



Article

Self-Similar Growth and Synergistic Link Prediction in Technology-Convergence Networks: The Case of Intelligent Transportation Systems

Yuxuan Xiu ^{1,2} , Kexin Cao ^{1,2} , Xinyue Ren ^{1,2} , Bokui Chen ^{1,3,4} and Wai Kin (Victor) Chan ^{1,2,4,*} ¹ Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China² Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen 518055, China³ Department of Network Intelligence, Peng Cheng Laboratory, Shenzhen 518055, China⁴ International Science and Technology Information Center, Shenzhen 518055, China

* Correspondence: chanw@sz.tsinghua.edu.cn

Abstract: Self-similar growth and fractality are important properties found in many real-world networks, which could guide the modeling of network evolution and the anticipation of new links. However, in technology-convergence networks, such characteristics have not yet received much attention. This study provides empirical evidence for self-similar growth and fractality of the technology-convergence network in the field of intelligent transportation systems. This study further investigates the implications of such fractal properties for link prediction via partial information decomposition. It is discovered that two different scales of the network (i.e., the micro-scale structure measured by local similarity indices and the scaled-down structure measured by community-based indices) have significant synergistic effects on link prediction. Finally, we design a synergistic link prediction (SLP) approach which enhances local similarity indices by considering the probability of link existence conditional on the joint distribution of two scales. Experimental results show that SLP outperforms the benchmark local similarity indices in most cases, which could further validate the existence and usefulness of the synergistic effect between two scales on link prediction.

Keywords: intelligent transportation systems; international patent classification; technology convergence; patent; fractal analysis; self-similarity; partial information decomposition; emergence; link prediction



Citation: Xiu, Y.; Cao, K.; Ren, X.; Chen, B.; Chan, W.K. Self-Similar Growth and Synergistic Link Prediction in Technology-Convergence Networks: The Case of Intelligent Transportation Systems. *Fractal Fract.* **2023**, *7*, 109. <https://doi.org/10.3390/fractalfract7020109>

Academic Editors: Matej Babič and Andrej Srakar

Received: 20 September 2022

Revised: 5 January 2023

Accepted: 10 January 2023

Published: 20 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Technology innovation is the major competitiveness of enterprises and an essential driving force for the high-quality development of the national economy. The innovation of technology could be broadly classified into two categories: (1) technology substitution, which refers to the innovative breakthrough of new technologies that replace the existing ones, and (2) technology convergence, which refers to the transfer and combination of existing knowledge among multiple areas of technology [1,2]. The former is a linear, step-by-step approach along with challenging difficulties and lengthy transition periods, whereas the latter is a nonlinear and complementary method of technology fusion that brings with it additional cooperation and inspiration. Technology convergence is regarded as an increasingly significant characteristic of current trends of technology innovations due to more realization paths and faster outcome emergence [3–5].

A prominent research strand of technology convergence is its anticipation, which aims to predict potential technology convergence or identify convergence movements at an early stage [6–8]. In many studies, the emergence of new technology convergence is defined as the first co-occurrence of two International Patent Classification (IPC) subclasses in the patent data [9,10], and the anticipation is formulated as a link-prediction problem on an IPC co-occurrence network [11–13]. Two main objectives of the literature for

anticipation of technology convergence are developing new methods and their practical applications [8]. From the perspective of network science, these anticipation approaches could be categorized into three categories: (1) topology-based [9,11–14], (2) ensemble-learning-based [7,15,16], and (3) multi-modal-based [17–19]. These methods are successfully applied in a wide range of industries, such as power systems [15], smartphones [11], electric vehicles [13], manufacturing [14], and telecommunications [17].

Despite the success of applying link-prediction algorithms to the anticipation of technology convergence, as noted in a recent milestone review by Sick et al. [8], there is relatively little research on theoretical underpinnings for such applications, especially from the perspective of network science. A major research topic that has not yet been adequately explained is particularly addressed in this study: As a time-evolving network, what evolutionary process may the technology-convergence network undergo? This question is extremely essential for the anticipation of technology convergence, since studying the evolution process of a network could help us understand the growth mechanism of its links and hence develop better link-prediction algorithms [20–25].

To answer the above question, we employed the theory of fractal networks as a tool for analysis. According to the dynamic evolutionary process, fractal networks could be generated by self-similar growth [26–28]. From the viewpoint of static topological structure, typical properties of fractal networks include self-similarity and scale invariance [29–32]. Using historical patent data in the ITS field, this study first provides empirical evidence for the self-similar growth of the time-evolving technology-convergence network based on the commonly adopted rescaled network statistics [28,33–37] and then verifies the fractality of each snapshot using a community-structure-based approximation of the box-cover algorithm [38].

After discovering the fractal property of the technology-convergence network, another intriguing research question emerges: How does the microscopic structure of the original technology-convergence network and its scaled-down replica provide information for the prediction of future technology convergence? This research question has some significant theoretical implications, since it could give insights into the multi-scale modeling and prediction of technology convergence, which has not yet been thoroughly studied.

As an attempt to answer the issue raised, this study adopts the partial information decomposition (PID) framework for the link-prediction problem. In particular, we discovered that the IPC hierarchy naturally provides community partitions for the original network, based on which the scaled-down network could be obtained. We further define a microscopic variable and a macroscopic variable based on two-hop link predictors and IPC hierarchy, respectively, to represent the two scales of the network. With PID, the joint contribution of the original and the corresponding scaled-down networks is decomposed into four components: the synergistic and redundant information between two scales, and the unique information provided by each scale. It is found that the joint contribution is mainly in the form of synergistic information and much redundant information.

The synergistic effect of two different scales on link prediction implies that the joint distribution of the two scales could be more informative than the marginal distributions of each scale. To further validate the existence and usefulness of the synergistic effect, we designed a synergistic link prediction (SLP) approach, which is essentially the probability of link existence on the condition of both the microscopic and the macroscopic variables. We adopted a cumulative distribution function (CDF)-like formula to score the likelihood of link existence between two nodes to address the sparsity issue of the network, while ensuring that the link-prediction score is monotonically increasing with the likelihood of link existence. SLP could be categorized as the link/community-based strategy [39–47], which enhances local similarity measures based on community information (see Section 2.4). Compared with existing link/community strategies, SLP could be used more easily in a plug-and-play manner, since it does not require manual adjustment of the formulae for local similarity measures.

Experiments were conducted by using the ITS-related patent data from November 1995 to July 2022. A time-evolving technology-convergence network was constructed on the basis of the IPC co-occurrence relationships obtained from the historical patent applications. Further, the new technology convergence of the next year is anticipated based on annual snapshots. Experimental results show that, in most cases, our designed SLP algorithm outperforms the benchmark algorithms that only consider one of the two scales. As a very interesting phenomenon, the link-prediction performance is highly consistent with PID results. The PID framework provides an appropriate theoretical underpinning for deeply interpreting the link-prediction performance. In most cases where SLP outperforms the benchmark algorithms, synergistic information also dominates the interaction between two scales of the network in the PID results. In the few cases where SLP has not managed to outperform the benchmark algorithms, significant redundant or unique information can be observed in the PID results. These results imply that the performance improvement provided by multi-scale fusion for link prediction is primarily due to synergy between different scales, and the presence of significant redundant or independent information may weaken the performance and necessitates special treatment. Our findings shed light on the development of multi-scale link-prediction algorithms for the anticipation of technology convergence and demonstrate that PID could be a powerful tool for exploring the evolutionary mechanisms of networks.

In summary, the main contributions of this study could be summarized as follows:

- We provide empirical evidence for the fractal characteristic of the technology-convergence network in the ITS field. In terms of the time-evolving mechanism, the technology-convergence networks grow in a self-similar paradigm, in which the rescaled topological properties remain stable. In terms of spatial properties, the annual snapshots of the technology-convergence networks are identified as fractal networks.
- We discovered that the structural information at two different scales has a synergistic effect on link prediction. The structural information at two different scales is measured by local similarity measures and community-based indices, respectively. Therefore, this discovery implies that the joint distribution of the two could be more informative than the marginal distributions of either the local or the community-based indices.
- We designed a link-prediction approach, namely, the SLP approach, based on the joint conditional probability of link existence given both the local and the community-based indices. Experimental results show that the SLP approach could enhance the corresponding local similarity measures by incorporating community structures, which further validates the existence and usefulness of the synergistic effect on link prediction between two scales.

The rest of this article is organized as follows. Theoretical background and related work for the research in this paper are provided in Section 2. Section 3 details the methodology of this study, including the empirical criteria for self-similar growth and fractality, the use of PID, and our developed SLP algorithm. Section 4 presents experimental results obtained from applying the new methodology and the benchmark algorithms. Finally, Section 5 presents the conclusions drawn from this research and directions for future research.

2. Related Work

The aims of this study were to analyze properties of the technology-convergence network based on fractal network theory and PID, and to propose a link-prediction algorithm for the anticipation of future technology convergence. In this section, related work in the following four research areas is discussed: technology convergence and its anticipation, fractal analysis of complex networks, PID, and link prediction.

2.1. Technology Convergence and Its Anticipation

The concept of technology convergence could be traced back to Rosenberg's study on the technology changes in the machine tool industry in the 1960s [48]. However, little attention was paid to technology convergence until the 1980s, when information

and communication technology (ICT) was extended to and integrated with various other industrial sectors [49–51]. In 1992, Kodama’s milestone of an article [49] symbolized that technology convergence began to attract attention. Some prior studies on convergence emerged in the 2000s, including theoretical definitions of convergence [52], the distinction between demand-side and supply-side convergence [53], and the partitioning of the four stages (i.e., science, technology, market, and industry) of the convergence process [52]. This study follows the theoretical definition of convergence as, “The blurring of boundaries between at least two hitherto disjoint areas” [54], and addresses the stages of technology convergence from a supply-side perspective.

Data-driven research on technology convergence has gained appeal since the 2010s, as publicly available data have drastically multiplied. An important data source for technology convergence studies is patent data, which are considered an up-to-date and reliable indicator of creative activities and knowledge accumulation within technology fields [55–58]. In many early studies of technology convergence, citation relationships among patents were used to measure technology convergence [55,59,60]. However, a recent argument is that such citation-based measurements focus on modeling the stretching process of the knowledge from one technology field to another and may not clearly characterize the events that represent the occurrence of technology convergence [10]. The emergence of new technology convergence could be clearly defined as the first co-occurrence of two IPC subclasses in the patent data based on the International Patent Classification (IPC) of patents [9,10]. This definition of technology convergence was widely adopted in subsequent studies [5,13,17,61–63].

Benefiting from the IPC-based definition of the technology convergence, several studies anticipating prospective technology convergence based on IPC co-occurrence networks have emerged in recent years [9,13,17]. This branch of research constructs networks based on IPC co-occurrence relationships and forecasts potential technology convergence as missing links in the networks. A wide range of studies are based on topological similarity indices, such as the weighted resource allocation (RA) index [13], the weighted common neighbor (WCN) index [11], the Adamic-Adar index [9], and modified versions of these indices. Another stream of research applies ensemble approaches to link prediction, which employs supervised classification models to learn combinations of several individual link predictors [7,15,16]. A novel research paradigm is to develop machine learning algorithms that predict missing links based on a combination of multi-modal information, such as topology, bibliometric, and semantic information [17–19].

Despite the successful implementations of the proposed methods in a variety of industries, there is a paucity of research on the theoretical underpinnings of anticipating technology convergence [8]. In particular, studies that analyze the properties and evolutionary processes of technology-convergence networks from a network science perspective are very scarce. Hence, the key contributions of this paper are considering the technology-convergence network as a time-evolving network and providing its three properties: (1) the self-similar growth of the evolutionary process, (2) the fractality of the snapshots, and (3) the synergy between different scales in link prediction.

2.2. Fractal Analysis of Complex Networks

The complex network provides a powerful theoretical tool for the abstract characterization of many real-world systems composed of various objects and mutual relationships, such as technological [13,64,65], biological [66–69], and social systems [70–74]. A wide range of real-world networks have the fractal property, which could be roughly described as, “The network looks similar under different magnification levels” [75].

The rigorous definition of fractality in a complex network is based on the box-covering approach. Given a box of size l_B , the box-covering approach groups a set of nodes into a box b_r such that the shortest distance between each pair of nodes in b_r is less than l_B . By partitioning all the nodes into boxes, a network is covered, and the minimum number of boxes required is denoted as $N_B(l_B)$. The circumstance where a network is covered by

$N_B(l_B)$ boxes is referred to as the optimal covering. If the scaling of $N_B(l_B)$ of the optimal covering follows a power law as

$$N_B \sim l_B^{-d_B}, \quad (1)$$

then the network is identified as a fractal network and d_B is denoted as its fractal dimension.

