of economic environments, suggesting a need for more nuanced studies that evaluate the effectiveness of network methodologies within economic frameworks.

Fourth, strategies developed to enhance network robustness often do not take into account the unique characteristics and interdependencies of economic networks, particularly those defined by industry-specific interactions. This oversight underscores the importance of devising robustness strategies that are not only theoretically sound but also practically applicable, taking into consideration the distinct nature of economic networks.

To address these limitations, future research should focus on developing dynamic modeling approaches that accurately reflect economic network operations and their responses to various shocks. Additionally, there is a critical need for studies that assess the suitability of network analysis methods for economic contexts, ensuring that these tools can capture the complexity of economic scenarios. Tailoring strategies for robustness to address the unique characteristics of economic networks will enhance our comprehension and management of these systems efficiently. By tackling these areas, subsequent research can significantly advance the application of complex network theory in economic studies, leading to more robust and applicable insights for policy-making and strategic planning [40–47].

## **Data Availability**

The China City-level MRIO Table data used to support the findings of this study may be released upon application to the Carbon Emission Accounts ~~~~~~~amp; Datasets, who can be contacted at shanyuli@outlook.com or guandabo@ hotmail.com or zhuliu@tsinghua.edu.cn.

## **Conflicts of Interest**

The authors declare that they have no conflicts of interest.

## References

- D. J. Watts and S. H. Strogatz, "Collective dynamics of 'smallworld'networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [2] A. L. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica A: Statistical Mechanics* and Its Applications, vol. 272, no. 1-2, pp. 173–187, 1999.
- [3] M. Newman, *Networks*, Oxford university press, Oxford, UK, 2018.
- [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: structure and dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.
- [5] S. Iyer, T. Killingback, B. Sundaram, and Z. Wang, "Attack robustness and centrality of complex networks," *PLoS One*, vol. 8, no. 4, 2013.
- [6] Y. Casali and H. R. Heinimann, "A topological analysis of growth in the Zurich road network," *Computers, Environment* and Urban Systems, vol. 75, pp. 244–253, 2019.
- [7] A. Ficara, F. Curreri, G. Fiumara, and P. De Meo, "Human and social capital strategies for Mafia network disruption," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 1926–1936, 2023.
- [8] A. Ficara, F. Curreri, G. Fiumara, P. De Meo, and A. Liotta, "Covert network construction, disruption, and resilience: a survey," *Mathematics*, vol. 10, no. 16, p. 2929, 2022.
- [9] R. E. Miller and P. D. Blair, *Input-output Analysis: Foundations and Extensions*, Cambridge university press, Cambridge, UK, 2009.

- [10] M. Á. Serrano, M. Boguñá, and A. Vespignani, "Patterns of dominant flows in the world trade web," *Journal of Economic Interaction and Coordination*, vol. 2, no. 2, pp. 111–124, 2007.
- [11] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.
- [12] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American power grid," *Physical Review*, vol. 69, no. 2, 2004.
- [13] R. Albert, H. Jeong, and A. L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.
- [14] R. Kumar and A. Singh, "Robustness in multilayer networks under strategical and random attacks," *Procedia Computer Science*, vol. 173, pp. 94–103, 2020.
- [15] X. Sun, V. Gollnick, and S. Wandelt, "Robustness analysis metrics for worldwide airport network: a comprehensive study," *Chinese Journal of Aeronautics*, vol. 30, no. 2, pp. 500–512, 2017.
- [16] R. V. Donner, M. Small, J. F. Donges et al., "Recurrence-based time series analysis by means of complex network methods," *International Journal of Bifurcation and Chaos*, vol. 21, no. 4, pp. 1019–1046, 2011.
- [17] C. Christensen and R. Albert, "Using graph concepts to understand the organization of complex systems," *International Journal of Bifurcation and Chaos*, vol. 17, no. 7, pp. 2201–2214, 2007.
- [18] X. Cheng and J. M. A. Scherpen, "Model reduction methods for complex network systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, pp. 425–453, 2021.
- [19] F. Emmert-Streib, A. Musa, K. Baltakys et al., "Computational analysis of the structural properties of economic and financial networks," 2017, https://arxiv.org/abs/1710.04455.
- [20] O. Polishchuk, "About network structures and systems complexity," 2018, https://arxiv.org/abs/1807.06044.
- [21] J. Dshalalow and R. White, "On Reliability of Stochastic Networks," 2019, https://arxiv.org/abs/1901.07137.
- [22] M. Fabris and D. Zelazo, "A robustness analysis to structured channel tampering over secure-by-design consensus networks," *IEEE Control Systems Letters*, vol. 7, pp. 2011–2016, 2023.
- [23] C. Sarraute, F. Miranda, and J. Orlicki, "Simulation of Computer Network Attacks," 2007, https://arxiv.org/abs/ 1006.2407.
- [24] P. Y. Chen and S. M. Cheng, "Sequential defense against random and intentional attacks in complex networks," *Physical Review*, vol. 91, no. 2, 2015.
- [25] O. Bel, J. Kim, W. J. Hofer, M. Maharjan, S. Purohit, and S. Niddodi, "Co-simulation framework for network attack generation and monitoring," 2023, https://arxiv.org/abs/2307. 09633.
- [26] J. Platig, E. Ott, and M. Girvan, "Robustness of network measures to link errors," *Physical Review*, vol. 88, no. 6, 2013.
- [27] V. Ufimtsev, S. Sarkar, A. Mukherjee, and S. Bhowmick, "Understanding stability of noisy networks through centrality measures and local connections," in *Proceedings of the 25th* ACM International on Conference on Information and Knowledge Management, pp. 2347–2352, Indianapolis, IN, USA, October 2016.
- [28] C. Saxena, M. N. Doja, and T. Ahmad, "Stability of local information-based centrality measurements under degree preserving randomizations," in *Proceedings of the Intelligent*