The optimal covering problem is known to be NP-hard [38,76]. Many heuristic algorithms have been proposed to solve this problem approximately, such as the greedy coloring algorithm [76], the compact-box-burning algorithm [76], the max-excluded mass-burning algorithm [76], the overlapping-box-covering algorithm [77], and the MCWR algorithm [78]. For a comparative analysis of the aforementioned box-covering algorithms, see the study in [79].

Most existing box-covering algorithms are designed for unweighted networks, whereas the technology-convergence network constructed in this study is a weighted network. For the box-covering problem of weighted networks, Wei et al. [80] proposed a method that transforms edge weights using a power function. Nevertheless, the partitioning of boxes in this method is affected not only by l_B , but also by the exponent of the power function, which increases the difficulty of modifying parameters in actual use. Hence, selecting a proper exponent in the practical application remains a problem. Another recent step was the community-structure-based method proposed by Giudicianni et al. [38], which approximates the box-covering problem with community detection algorithms. Since there are many off-the-shelf algorithms for community detection on weighted networks, this method provides many reliable options for the coarse-graining of weighted networks. Therefore, we chose to apply this community-structure-based approach to analyze the fractal property of the technology-convergence network.

The inverse process of the box-covering algorithm mirrors the evolution of fractal networks to some extent [26,75,81]. For a time-evolving network, if its growth mechanism approximates the inverse process of the box-covering algorithm, the snapshot at time $t - 1$ could be considered as a scaled-down replica of the snapshot at time t . As a result, we could observe self-similar behaviors of the network statistics on the rescaled snapshots. Commonly adopted empirical evidence for the self-similar growth is the curve overlapping of rescaled network statistics, including rescaled degree distribution, rescaled clustering coefficient, rescaled degree-degree correlation, and the community structure [28,33–37]. Following the convention in the literature, this study provides empirical evidence for the self-similar growth of the time-evolving technology-convergence network based on the aforementioned rescaled network statistics.

2.3. Partial Information Decomposition

Information theory, pioneered by Claude Shannon [82], is widely used to analyze the interactions between components in complex systems. However, the commonly applied mutual information only concerns the interaction between two variables, as it measures the information provided by a single variable about another. Instead, this study considers a three-way system involving the original technology-convergence network and its scaled-down replica as two inputs and the future technology convergence as an output.

A recent extension of information theory, the partial information decomposition (PID), provides a powerful tool for analyzing the three-way system with two inputs, X_1 and X_2 , and one output, Y [83]. Specifically, the information that X_1 and X_2 jointly provide for Y is defined as

$$I(Y; X_1, X_2) = \sum_{x_1} p_{Y, X_1, X_2}(y, x_1, x_2) \left[\log \frac{1}{p_Y(y)} - \log \frac{1}{p_{Y|X_1, X_2}(y | x_1, x_2)} \right]. \quad (2)$$

The joint mutual information $I(Y; X_1, X_2)$ is further decomposed into four terms:

$$I(Y; X_1, X_2) = I_{\text{red}}(Y; X_1, X_2) + I_{\text{unq}}(Y; X_1 | X_2) + I_{\text{unq}}(Y; X_2 | X_1) + I_{\text{syn}}(Y; X_1, X_2), \quad (3)$$

where I_{red} , I_{unq} , and I_{syn} are named redundant information, unique information and synergistic information, respectively.

I_{red} is defined as “the expected value of the minimum information provided by either X_1 or X_2 about each possible outcome of Y ” [83], which is formally written as

$$I_{\text{red}}(Y; X_1, X_2) = \sum_y p_Y(y) \min\{I(Y = y; X_1), I(Y = y; X_2)\}. \quad (4)$$

Given the definition of I_{red} , the other three terms could be induced as

$$I_{\text{unq}}(Y; X_1 | X_2) = I(Y; X_1) - I_{\text{red}}(Y; X_1, X_2), \quad (5)$$

$$I_{\text{unq}}(Y; X_2 | X_1) = I(Y; X_2) - I_{\text{red}}(Y; X_1, X_2), \quad (6)$$

$$I_{\text{syn}}(Y; X_1, X_2) = I(Y; X_1, X_2) - I_{\text{unq}}(Y; X_1 | X_2) - I_{\text{unq}}(Y; X_2 | X_1) - I_{\text{red}}(Y; X_1, X_2). \quad (7)$$

The practical meanings of these four terms are as follows:

- $I_{\text{red}}(Y; X_1, X_2)$ represents the information that both X_1 and X_2 could provide for Y .
- $I_{\text{unq}}(Y; X_1 | X_2)$ represents the information about Y that could be provided by X_1 but not by X_2 .
- $I_{\text{unq}}(Y; X_2 | X_1)$ represents the information about Y that could be provided by X_2 but not by X_1 .
- $I_{\text{syn}}(Y; X_1, X_2)$ represents the information about Y that could only be obtained by jointly considering both X_1 and X_2 .

PID has been applied in various research fields, such as neuroscience [84–86], artificial intelligence [87–90], and emergent behavior recognition and analysis [91–93]. Among all these, the most relevant is the information-theoretic framework for causal emergence proposed by Rosas et al. [91].

In the proposed information-theoretic framework, a complex system composed of n microscopic states $\mathbf{X}_t = [X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(n)}]$ is taken into account. It is assumed that the system conforms to Markovian stochastic dynamics with the transition probability $p_{\mathbf{X}_{t'}|\mathbf{X}_t}$, where t and t' are two time points satisfying the condition $t < t'$. An emergent behavior is deemed to have occurred if some macroscopic states V_t of the system could provide unique information beyond that provided by microscopic states \mathbf{X}_t separately. The unique information provided by V_t is denoted as $I_{\text{unq}}(V_t; \mathbf{X}_{t'} | \mathbf{X}_t)$, and V_t demonstrates causal emergence if

$$I_{\text{unq}}(V_t; \mathbf{X}_{t'} | \mathbf{X}_t) > 0, \quad (8)$$

(see Definition 2 in [91]).

By analogy with the above information-theoretic framework, the original technology-convergence network corresponds to the microscopic state, and its scaled-down replica corresponds to the macroscopic state. However, defining microscopic and macroscopic variables is not a trivial or intuitive task. The microscopic variables of a system should not contain information about the high-order interactions among its components, and the definition of macroscopic variables relies on manually designed coarse-graining functions. Since there is no universal methodology for defining microscopic variables and the coarse-graining functions for deriving macroscopic variables, this study contributes to the PID literature by proposing definitions of microscopic variables and coarse-graining functions in complex networks from the perspective of link prediction. This study could be considered as an attempt to apply PID to the field of link prediction in complex networks. In addition,

from a machine learning perspective, the link-prediction method proposed in this study could be viewed as a PID-inspired feature fusion algorithm. Compared with previous studies on PID-inspired feature engineering [88,89], this study proposes a novel feature fusion method based on the cumulative distribution function (CDF).

2.4. Link Prediction

Link prediction, a basic issue in the study of complex networks, aims to predict unobserved missing links based on the the observed current network structure. The formal settings of link prediction are given as follows.

An observed network $G(\mathbf{N}, \mathbf{E})$ is considered, where \mathbf{N} and \mathbf{E} represent the node set and the edge set, respectively. It is assumed that $G(\mathbf{N}, \mathbf{E})$ is an undirected network without multi-edges and self-loops. The number of nodes is denoted as $N = |\mathbf{N}|$, and the number of the observed edges is denoted as $E = |\mathbf{E}|$. The adjacency matrix of $G(\mathbf{N}, \mathbf{E})$ is denoted as $\mathbf{A} = [a_{ij}]$, which is an $N \times N$ real symmetric matrix with $a_{i,i} = 0$. The entry $a_{ij} = 0$ indicates that no link is observed between node i and node j —i.e., $(i, j) \notin \mathbf{E}$, whereas $a_{ij} > 0$ indicates that a link is observed between node i and node j with a weight of a_{ij} . Note that only the undirected weighted network with positive edge weights is considered in this study.

In such a network, the total number of node pairs is $\frac{N(N-1)}{2}$, and the number of unconnected node pairs is $\frac{N(N-1)}{2} - E$. We use \mathbf{U} to denote the set of unconnected node pairs. The link-prediction problem assumes that there are unobserved or future links between some node pairs $(i, j) \in \mathbf{U}$, the set of which is denoted as \mathbf{E}' . The task of link prediction is to find the missing links in \mathbf{E}' based on the observed network $G(\mathbf{N}, \mathbf{E})$. Generally, a link-prediction algorithm f , also known as a link predictor, computes a score s_{ij}^f for each node pair $(i, j) \in \mathbf{U}$ based on the adjacency matrix \mathbf{A} , which could also be written in a matrix form as

$$\mathbf{S}^f = [s_{ij}^f]_{N \times N} = f(\mathbf{A}). \quad (9)$$

The node pairs with higher scores are the most likely to be the missing links.

The existing link predictors could be broadly classified into three categories: (1) similarity-based indices, which score node pairs according to their topological similarities [94–97]; (2) maximum likelihood methods, where some parametric network models are presumed and the parameters of the models are estimated by maximizing the likelihood of the observed network structure [98–100]; (3) machine-learning-based methods, in which the link-prediction problem is regarded as a binary classification problem and is solved through machine learning algorithms (e.g., embedding learning [101–103], ensemble learning [104–106], and graph neural networks [107–109]). For the latest advances in the field of link prediction, two seminal reviews were published in 2020 [110] and 2021 [111].

The similarity-based indices are the most widely used approaches in the anticipation of technology convergence [9,11–14], mainly due to their low computational complexity and high interpretability. Widely applied similarity-based indices include the common neighbor (CN) index [94], the preferential attachment (PA) index [112], the Jaccard (JC) index [95], the resource allocation (RA) index [96], and the Adamic-Adar (AA) index [97]. Since the technology-convergence networks are essentially weighted networks, the weighted similarity-based indices have attracted much attention, such as the weighted resource allocation (WRA) index [13] and the weighted common neighbor (WCN) index [11]. Other weighted similarity-based indices that could be applied or extended for technology-convergence networks include the weighted Jaccard coefficient (WJC) [113], the weighted preferential attachment (WPA) index [113], the selectivity index [114], and the inverse selectivity index [114]. In particular, Martinčić-Ipšić et al. [114] provided a clear and comprehensive example of how to design and evaluate similarity-based indices for real-world weighted networks.

The above-mentioned mainstream methods could be summarized as local similarity measures, since they mainly exploit the information of 2-hop paths in complex networks.

However, due to the hierarchical structure of the IPC system, technology-convergence networks naturally have a priori community structures (i.e., IPC subclasses belonging to the same IPC class can be considered as nodes from the same community), which could provide information beyond local similarity measures. Therefore, link prediction indices that jointly consider the local similarity and the community structure could be particularly suited for the anticipation of technology convergence.

In the link prediction literature, the methods that enhance local similarity measures based on community information are known as the link/community-based strategies [39–42,44,45,47]. A pioneering work in this research was [39], which was based on the assumption that two nodes are more likely to connect if their common-first-neighbors form a strongly inner-linked cohort (i.e., a local-community). Based on this assumption, the CN, JC, AA, and RA indices were extended to the CAR, CJC, CAA, and CRA indices. Subsequently, Ding et al. [41] proposed a link/community-based method that is suitable for sparse networks. The method first performs community detection based on local information only, and then scores each node pair based on the relevance between the communities the two nodes belong to. The community relevance indices, namely, CRCN, CRJC, CRAA, and CRRRA, could be regarded as community-based extensions of the CN, JC, AA, and RA indices. Other methods that enhance local similarity measures based on community structures include [40,42,44,45,47]. Other seminal studies that exploit community structures to address the sparsity issue are [43,46], which propose cluster-based meta-paths for the link-prediction problem in multiplex networks.

From the perspective of link prediction, the SLP method proposed in this study could be regarded as a link/community-based strategy that enhances local similarity measures by incorporating community structures. The motivation and novelty of the SLP approach originate from our discovery of the synergistic information between local similarity measures and community structures in the technology-convergence network of ITS. It is the existence of the synergistic information that makes the joint distribution of the two more informative than either of the marginal distributions. Most of the existing link/community-based strategies require manual adjustment of the formulae for local similarity measures. In comparison, one advantage of the SLP method is that it could be used as a plug-and-play method without the need to manually design new computational formulas. In addition, the community information that could be integrated by the SLP method is not limited to link density. Other community-based indices (e.g., CRCN, CRJC, CRAA, and CRRRA) can be directly applied to the SLP method to enhance the local similarity measures.

Moreover, the SLP approach proposed in this study is quite intuitive, lacking a careful design for the formula of the link prediction index. In particular, the sparsity of the links in real-world networks could cause serious issues for the estimation of probabilities, although this study tries to tackle this issue, at least partially, via a CDF-like formula. In the future, we would like to extend this study from the perspective of link prediction. First, we would like to examine whether the synergy between community structures and local similarity measures is widespread in real-world networks. Second, a better-designed link/community-based strategy could be proposed to make greater use of the synergistic effect. Third, extensive comparisons with different link/community-based strategies could be conducted in a broad diversity of real-world networks.

3. Methodology

This section outlines the methodology of this study, which includes four steps: (1) constructing the technology-convergence network based on IPC co-occurrence relationships obtained from patent data; (2) empirically analyzing the self-similar growth and fractality of the technology-convergence network; (3) applying the PID framework to the link-prediction problem to analyze the contributions of the original and the scaled-down networks to the anticipation of technology convergence; (4) proposing a synergistic link-prediction algorithm to predict the future technology convergence. Each of these steps is described in

detail in the following four subsections, and the research framework and main findings are shown in Figure 1.

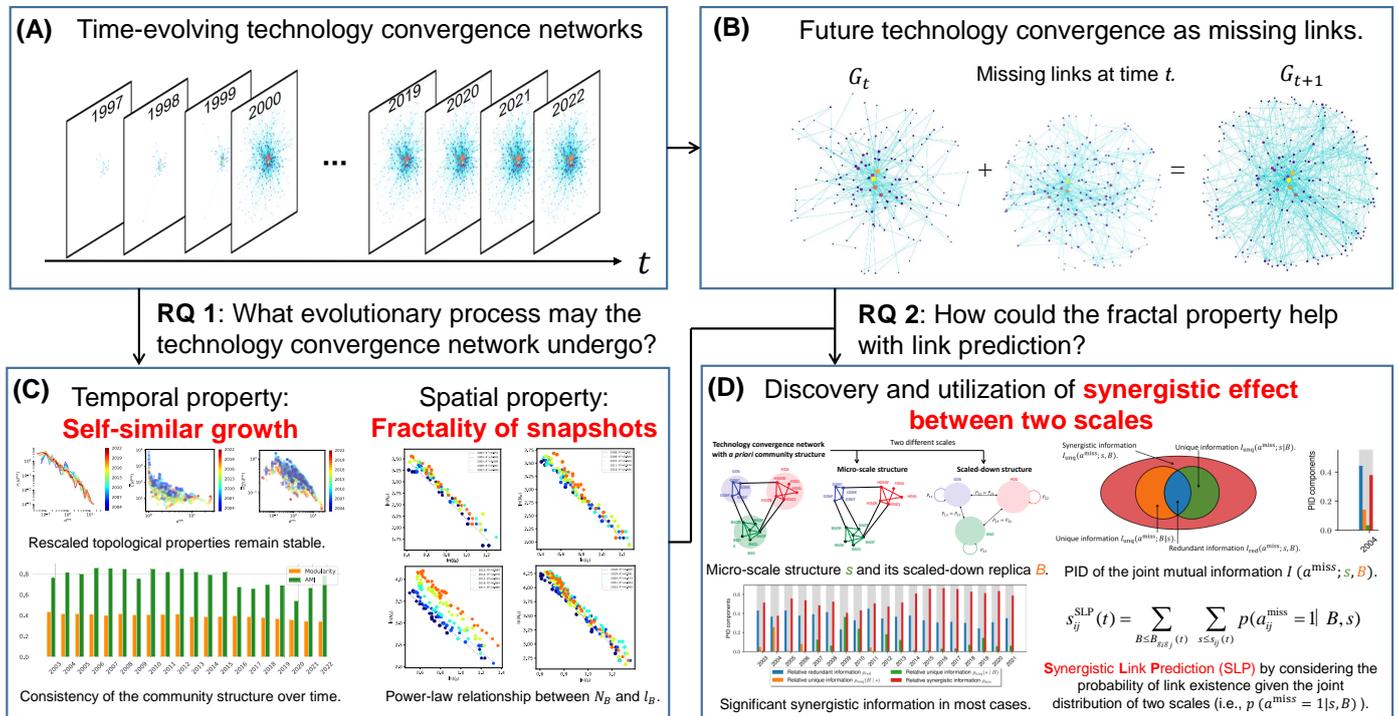


Figure 1. The research framework and main findings of this study. Sub-figure (A) shows the time-evolving technology-convergence network, in which each year’s snapshot is constructed based on patent applications up to that year; that is, the future technology convergence at time $t + 1$ could be regarded as missing links of the snapshot at time t . Sub-figure (B) shows the formulation of technology convergence anticipation as a link-prediction problem. Two main research questions of this study (i.e., RQ 1 and RQ 2) are answered in sub-figures (C,D), respectively. Sub-figure (C) presents the empirical evidence for the self-similar growth and the fractality of snapshots. Sub-figure (D) demonstrates that a synergistic effect between two scales on link prediction is discovered based on PID, and this synergistic effect is utilized via a SLP approach by considering the probability of a link’s existence being conditional on the joint distribution of two scales.

3.1. Constructing Technology-Convergence Networks

The first step in constructing a technology-convergence network for a particular industry is obtaining the relevant patents. Patent data were collected from the Derwent Innovation database, a widely adopted tool for patent analysis [13,115,116], in this study. Then, a search string was created based on the keywords of the industry, and all the query results were exported. The retrieved patent data included the title, application number, application date, IPC code, etc.

Subsequently, the annual technology-convergence network was constructed based on the collected patent data. All the IPC subclasses appearing in the patent data were first extracted as nodes in the node set in the technology-convergence network. Each IPC subclass is represented by a 4-digit IPC code (IPC4) in the patent data. An annual IPC co-occurrence network was then derived based on the patent applications filed each year. For a patent application assigned to two IPC subclasses, i and j , in year t , an IPC co-occurrence (i, j) will be recorded. For a patent with three IPC subclasses i, j , and k , three additional IPC co-occurrences are recorded as (i, j) , (j, k) , and (i, k) . By scanning all the patent applications in year t , the annual IPC-occurrence network G_t^{annl} is constructed. The adjacency matrix of G_t^{annl} is denoted as $A_t^{annl} = [a_{ij}^{annl}(t)]$, where $a_{ij}^{annl}(t) > 0$ indicates the times of co-

occurrences between i and j in year t , and $a_{ij}^{annl}(t) = 0$ indicates no co-occurrence in year t .

Finally, the historically cumulative IPC co-occurrence network G_t^{hist} is constructed with adjacency matrix defined as $A_t^{hist} = \sum_{\tau=1}^t A_{\tau}^{annl}$. From the perspective of the time-evolving network, G_t^{hist} is the snapshot of the technology-convergence network taken at the end of year t . With all the snapshots stacked in a chronological order, the time-evolving technology-convergence network could be represented by a sequence of snapshots as $\mathcal{G}^{hist} = \{G_1^{hist}, G_2^{hist}, \dots, G_T^{hist}\}$. The following analyses in this study were conducted based on \mathcal{G}^{hist} . The complete procedure for the technology-convergence network construction is depicted in Figure 2.

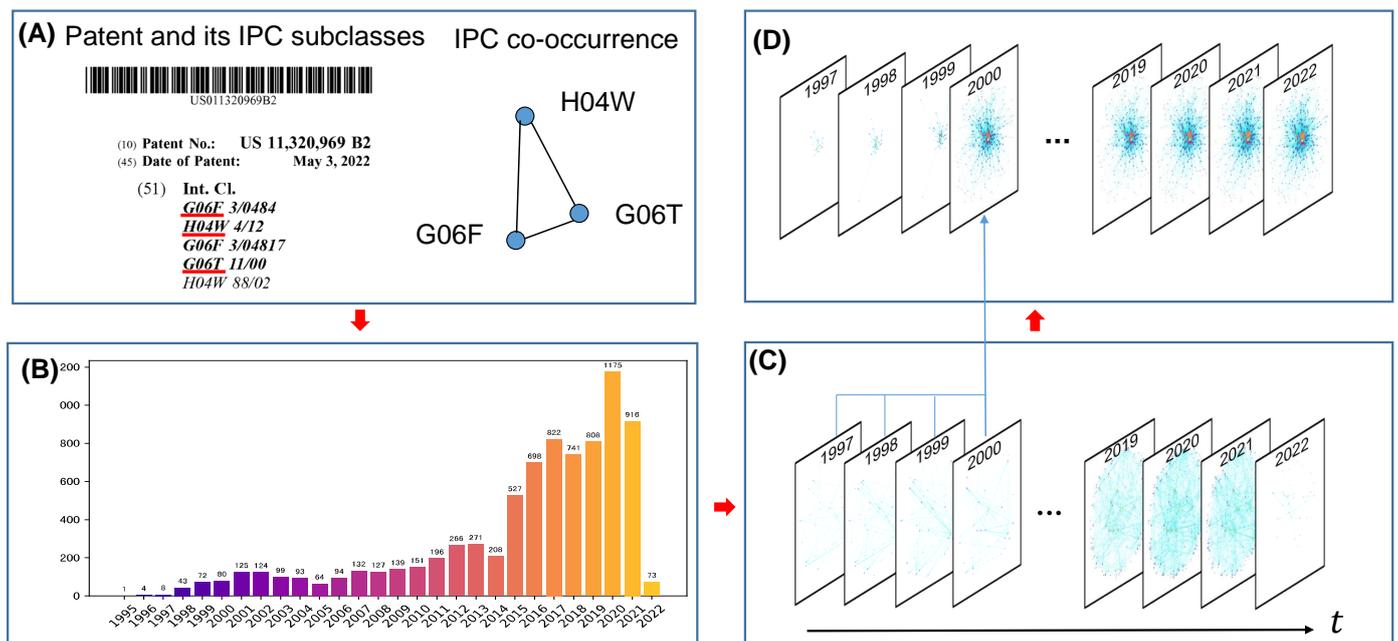


Figure 2. The complete procedure for the technology-convergence network construction. Sub-figure (A) shows an example of the patent data, to which three IPC4 codes were assigned. The three IPC4 codes are highlighted by red underlines. The patent in sub-figure (A) corresponds to three pairwise co-occurrence relationships. Sub-figure (B) shows the number of patent applications in each year. Sub-figures (C,D) show the annual and historically cumulative IPC co-occurrence networks, respectively. This study focuses on the historically cumulative IPC co-occurrence network, which is essentially a time-evolving network. The emerging IPC co-occurrence relationships could be considered as new links in the historically cumulative network.

3.2. Empirical Analysis of Self-Similar Growth and Fractality

In this subsection, the growth mechanism of the IPC co-occurrence network based on fractal analysis is examined. This study first identifies a self-similar growth process in network evolution, and then analyzes the fractality of the historical data. The self-similar growth provides the inspiration to predict the future evolution of the network based on the historical data, and the fractality of the historical data offers a motivation to explore the interaction between the original network and its scaled-down replica. The above findings inspire the PID analysis for link prediction in the next subsection.

The self-similar growth of a time-evolving network is characterized by the fact that its topological property remains in a steady state while its average degree increases moderately over time [28]. Based on the analysis framework widely adopted in the previous literature [28,33–37], this study first illustrates the evolution of the average degree in \mathcal{G}^{hist} , and then demonstrates the stability of five topological properties: (1) complementary cumulative distribution of the rescaled degree $\mathbb{P}_c(d^{res})$; (2) clustering coefficient over the

rescaled degree $\bar{C}(d^{\text{res}})$; (3) normalized average degree of neighbors over the rescaled degree $\bar{d}_{\text{nn}}^{\text{res}}(d^{\text{res}})$; (4) modularity Q of the community partitions by the Louvain method; (5) adjusted mutual information (AMI) between the community partitions of two consecutive snapshots. The detailed definitions of the average degree and the five rescaled network statistics are presented in Appendix A.

The concept of self-similarity is closely related to fractality [29,117]. In particular, Figure 6a shows that the rescaled degree of the snapshots in $\mathcal{G}^{\text{hist}}$ approximately follows a power-law distribution, which could be regarded as an evidence of fractal scaling. Therefore, this study further explores the fractality of each snapshot in $\mathcal{G}^{\text{hist}}$.