*Computing and Information and Communication*, pp. 395–403, Springer Singapore, Rajasthan, India, November 2018.

- [29] S. Oldham, B. Fulcher, L. Parkes, A. Arnatkevičiūtė, C. Suo, and A. Fornito, "Consistency and differences between centrality measures across distinct classes of networks," *PLoS One*, vol. 14, no. 7, 2019.
- [30] N. Gupta, A. Singh, and H. Cherifi, "Centrality measures for networks with community structure," *Physica A: Statistical Mechanics and Its Applications*, vol. 452, pp. 46–59, 2016.
- [31] H. Zheng, Y. Bai, W. Wei et al., "Chinese provincial multiregional input-output database for 2012, 2015, and 2017," *Scientific Data*, vol. 8, no. 1, p. 244, 2021.
- [32] G. Loaiza-Ganem, Y. Gao, and J. P. Cunningham, "Maximum Entropy Flow Networks," 2017, https://arxiv.org/abs/1701. 03504.
- [33] C. Metzig and C. Colijn, "A maximum entropy method for the prediction of size distributions," *Entropy*, vol. 22, no. 3, p. 312, 2020.
- [34] H. Zenil, N. A. Kiani, and J. Tegnér, "The thermodynamics of network coding, and an algorithmic refinement of the principle of maximum entropy," *Entropy*, vol. 21, no. 6, p. 560, 2019.
- [35] J. Jalving, Y. Cao, and V. M. Zavala, "Graph-based modeling and simulation of complex systems," *Computers & Chemical Engineering*, vol. 125, pp. 134–154, 2019.
- [36] L. Torres, A. S. Blevins, D. Bassett, and T. Eliassi-Rad, "The why, how, and when of representations for complex systems," *SIAM Review*, vol. 63, no. 3, pp. 435–485, 2021.
- [37] F. Grando, L. Z. Granville, and L. C. Lamb, "Machine learning in network centrality measures: tutorial and outlook," ACM Computing Surveys, vol. 51, no. 5, pp. 1–32, 2018.
- [38] S. Rajeh, M. Savonnet, E. Leclercq, and H. Cherifi, "Investigating centrality measures in social networks with community structure," in *Complex Networks & Their Applications IX: Volume 1*, pp. 211–222, Springer International Publishing, Cham, 2021.
- [39] A. Ficara, G. Fiumara, P. De Meo, and A. Liotta, "Correlation analysis of node and edge centrality measures in artificial complex networks," in *Proceedings of theSixth International Congress on Information and Communication Technology: ICICT 2021*, pp. 901–908, London, UK, September 2021.
- [40] Y. Fan, H. Liu, and S. He, "Mitigating Congestion in Complex Transportation Networks via Maximum Entropy," 2017, https://arxiv.org/abs/1701.04974.
- [41] V. Venkatasubramanian, D. N. Politis, and P. R. Patkar, "Entropy maximization as a holistic design principle for complex optimal networks," *AIChE Journal*, vol. 52, no. 3, pp. 1004–1009, 2006.