To analyze the fractality of the technology-convergence network, we employed a recently proposed community-structure-based approximation of the box-covering algorithm [38], whose detailed procedures are presented in Appendix B. The basic idea of the community-structure-based box-covering algorithm is to perform community detection η times by setting η different resolutions for the Louvain algorithm. The average community size obtained from each time of community detection is denoted as l_{B} , and the corresponding number of communities is denoted as N_{B} . Based on the sequences $\mathbf{I}_{\text{B}} = [l_{\text{B}}]$ and $\mathbf{N}_{\text{B}} = [N_{\text{B}}]$, the relationship between l_{B} and N_{B} could be investigated. If the scaling of $N_{\text{B}}(l_{\text{B}})$ over l_{B} approximately follows a power-law distribution as Equation (1), the snapshot G_t^{hist} could be identified as a fractal network, and d_{B} is referred to as its fractal dimension.

3.3. Link-Prediction-Motivated Partial Information Decomposition for Technology-Convergence Networks

The fractality of complex networks reveals the structural similarity between the original network and its scaled-down replica. This study uses PID to investigate the implication of such self-similarity, in particular, to answer the following research question: How does the microscopic structure of the original network and its scaled-down replica provide information for link prediction? Intuitively, the structural self-similarity implies the existence of redundant information between the original network and its scaled-down replica. In addition, the scaling-down process may result in information loss, allowing the original network to provide unique information for link prediction. By adopting the PID framework to the link-prediction problem, this paper quantifies such redundant, unique, and synergistic information, and analyzes the role that the scaled-down replica plays in link prediction. Figure 3 provides an overview of the link-prediction-motivated PID framework for the technology-convergence network.

In this subsection, microscopic variables represent the microscopic structure of the original network, and macroscopic variables represent the scaled-down replica. Subsequently, the link-prediction formulation is presented for the anticipation of future technology convergence and is incorporated into the PID framework. The terminology used in this subsection (e.g., microscopic variables and macroscopic variables) follows the convention in the PID-related literature [91].

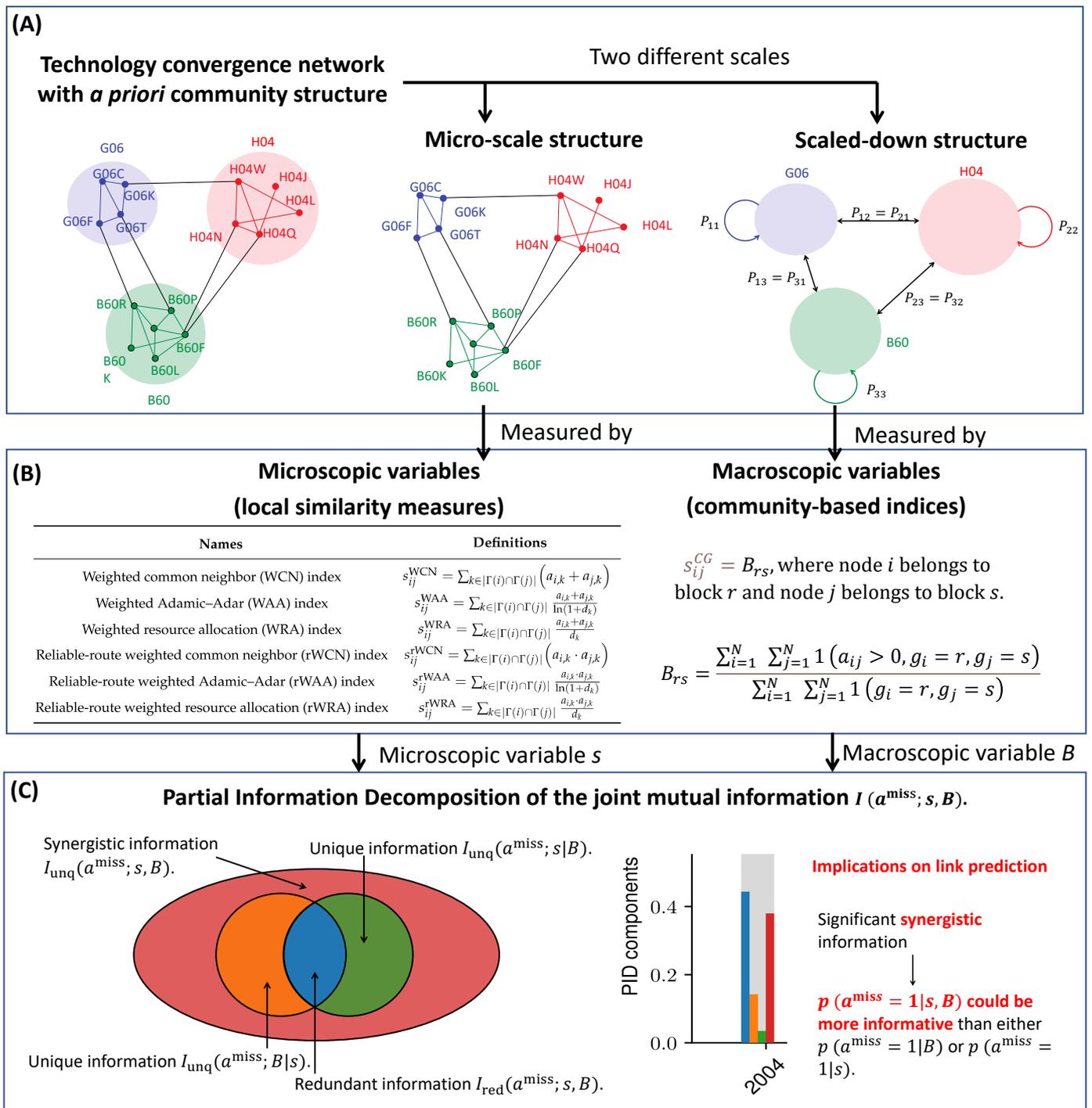


Figure 3. Overview of the link-prediction-motivated PID framework. Sub-figure (A) shows the motivation of this framework: A priori community structures in the technology-convergence network naturally provide a scaled-down replica of the original network, and we would like to explore how this multi-scale property contributes to link prediction. Sub-figure (B) shows that the micro-scale structure and its scaled-down replica are measured by local similarity measures and community-based indices, respectively. Sub-figure (C) demonstrates the PID of the joint mutual information $I(a^{miss}; s, B)$, in which significant synergistic information could imply that the joint conditional probability, given two scales, could be more informative than either scale alone. Note that, in the right image of sub-figure (C), the colors blue, red, orange and green represent the relative redundant information ρ_{red} , the relative synergistic information ρ_{syn} , and two types of relative unique information, $\rho_{unq}(B | s)$ and $\rho_{unq}(s | B)$, respectively.

3.3.1. Definitions of Microscopic Variables Based on Two-Hop Link Predictors

A key issue for the definitions of microscopic variables is to ensure that they do not contain information beyond microscopic structures in the network. This study considers the most fundamental microscopic structure, i.e., the two-hop paths in the network. From the perspective of link prediction, two-hop paths could be quantified by the scores of two-hop link predictors such as CN, AA, and RA. Therefore, microscopic variables in the network are defined based on the scores of two-hop link predictors. The detailed procedure is given as follows.

For a snapshot whose adjacency matrix is denoted as \mathbf{A} , first, a score matrix $\mathbf{S} = [s_{ij}]$ is calculated based on \mathbf{A} using a two-hop-based link predictor $f(\cdot)$, which is formally written as

$$\mathbf{S} = [s_{ij}] = f(\mathbf{A}). \tag{10}$$

Then, the score s_{ij} for a randomly chosen node pair (i, j) is considered as a random variable, and each element in the matrix \mathbf{S}^f is a sample of s_{ij} . The random variable s_{ij} is defined as the microscopic variable in the technology-convergence network.

It is worth noting that the link predictor f is restricted to be based only on two-hop paths to filter out the information provided by the paths with three hops or more. This study considers six commonly used two-hop link predictors for weighted networks, which are listed in Table 1. To simplify the expression, the time t is omitted, and the adjacency matrix is denoted as \mathbf{A} . In Table 1, $\Gamma(i)$ and $\Gamma(j)$ denote the neighbor sets of nodes i and j . The element in the i -th row and k -th column of \mathbf{A} is denoted as $a_{i,k}$. Note that $a_{i,k}$ is a positive integer rather than a binary variable, since the networks are weighted. The definition of d_k is given in Equation (A2).

Table 1. Six local similarity measures based on two-hop paths .

Names	Definitions
Weighted common neighbor (WCN) index [118]	$s_{ij}^{\text{WCN}} = \sum_{k \in \Gamma(i) \cap \Gamma(j) } (a_{i,k} + a_{j,k})$
Weighted Adamic–Adar (WAA) index [118]	$s_{ij}^{\text{WAA}} = \sum_{k \in \Gamma(i) \cap \Gamma(j) } \frac{a_{i,k} + a_{j,k}}{\ln(1 + d_k)}$
Weighted resource allocation (WRA) index [118]	$s_{ij}^{\text{WRA}} = \sum_{k \in \Gamma(i) \cap \Gamma(j) } \frac{a_{i,k} + a_{j,k}}{d_k}$
Reliable-route weighted common neighbor (rWCN) index [119]	$s_{ij}^{\text{rWCN}} = \sum_{k \in \Gamma(i) \cap \Gamma(j) } (a_{i,k} \cdot a_{j,k})$
Reliable-route weighted Adamic–Adar (rWAA) index [119]	$s_{ij}^{\text{rWAA}} = \sum_{k \in \Gamma(i) \cap \Gamma(j) } \frac{a_{i,k} \cdot a_{j,k}}{\ln(1 + d_k)}$
Reliable-route weighted resource allocation (rWRA) index [119]	$s_{ij}^{\text{rWRA}} = \sum_{k \in \Gamma(i) \cap \Gamma(j) } \frac{a_{i,k} \cdot a_{j,k}}{d_k}$

3.3.2. Definitions of Macroscopic Variables Based on Coarse-Graining

Macroscopic variables are obtained by a scaling-down (i.e., coarse-graining) procedure of the network, which consists of two parts: (1) a many-to-one mapping from the nodes of the original network to those of the scaled-down network; (2) a function that computes the properties of the scaled-down network based on the original network. The following part first discusses the reasons why existing coarse-graining methods in the fractal analysis may not be fully applicable to the technology-convergence network, then identifies the hierarchical structure of IPC as a natural way for coarse-graining the network. Finally, a coarse-graining function is defined based on the edge density of the original network.

A critical challenge for the coarse-graining procedure is the possible incomplete connection of the technology-convergence network. In fact, the evolution of the technology-convergence network is accompanied by the addition of new nodes (i.e., technology areas). If these new nodes are ignored and only the maximum connected component of a snapshot G^{hist} is considered, some important potential technology convergence may be omitted. Therefore, all the nodes for each year need to be considered to avoid such omissions. However, this may result in an incomplete connection of G^{hist} and many unconnected nodes.

Most existing box-covering algorithms have not elaborately dealt with the unconnected nodes. A conventional approach is to assign each unconnected node to a random box (e.g., Algorithm 2 in [30]). However, due to large amounts of unconnected nodes existing in the technology-convergence network (i.e., the IPC co-occurrence network), especially in the early stage of ITS development, completely random coarse-graining of unconnected nodes, which may entail much original network information loss, would lack some feasibility in this study.

The hierarchical structure of the IPC system naturally provides a criterion for coarse-graining the IPC co-occurrence network, which consists of four levels: section, class, subclass, and group. For an IPC subclass represented by a 4-digit IPC code (IPC4), the first digit and the first three digits in IPC4 represent its section and class, respectively. For example, for the IPC subclass “G06F”, “G” stands for its section “PHYSICAL”, and “G06” represents its class “COMPUTING; CALCULATING OR COUNTING”. As mentioned previously in Section 3.1, IPC subclasses are selected as nodes in the IPC co-occurrence network in this study. By grouping all the nodes (i.e., IPC subclasses) belonging to the same IPC class into a box b_r , the original network is naturally coarse-grained. The number of boxes N_B equals the number of IPC classes that appear in ITS-related patents.

From another perspective, the coarse-grained (i.e., scaled-down) network G^{CG} could be viewed as a network constructed with IPC classes as nodes. However, unlike the network constructed directly based on IPC classes, each node and edge in G^{CG} contains a property calculated by a coarse-grained function $F_{CG}(\cdot)$ based on the original network G^{hist} . To simplify the expression of macroscopic variables, a scalar variable B_{rs} is used to represent the properties of each node ($r = s$) or edge ($r \neq s$) in the coarse-grained network G^{CG} . The definition of $F_{CG}(\cdot)$ is given in the following paragraph.

The coarse-graining function $F_{CG}(\cdot)$ is a deterministic function that maps the $N \times N$ matrix \mathbf{A} to a $N_B \times N_B$ matrix. The vector $\mathbf{g} = [g_1, g_2, \dots, g_N]$ denotes the mapping relationship from nodes of the original network to nodes of the coarse-grained network, where $g_i = r$ indicates that node i in the original network is covered by box b_r , and thereby mapped to node r in the coarse-grained network. Then, $F_{CG}(\cdot)$ is defined as

$$B_{rs} = F_{CG}(\mathbf{A}, r, s) = \frac{\sum_{i=1}^N \sum_{j=1}^N \mathbb{1}(a_{ij} > 0, g_i = r, g_j = s)}{\sum_{i=1}^N \sum_{j=1}^N \mathbb{1}(g_i = r, g_j = s)}, \quad (11)$$

where $\mathbb{1}(\cdot)$ is the indicator function. The denominator and numerator of Equation (11) represent, respectively, the numbers of node pairs (i, j) and edges (i, j) in the original network with node i covered by box b_r and node j covered by box b_s .

Finally, B_{rs} is selected as the macroscopic variable, which represents the value of a randomly chosen entry (r, s) in the matrix $\mathbf{B} = [B_{rs}]$. Without loss of generality, this study assumes that all the edge weights in the original network are positive. Therefore, with Equation (11), B_{rr} represents the density of edges in box b_r of the original network, and B_{rs} is the density of edges between box b_r and b_s .

3.3.3. Problem Formulation and Partial Information Decomposition

In this study, the anticipation of technology convergence is formulated as a link-prediction problem. The potential technology convergence relationship is defined as the IPC co-occurrence which is nonexistent before time t but will appear at a future time point t' . The problem is simplified by restricting $t' = t + 1$. For each year t , the information contained in G_t^{hist} is used to predict the links in G_{t+1}^{hist} that do not exist in G_t^{hist} . These links are referred to as the missing links in year t . The set of missing links in year t is denoted as an $N \times N$ matrix $\mathbf{A}_t^{miss} = [a_{ij}^{miss}(t)]$, where $a_{ij}^{miss}(t)$ is formally defined as

$$a_{ij}^{miss}(t) = \begin{cases} 1, & \text{if } a_{ij}^{hist}(t) = 0 \text{ and } a_{ij}^{hist}(t+1) > 0, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

This study aims to analyze how the microscopic structure of the original network s_{ij} and its scaled-down replica $B_{g_i g_j}$ provide information for the prediction of future technology convergence a_{ij}^{miss} . Therefore, the PID of the joint mutual information $I(a_{ij}^{\text{miss}}; s_{ij}, B_{g_i g_j})$ is considered and given as

$$I(a_{ij}^{\text{miss}}; s_{ij}, B_{g_i g_j}) = I_{\text{red}}(a_{ij}^{\text{miss}}; s_{ij}, B_{g_i g_j}) + I_{\text{unq}}(a_{ij}^{\text{miss}}; s_{ij} | B_{g_i g_j}) + I_{\text{unq}}(a_{ij}^{\text{miss}}; B_{g_i g_j} | s_{ij}) + I_{\text{syn}}(a_{ij}^{\text{miss}}; s_{ij}, B_{g_i g_j}), \quad (13)$$

where I_{red} , I_{unq} , and I_{syn} are the redundant information, unique information, and synergistic information, respectively.

First, the redundant information, I_{red} , is calculated based on Equation (4) as

$$I_{\text{red}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}) = p(a_{ij}^{\text{miss}} = 1) \times \min(I(a_{ij}^{\text{miss}} = 1, B_{g_i g_j}), I(a_{ij}^{\text{miss}} = 1, s_{ij})) + p(a_{ij}^{\text{miss}} = 0) \times \min(I(a_{ij}^{\text{miss}} = 0, B_{g_i g_j}), I(a_{ij}^{\text{miss}} = 0, s_{ij})), \quad (14)$$

where $\min(\cdot)$ chooses the smaller of the two types of mutual information, and $I(a_{ij}^{\text{miss}} = 1, B_{g_i g_j})$ is defined as

$$I(a_{ij}^{\text{miss}} = 1, B_{g_i g_j}) = \sum_{B_{g_i g_j}} p(B_{g_i g_j} | a_{ij}^{\text{miss}} = 1) \log_2 \frac{p(a_{ij}^{\text{miss}} = 1 | B_{g_i g_j})}{p(a_{ij}^{\text{miss}} = 1)}, \quad (15)$$

(see Equations (1)–(3) in [83]). The other three types of mutual information in (14) follow similar definitions as (15).

The unique information provided by $B_{g_i g_j}$ beyond s_{ij} is given as

$$I_{\text{unq}}(a_{ij}^{\text{miss}}; B_{g_i g_j} | s_{ij}) = I(a_{ij}^{\text{miss}}; B_{g_i g_j}) - I_{\text{red}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}), \quad (16)$$

where $I(a_{ij}^{\text{miss}}; B_{g_i g_j})$ denotes the mutual information of a_{ij}^{miss} and $B_{g_i g_j}$,

$$I(a_{ij}^{\text{miss}}; B_{g_i g_j}) = \sum_{B_{g_i g_j}} p(a_{ij}^{\text{miss}} = 1, B_{g_i g_j}) \log \frac{p(a_{ij}^{\text{miss}} = 1, B_{g_i g_j})}{p(a_{ij}^{\text{miss}} = 1)p(B_{g_i g_j})} + \sum_{B_{g_i g_j}} p(a_{ij}^{\text{miss}} = 0, B_{g_i g_j}) \log \frac{p(a_{ij}^{\text{miss}} = 0, B_{g_i g_j})}{p(a_{ij}^{\text{miss}} = 0)p(B_{g_i g_j})}. \quad (17)$$

Similarly, the unique information provided by s_{ij} beyond $B_{g_i g_j}$ is given as

$$I_{\text{unq}}(a_{ij}^{\text{miss}}; s_{ij} | B_{g_i g_j}) = I(a_{ij}^{\text{miss}}; s_{ij}) - I_{\text{red}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}). \quad (18)$$

Furthermore, the synergistic information is calculated as

$$I_{\text{syn}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}) = I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}) - I(a_{ij}^{\text{miss}}; s_{ij}) - I_{\text{unq}}(a_{ij}^{\text{miss}}; B_{g_i g_j} | s_{ij}), \quad (19)$$

which quantifies the joint contribution of $B_{g_i g_j}$ and s_{ij} to link prediction, and can be further formulated as

$$I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}) = \sum_{B_{g_i g_j}} \sum_{s_{ij}} p(a_{ij}^{\text{miss}} = 1, B_{g_i g_j}, s_{ij}) \log \frac{p(a_{ij}^{\text{miss}} = 1 | B_{g_i g_j}, s_{ij})}{p(a_{ij}^{\text{miss}} = 1)}$$

$$+ \sum_{B_{g_i g_j}} \sum_{s_{ij}} p(a_{ij}^{\text{miss}} = 0, B_{g_i g_j}, s_{ij}) \log \frac{p(a_{ij}^{\text{miss}} = 0 | B_{g_i g_j}, s_{ij})}{p(a_{ij}^{\text{miss}} = 0)}. \quad (20)$$

Since the missing links are rather sparse at each time point t , $p(a_{ij}^{\text{miss}} = 1)$ will be very small, and $p(a_{ij}^{\text{miss}} = 0)$ will be close to one. As a result, the value of mutual information will be quite small. To highlight the relative contribution of each PID component, this study proposes three relative PID measures, namely, the relative redundant information, the relative unique information, and the relative synergistic information, which are defined as the corresponding PID component normalized by the mutual information $I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})$. The formal definitions of the relative PID measures are given as

$$\rho_{\text{red}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}) = \frac{I_{\text{red}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})}{I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})}, \quad (21)$$

$$\rho_{\text{unq}}(a_{ij}^{\text{miss}}; B_{g_i g_j} | s_{ij}) = \frac{I_{\text{unq}}(a_{ij}^{\text{miss}}; B_{g_i g_j} | s_{ij})}{I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})}, \quad (22)$$

$$\rho_{\text{unq}}(a_{ij}^{\text{miss}}; s_{ij} | B_{g_i g_j}) = \frac{I_{\text{unq}}(a_{ij}^{\text{miss}}; s_{ij} | B_{g_i g_j})}{I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})}, \quad (23)$$

$$\rho_{\text{syn}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}) = \frac{I_{\text{syn}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})}{I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})}. \quad (24)$$

For convenience of following expression, the above four measures will be simplified to ρ_{red} , $\rho_{\text{unq}}(B | s)$, $\rho_{\text{unq}}(s | B)$, and $\rho_{\text{syn}}(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})$, respectively. Notice that $\rho_{\text{red}} + \rho_{\text{unq}}(B | s) + \rho_{\text{unq}}(s | B) + \rho_{\text{syn}} = 1$.

As the final essential issue in PID, the estimate of mutual information will be discussed. The mutual information is estimated using the most direct binning-based approach in this study. $\mathcal{B} = [\min(\mathbf{B}), \max(\mathbf{B})]$ and $\mathcal{S} = [\min(\mathbf{S}), \max(\mathbf{S})]$, which denote the ranges of values of B_{ij} and s_{ij} , are divided into sets of bins of finite size. All the probabilities used for the calculation of mutual information are estimated by counting the number of samples falling into the bins [120,121]. $n_B^+(x)$ and $n_B^-(x)$ represent the numbers of positive samples (i.e., node pairs (i, j) with $a_{ij}^{\text{miss}} = 1$) and negative samples (i.e., node pairs (i, j) with $a_{ij}^{\text{miss}} = 0$) falling into the x -th bin of \mathcal{B} . Since there are $\frac{N(N-1)}{2}$ node pairs in an undirected network,

$$p(a^{\text{miss}} = 1, B_{g_i g_j}) = \frac{2 \cdot n_B^+(x)}{N(N-1)}, \quad (25)$$

$$p(B_{g_i g_j}) = \frac{2 \cdot (n_B^+(x) + n_B^-(x))}{N(N-1)}, \quad (26)$$

where the value of $B_{g_i g_j}$ falls into the x -th bin of \mathcal{B} —i.e., $B_{g_i g_j} \in \text{bin}_x$.

We could further calculate $p(B_{g_i g_j} | a^{\text{miss}} = 1)$ and $p(a^{\text{miss}} = 1 | B_{g_i g_j})$ based on Equations (25) and (26). Note that $p(a^{\text{miss}} = 1 | B_{g_i g_j})$ is set to 0 when its denominator is zero. Other distributions, such as $p(s_{ij})$ and $p(s_{ij}, a_{ij}^{\text{miss}} = 0)$, could be estimated in the same way, so their detailed expressions are omitted.

The joint distribution is estimated by counting the intersection of samples in the x -th bin of \mathcal{B} and the y -th bin of \mathcal{S} . The number of positive samples falling into the intersection of the bins x and y is denoted as $n_{B,s}^+(x, y)$, which is given as

$$n_{B,s}^+(x, y) = \sum_{i=1}^N \sum_{j=1}^N \mathbb{1}(a_{ij}^{\text{miss}} > 0, B_{g_i g_j} \in \text{bin}_x, s_{ij} \in \text{bin}_y). \quad (27)$$

The number of negative samples in the intersection of x and y is denoted as $n_{B,s}^-(x, y)$. The joint probability $p(a_{ij}^{\text{miss}}, B_{g_i g_j}, s_{ij})$ is estimated as

$$p(a_{ij}^{\text{miss}} = 1, B_{g_i g_j}, s_{ij}) = \frac{2 \cdot n_{B,s}^+(x, y)}{N(N-1)}, \quad (28)$$

$$p(a_{ij}^{\text{miss}} = 0, B_{g_i g_j}, s_{ij}) = \frac{2 \cdot n_{B,s}^-(x, y)}{N(N-1)}. \quad (29)$$

By plugging the aforementioned distributions into Equations (15), (17) and (20), the mutual information of $I(a_{ij}^{\text{miss}} = 1, B_{g_i g_j})$, $I(a_{ij}^{\text{miss}}; B_{g_i g_j})$, and $I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})$ is obtained. Other mutual information, such as $I(a_{ij}^{\text{miss}} = 0, B_{g_i g_j})$ and $I(a_{ij}^{\text{miss}}; s_{ij})$, can be obtained using a similar approach. Based on the mutual information estimated above, the relative PID measures defined in this research, ρ_{red} , $\rho_{\text{unq}}(B | s)$, $\rho_{\text{unq}}(s | B)$, and ρ_{syn} , can finally be calculated.

3.4. Synergistic Link-Prediction Approach for the Anticipation of Technology Convergence

Based on the PID of $I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})$ in Equation (13), it can be observed that if $B_{g_i g_j}$ could provide unique information beyond s_{ij} or synergistic information in conjunction with s_{ij} , we have

$$I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij}) > I(a_{ij}^{\text{miss}}; s_{ij}), \quad (30)$$

which implies that considering both $B_{g_i g_j}$ and s_{ij} may achieve better performance than considering s_{ij} alone. In this study, experimental results show that the contribution of $B_{g_i g_j}$ mainly takes the form of the synergistic information. Therefore, a link-prediction algorithm is proposed to combine the information provided by $B_{g_i g_j}$ and s_{ij} , which is named the synergistic link prediction (SLP) approach.

Consider a snapshot G_t^{hist} at time t . The mutual information could be written as $I(a_{ij}^{\text{miss}}(t); B_{g_i g_j}(t), s_{ij}(t))$. From Equation (20), we discover that $p(a_{ij}^{\text{miss}}(t) = 1 | B_{g_i g_j}(t), s_{ij}(t))$ naturally provides a link-prediction score for each node pair. For a node pair (i, j) , a higher $p(a_{ij}^{\text{miss}}(t) = 1 | B_{g_i g_j}(t), s_{ij}(t))$ means that a missing link is more likely to exist.

However, it is worth noting that $a_{ij}^{\text{miss}}(t)$ is unknown at time t , since $a_{ij}^{\text{hist}}(t+1)$ can only be observed at time $t+1$ according to the definition in Equation (12). Therefore, the joint probability $p(a_{ij}^{\text{miss}} = 1 | B_{g_i g_j}, s_{ij})$ cannot be estimated directly at time t .

To solve this problem, we introduce the historically cumulative missing links, which is an $N \times N$ matrix $\mathbf{A}_t^{\text{train}} = [a_{ij}^{\text{train}}(t)]$ defined as

$$a_{ij}^{\text{train}}(t) = \begin{cases} 1, & \text{if } \sum_{\tau=1}^{t-1} a_{ij}^{\text{miss}}(\tau) > 0, \\ 0, & \text{otherwise.} \end{cases} \quad (31)$$

The matrix $\mathbf{A}_t^{\text{train}}$ is the aggregation of all the historical missing links $\mathbf{A}_\tau^{\text{miss}}$ from $\tau = 1$ to $\tau = t-1$. Since $\mathbf{A}_t^{\text{train}}$ does not contain any future information, it could be used as our training labels. Based on the binning-approach described in the previous subsection, $p(a_{ij}^{\text{train}}(t) = 1 | B_{g_i g_j}(t), s_{ij}(t))$ is estimated as

$$p(a_{ij}^{\text{train}} = 1 | B_{g_i g_j}(t), s_{ij}(t)) = \frac{\sum_{i=1}^N \sum_{j=1}^N \mathbb{1}(a_{ij}^{\text{train}}(t) = 1, B_{g_i g_j}(t) \in \text{bin}_x, s_{ij}(t) \in \text{bin}_y)}{\sum_{i=1}^N \sum_{j=1}^N \mathbb{1}(B_{g_i g_j}(t) \in \text{bin}_x, s_{ij}(t) \in \text{bin}_y)}. \quad (32)$$

We set $p(a_{ij}^{\text{train}} = 1 | B_{g_i g_j}(t), s_{ij}(t)) = 0$ if the denominator is zero, which indicates that no sample falls into the intersection of the bin x and y .

It seems that an $N \times N$ score matrix could be constructed based on $p(a_{ij}^{\text{train}} = 1 | B_{g_i g_j}(t), s_{ij}(t))$ to quantify the likelihood for the existence of edges between node pairs (i, j) . However, such a direct approach may not perform well in practice, since the score matrix may be very sparse. In fact, for two node pairs (i_1, j_1) and (i_2, j_2) , if $B_{g_{i_1} g_{j_1}}(t) > B_{g_{i_2} g_{j_2}}(t)$ and $s_{i_1 j_1}(t) > s_{i_2 j_2}(t)$, the node pair (i_1, j_1) is more likely to be connected than the node pair (i_2, j_2) . Such a property is not reflected in Equation (32).

To compensate for the limitation of $p(a_{ij}^{\text{train}} = 1 | B_{g_i g_j}(t), s_{ij}(t))$, this study finally proposes the precise definition of the synergistic link-prediction score as

$$s_{ij}^{\text{SLP}}(t) = \sum_{B \leq B_{g_i g_j}(t)} \sum_{s \leq s_{ij}(t)} p(a_{ij}^{\text{train}} = 1 | B, s), \quad (33)$$

which could be regarded as the cumulative distribution function (CDF) of $p(a_{ij}^{\text{train}} = 1 | B_{g_i g_j}(t), s_{ij}(t))$.

4. Experiments

We chose ITS as the technological area for experimental research for three reasons. First, ITS is one of the most cutting-edge branches of transportation, providing innovative and sustainable solutions to notable transportation problems such as traffic congestion, traffic accidents, and high maintenance costs [122]. Second, ITS is an interdisciplinary technology field that involves sensors, communications, algorithms, vehicles, traditional transportation infrastructure, etc. Third, to the best of our knowledge, little work has been done to analyze the ITS development from a technology convergence perspective. This section describes our experiments based on the ITS-related patents, including data description, empirical evidence on self-similar evolution and fractality, PID analysis of the technology-convergence network, and performance evaluation of our proposed synergistic link prediction method.

4.1. Data Description

In this study, the patent data were collected from the Derwent Innovation database, one of the most widely used databases for patent analyses [5,13,116,123–125]. The patent data were selected from all the authorities worldwide on the date of 1 July 2022 and searched using the following query: *TAB = (intelligent ADJ transportation)*. This query retrieves patents with the phrase intelligent transportation in the title or abstract.

It is worth noting that the choice of keywords and query rules is a trade-off between accuracy and comprehensiveness. Overly broad query rules may lead to many completely irrelevant patents in the search results. For instance, the topic of a patent with the term “smart traffic” in the title could be “smart traffic scheduling in wireless networks”. On the contrary, excessively rigorous query rules may result in the absence of some relevant patents. There is no clear criterion for this trade-off, since it is rather challenging to strictly define which patents are “relevant” to ITS. This research applied strict search criteria to ensure high accuracy and avoid including utterly irrelevant patents.

The data collected cover 8057 ITS-related patent applications for the period from November 1995 to July 2022. Each patent’s data contain a title, application number, application date, International Patent Classification (IPC), etc. Figure 4 shows the number of patent applications in each year from 1995 to 2022. It can be observed that the number of

patent applications is generally increasing annually. Note that the numbers of applications in 2021 and 2022 may not be accurate, since there is a delay between patent application and patent publication, and some patent applications in 2021 and 2022 have not been made publicly available yet.

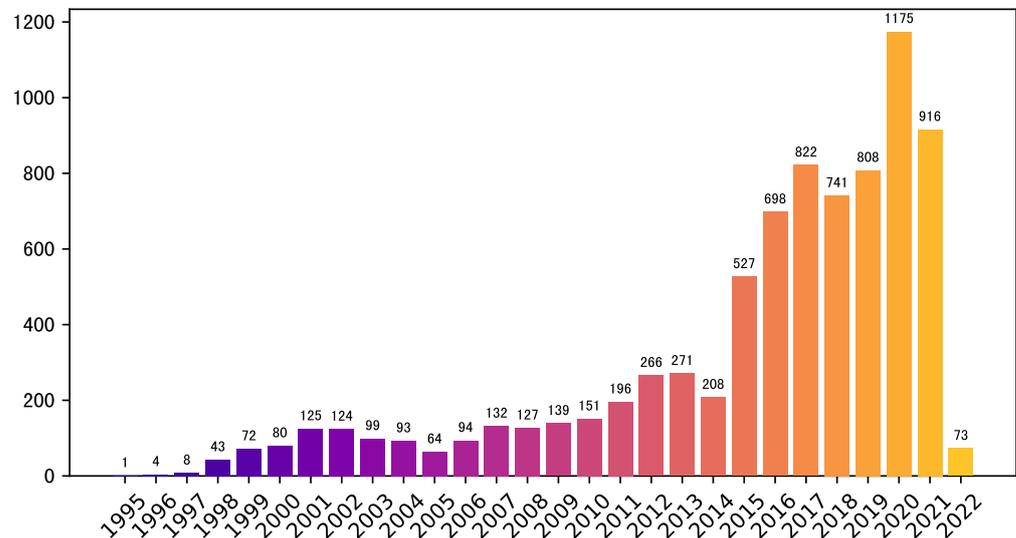


Figure 4. Number of patent applications per year. The number of patent applications has generally been increasing. The apparent decline in 2021 and 2022 could have been due to the delays from the application dates to the disclosure dates of the patents' applications.

By scanning all 8057 patents, 306 related technology fields were identified (i.e., IPC subclasses represented by IPC4). Among the 8057 patents, 3763 patents contain two or more IPC subclass codes, based on which the IPC co-occurrence network was constructed.

4.2. Empirical Evidence of the Self-Similar Growth and Fractality of the Technology-Convergence Network

This subsection presents the empirical evidence demonstrating the self-similar growth and fractality of the technology-convergence network. According to the existing literature [28,33–37], the main criterion for the self-similar growth of a time-evolving network is that its size increases moderately over time while its topological property remains in a steady state. In terms of the growth of network sizes, Figure 5 shows that the number of nodes and the average weighted degree of the historically cumulative IPC co-occurrence network $\mathcal{G}^{\text{hist}}$ moderately increase over time. In terms of the stability of topological properties, Figures 6 and 7 show that the rescaled global (i.e., degree distributions), local (i.e., clustering coefficients and degree of neighbors), and meso-scale (i.e., community structures) topological properties of $\mathcal{G}^{\text{hist}}$ remain in a steady state. Figure 8 demonstrates the fractality of the snapshots of $\mathcal{G}^{\text{hist}}$.

Since the IPC co-occurrence network may not be fully connected, Figure 5a presents the evolution of the node number in the maximum connected component per year, denoted as N_{LCC} . In addition, since this study models the IPC co-occurrence network as a weighted graph, Figure 5b illustrates the evolution of the average weighted degree \bar{d} in each snapshot.

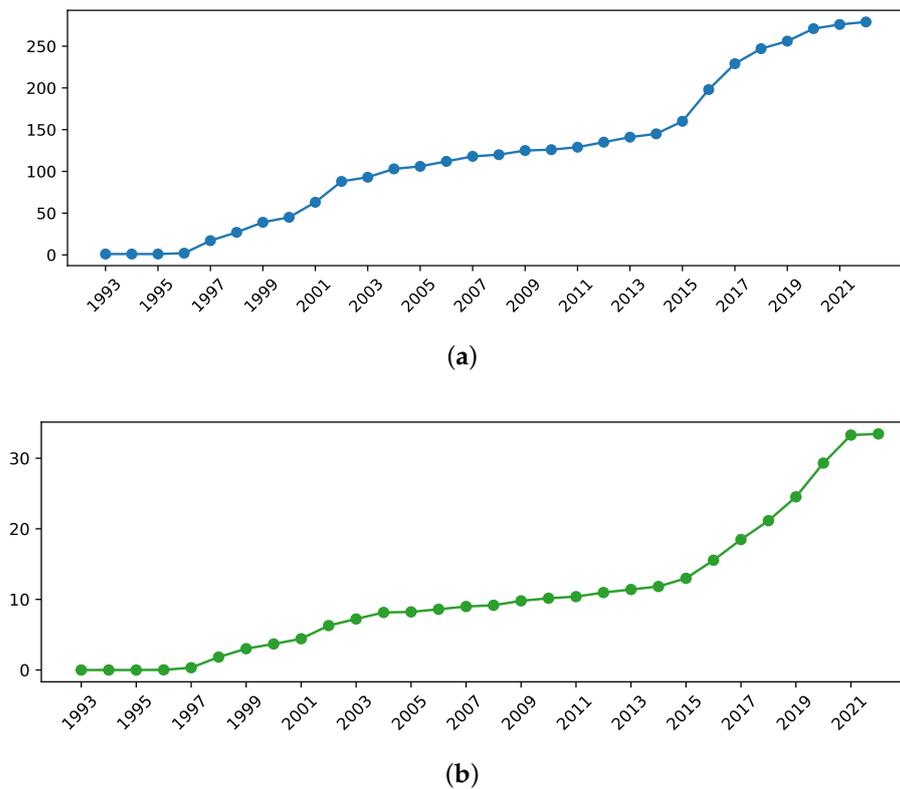


Figure 5. Evolution of network statistics of the technology-convergence network. It can be observed that both the number of nodes and the average degree moderately increase in most years, which matches with the characteristic of self-similar growth, in that the network size moderately grows over time. (a) Number of nodes in the largest connected component of $\mathcal{G}^{\text{hist}}$. (b) Average weighted degree of $\mathcal{G}^{\text{hist}}$.

It can be observed that N_{LCC} and \bar{d} increase over time, and both showed high growth rates in 1997–2003 and 2015–2019. During other years, N_{LCC} and \bar{d} grew relatively slowly and even exhibited flat behavior. This observation is consistent with the history of ITS. The study of ITS dates back to the 1970s, during which, the United States, Europe, and Japan were pioneers [122,126]. Some early representative achievements appeared in the 1990s, such as DEMO'97 in the United States [127] and the Vehicle Information and Communication System (VICS) in Japan [128]. Analogously, ITS-related patent applications began in the 1990s and increased significantly in 1997 and 1998, coinciding with the emergence of DEMO'97 and VICS. The remarkable growth of N_{LCC} and \bar{d} around 2015 matches the development of connected vehicles and autonomous vehicles [129]. Since snapshots before 2003 are relatively sparse and their growth became stable after 2003, we examined snapshots from 2003 and after.

Figure 6a–c present the complementary cumulative distributions of d^{res} , clustering coefficients over the rescaled degree $\bar{C}(d^{\text{res}})$, and normalized average degree of neighbors over the rescaled degree $\bar{d}_{\text{nn}}^{\text{res}}(d^{\text{res}})$, respectively. Each color in Figure 6 represents a snapshot. The curve overlaps of different snapshots demonstrate the self-similar growth of the IPC co-occurrence network. In addition, Figure 6a shows that the distribution of the rescaled degree d^{res} approximates a power-law distribution, which could be regarded as an evidence for the fractality of the snapshots.

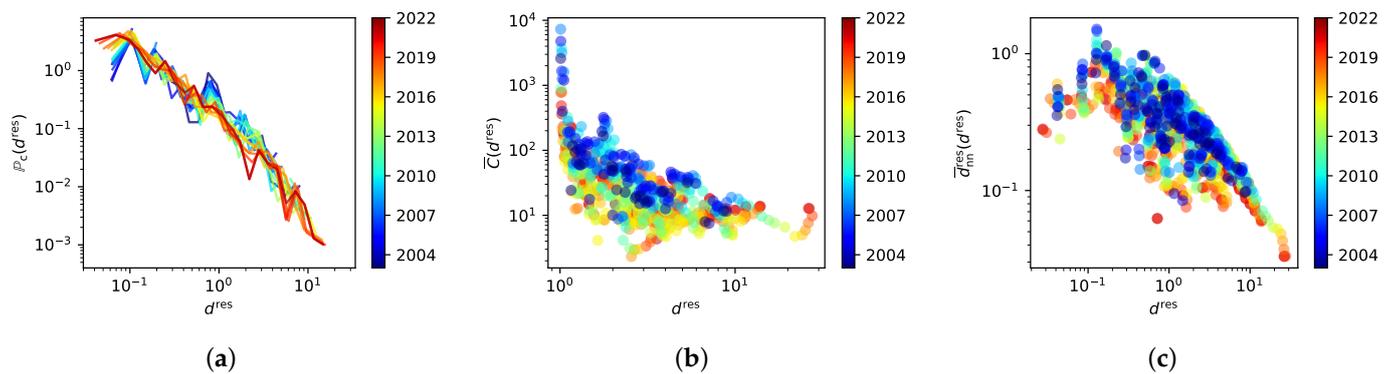


Figure 6. Self-similar evolution of the technology-convergence network. (a) The complementary cumulative distributions of d^{res} . (b) The clustering coefficients over the rescaled degree $\bar{C}(d^{\text{res}})$. (c) The normalized average degree of neighbors over the rescaled degree $\bar{d}_{\text{nn}}^{\text{res}}(d^{\text{res}})$. Each color corresponds to a different year. The curve overlaps demonstrate that the topological properties remain stable over time, which reflects the self-similarity characteristic of the growth process.

Figure 7 depicts the self-similarity of network evolution from the perspective of community structure. In the figure, the orange and green bars represent the modularity Q of community partitions in each snapshot and the AMI between the community partitions in the current and prior snapshots. The Louvain method [130] was used to obtain the community partitions. We can observe that the modularity Q remained stable over time, and the community partitions of two consecutive snapshots are highly overlapping. Such observations indicate that the community structure of the network does not change drastically during the evolution of the network.

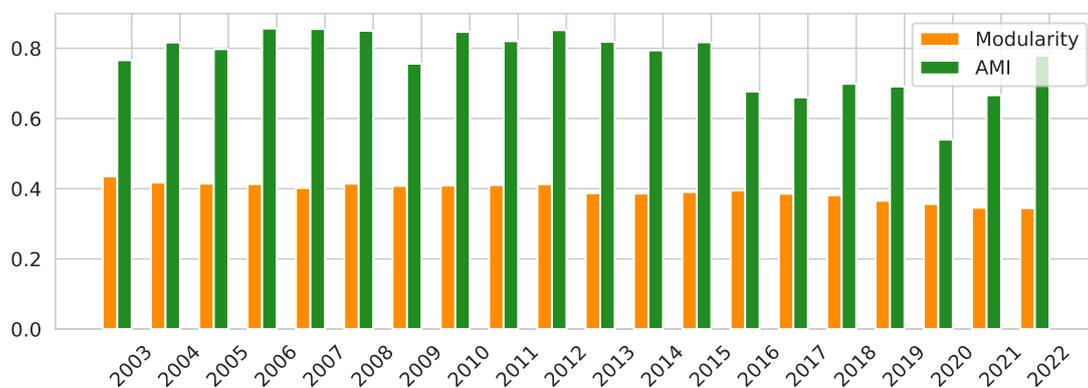


Figure 7. Consistency of the community structure during the evolution of the technology-convergence network. It can be observed that the modularity remains stable and the AMI between two consecutive snapshots is relatively high, which matches with the characteristic of self-similar growth: that the community structure of the network remains stable.

Finally, Figure 8 provides evidence for the fractality of snapshots in $\mathcal{G}^{\text{hist}}$, where each subplot corresponds to five snapshots. For each snapshot, the number of boxes N_B and the average size of boxes l_B obtained by Algorithm A1 are plotted in a logarithmic plane. Clear-cut linear relationships can be observed between $\ln(N_B)$ and $\ln(l_B)$. We further confirmed the linear relationship by performing linear regressions between $\ln(N_B)$ and $\ln(l_B)$ for each snapshot. The coefficient of determination, R^2 , of the linear regression of each snapshot is presented in Figure 8. With $R^2 > 0.95$ holding for all the snapshots, the power-law relationship between l_B and N_B is further confirmed, indicating the fractality of snapshots in the time-evolving technology-convergence network.

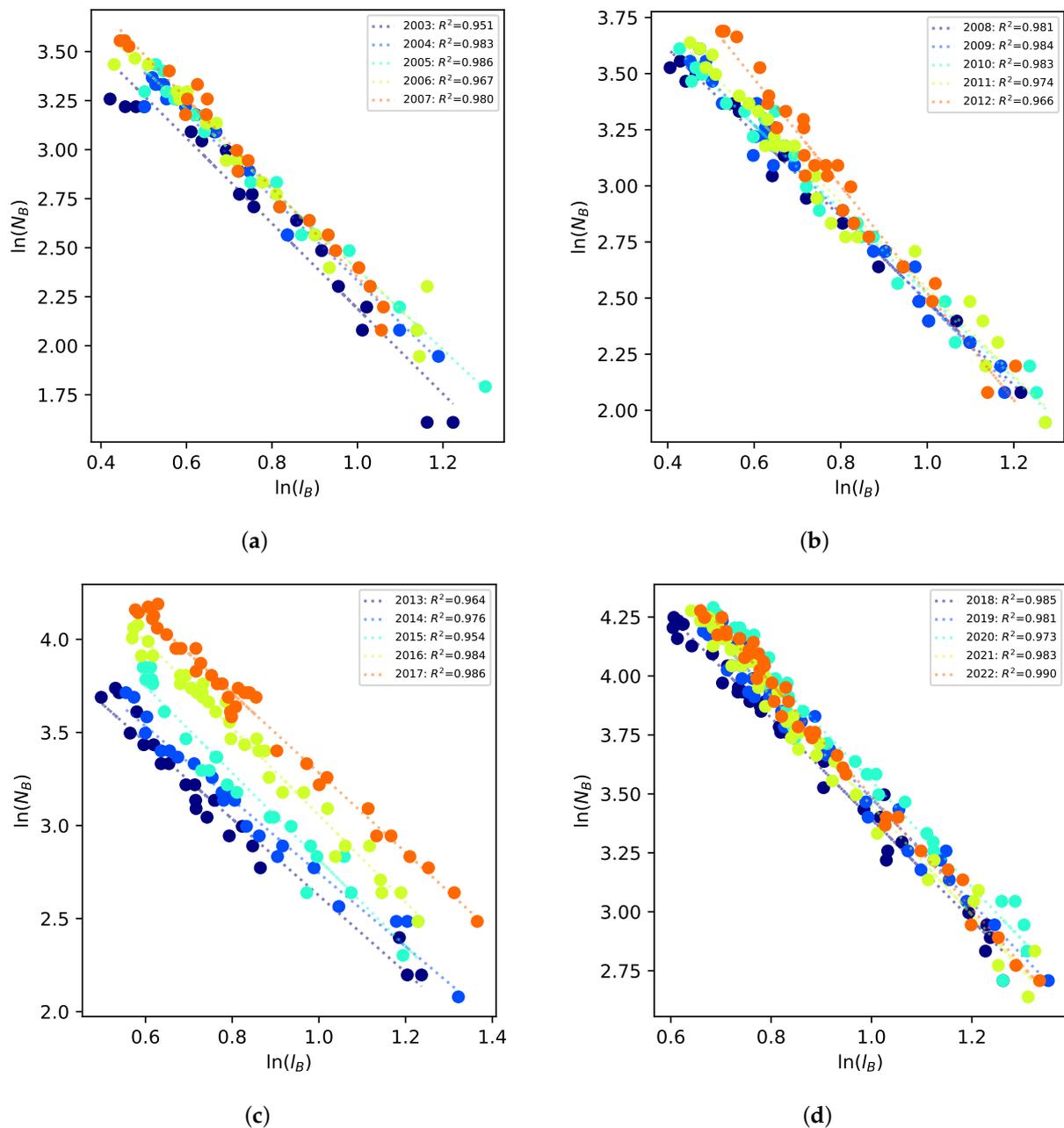


Figure 8. Power-law relationships between the number of boxes N_B and the average size l_B . Clear-cut linear relationships can be observed between $\ln(N_B)$ and $\ln(l_B)$, which reveal the fractality of the snapshots. (a) Snapshots of 2003–2007. (b) Snapshots of 2008–2012. (c) Snapshots of 2013–2017. (d) Snapshots of 2018–2022.

4.3. Partial Information Decomposition of the Technology–Convergence Network

In this subsection, we investigate how the microscopic structure of the original technology-convergence network and its scaled-down replica provide information for link prediction. The main results are shown in Figure 9. It can be observed that the synergistic information (represented by the red bars for each year) is significant in most cases, which implies that the joint distribution of the two scales could be more informative than the marginal distribution of each individual scale.

For a randomly chosen node pair (i, j) , the microscopic structure is represented by the microscopic variable s_{ij} , and the scaled-down replica is represented by the macroscopic

variable $B_{g_i g_j}$. It is worth noting that the calculation of s_{ij} depends on the specific two-hop link predictors. In this study, six two-hop link predictors, as shown in Table 1, were considered. For Figure 9a–f, the microscopic variables s_{ij} were calculated by the link predictors WCN, rWCN, WAA, rWAA, WRA, and rWRA, respectively.

Figure 9 shows the relative PID measures (i.e., ρ_{red} , $\rho_{\text{unq}}(B | s)$, $\rho_{\text{unq}}(s | B)$, and ρ_{syn}) of the joint mutual information $I(a_{ij}^{\text{miss}}; B_{g_i g_j}, s_{ij})$, which are colored in blue, orange, green, and red, respectively. The four relative PID measures belonging to the same year are marked by the gray regions. It can be observed that the interplay between s_{ij} and $B_{g_i g_j}$ is primarily represented by redundant and synergistic information, whereas s_{ij} could provide unique information beyond $B_{g_i g_j}$. In most circumstances, $B_{g_i g_j}$ cannot provide unique information beyond s_{ij} .

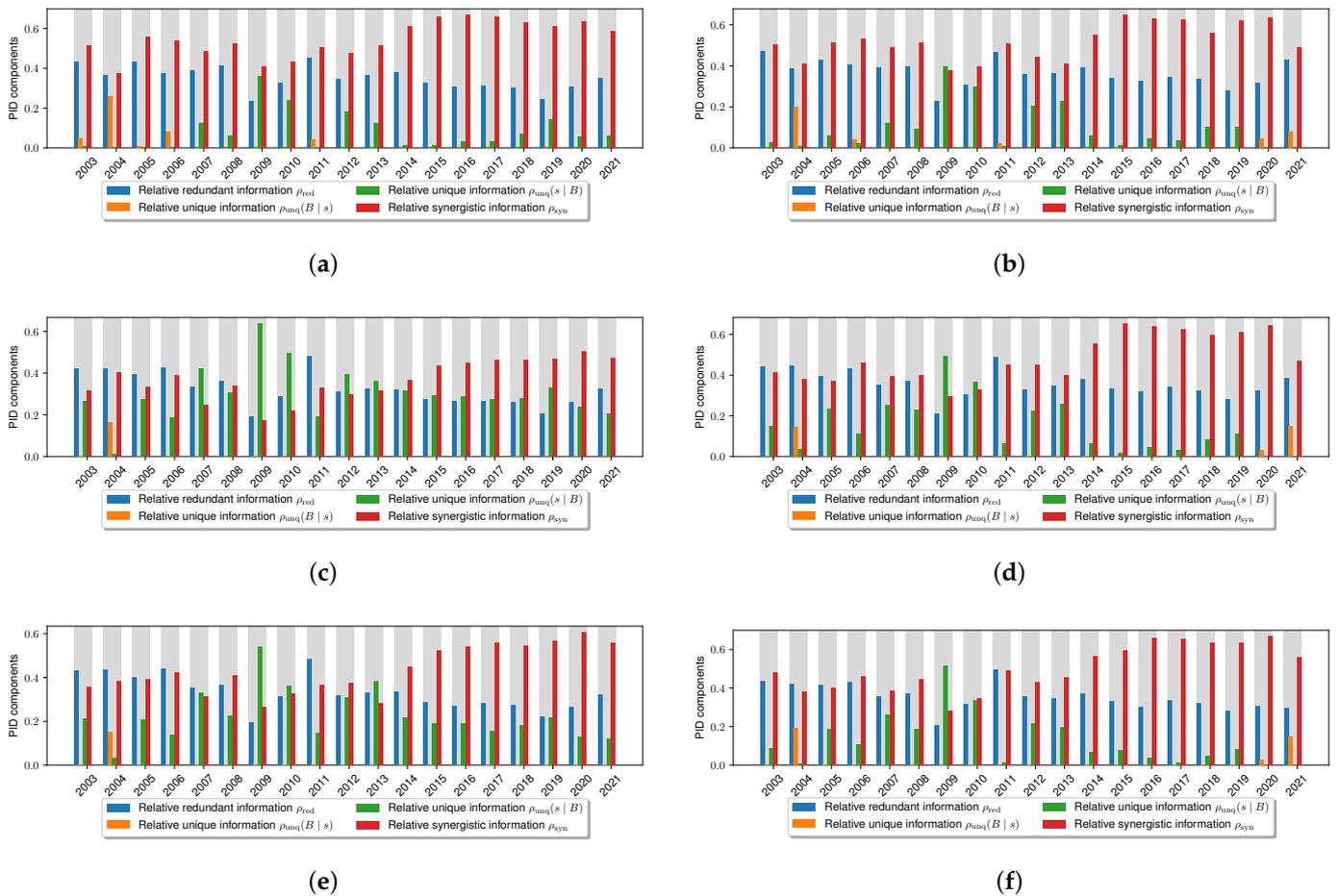


Figure 9. PID decomposition components with different two-hop link predictors. Shown are the relative PID measures (see legend) of the joint mutual information $I(a^{\text{miss}}; B, s)$, namely, the relative redundant information ρ_{red} , the relative synergistic information ρ_{syn} , and two types of relative unique information, $\rho_{\text{unq}}(B | s)$ and $\rho_{\text{unq}}(s | B)$. The microscopic variable s for each sub-figure was calculated based on a type of two-hop link predictor, and the macroscopic variable B for all the sub-figures was calculated in the same way (i.e., the link density of communities). It can be observed that, in most years, especially in the middle and late stages of the development of ITS, the relative synergistic information ρ_{syn} plays a dominant role. Such observation implies that the joint conditional distribution $p(a^{\text{miss}} | B, s)$ could be more informative than any of the two marginal conditional distributions, $p(a^{\text{miss}} | B)$ and $p(a^{\text{miss}} | s)$. (a) Two-hop link predictor: WCN. (b) Two-hop link predictor: rWCN. (c) Two-hop link predictor: WAA. (d) Two-hop link predictor: rWAA. (e) Two-hop link predictor: WRA. (f) Two-hop link predictor: rWRA.

4.4. Performance Evaluation of Synergistic Link Prediction

In this last set of experiments, we evaluated the performance of the SLP algorithm. We implemented six synergy-based link prediction indices by incorporating a community-based index into six local similarity measures via the SLP algorithm. Pairwise comparisons were conducted between the synergy-based index and its corresponding local similarity measure, and the results are shown in Table 2. It is observed that the synergy-based indices outperform both the community-based index and the local similarity measures in most cases, which could validate the existence and usefulness of the synergistic effect between two scales on link prediction.

The evaluation metric was the area under the receiver operating characteristic curve (AUC), which is widely adopted for the evaluation of link-prediction algorithms [20,25,111,131]. The AUC for the link predictor f in the snapshot of year t is calculated as

$$\begin{aligned}
 \text{AUC}(f, t) = & \mathbb{P}\left(s_{i_1, j_1}^f(t) > s_{i_2, j_2}^f(t) \mid a_{i_1, j_1}^{\text{miss}}(t) = 1, a_{i_2, j_2}^{\text{miss}}(t) = 0\right) \\
 & + \frac{1}{2} \mathbb{P}\left(s_{i_1, j_1}^f(t) = s_{i_2, j_2}^f(t) \mid a_{i_1, j_1}^{\text{miss}}(t) = 1, a_{i_2, j_2}^{\text{miss}}(t) = 0\right),
 \end{aligned}
 \tag{34}$$

where (i_1, j_1) and (i_2, j_2) are two randomly selected node pairs with $a_{i_1, j_1}^{\text{miss}}(t) = 1$ and $a_{i_2, j_2}^{\text{miss}}(t) = 0$. Intuitively, Equation (34) measures the probability that the link predictor f ranks a randomly selected positive sample (i_1, j_1) higher than a randomly selected negative sample (i_2, j_2) . See the two surveys [20,111] for additional details of the AUC.

Table 2. Performance evaluation of the synergistic link-prediction methods based on AUC. Six synergy-based indices were implemented in this study, namely, S-WCN, S-WAA, S-WRA, S-rWCN, S-rWAA, and S-rWRA, which aim at enhancing the corresponding local similarity measures (i.e., WCN, WAA, WRA, rWCN, rWAA, and rWRA) by incorporating the community-based index B_{g_i, g_j} . Pairwise comparisons were conducted between the synergy-based index and its corresponding local similarity measure. The one with the better performance is underlined. The best link predictor for each year is marked in bold. It is observed that the synergy-based indices outperformed both B_{g_i, g_j} and the corresponding local similarity measures in most cases, which could validate the existence and usefulness of the synergistic effect between two scales on link prediction.

Year	B_{g_i, g_j}	WCN	S-WCN	WAA	S-WAA	WRA	S-WRA	rWCN	S-rWCN	rWAA	S-rWAA	rWRA	S-rWRA
2003	0.7580	0.7192	<u>0.8013</u>	0.7194	<u>0.8028</u>	0.7199	<u>0.8030</u>	0.7192	<u>0.8016</u>	0.7193	<u>0.8021</u>	0.7198	0.8046
2004	0.7513	0.5926	<u>0.7500</u>	0.5926	<u>0.7500</u>	0.5922	<u>0.7489</u>	0.5926	<u>0.7497</u>	0.5926	<u>0.7506</u>	0.5925	<u>0.7501</u>
2005	0.6785	0.6962	<u>0.7693</u>	0.6964	<u>0.7697</u>	0.6968	0.7840	0.6957	<u>0.7683</u>	0.6960	<u>0.7533</u>	0.6963	<u>0.7520</u>
2006	0.7926	0.7145	<u>0.8454</u>	0.7142	<u>0.8233</u>	0.7134	<u>0.8220</u>	0.7148	0.8462	0.7145	<u>0.8238</u>	0.7137	<u>0.8252</u>
2007	0.7698	0.8121	<u>0.8844</u>	0.8130	<u>0.8872</u>	0.8143	0.8892	0.8106	<u>0.8823</u>	0.8112	<u>0.8860</u>	0.8129	<u>0.8874</u>
2008	0.7700	0.7694	<u>0.8844</u>	0.7697	0.8850	0.7697	<u>0.8712</u>	0.7689	<u>0.8849</u>	0.7690	<u>0.8841</u>	0.7692	<u>0.8713</u>
2009	0.6869	<u>0.8813</u>	0.8512	<u>0.8823</u>	0.8554	0.8841	0.8583	<u>0.8797</u>	0.8486	<u>0.8804</u>	0.8526	<u>0.8828</u>	0.8582
2010	0.6926	0.7940	<u>0.8295</u>	0.7936	0.8342	0.7914	<u>0.8072</u>	0.7935	<u>0.8313</u>	0.7933	<u>0.8299</u>	0.7916	<u>0.8304</u>
2011	0.7769	0.6914	<u>0.7658</u>	0.6914	<u>0.7628</u>	0.6904	<u>0.7661</u>	0.6904	<u>0.7646</u>	0.6904	<u>0.7647</u>	0.6899	<u>0.7666</u>
2012	0.6698	<u>0.7068</u>	0.7059	0.7062	0.7115	<u>0.7043</u>	0.6620	0.7069	<u>0.7090</u>	0.7066	<u>0.7075</u>	0.7051	<u>0.7096</u>
2013	0.7049	0.7301	<u>0.7525</u>	0.7312	<u>0.7572</u>	0.7335	0.7606	<u>0.7303</u>	0.7180	<u>0.7310</u>	0.7192	0.7334	<u>0.7582</u>
2014	0.7089	0.7165	<u>0.7476</u>	0.7176	<u>0.7429</u>	0.7199	0.7590	0.7153	<u>0.7450</u>	0.7160	<u>0.7402</u>	0.7181	<u>0.7498</u>
2015	0.6978	0.7045	<u>0.7539</u>	0.7052	<u>0.7474</u>	0.7064	<u>0.7560</u>	0.7036	0.7594	0.7041	<u>0.7559</u>	0.7056	<u>0.7522</u>
2016	0.7336	0.7259	<u>0.7743</u>	0.7282	<u>0.7767</u>	0.7301	<u>0.7774</u>	0.7245	<u>0.7738</u>	0.7260	<u>0.7785</u>	0.7284	0.7787
2017	0.6748	0.7555	<u>0.7635</u>	0.7565	0.7671	0.7553	<u>0.7573</u>	<u>0.7533</u>	0.7485	0.7540	<u>0.7544</u>	<u>0.7541</u>	0.7521
2018	0.7396	0.7678	<u>0.8108</u>	0.7699	<u>0.8148</u>	0.7712	<u>0.8164</u>	0.7662	0.8190	0.7677	<u>0.8082</u>	0.7703	<u>0.8157</u>
2019	0.7144	<u>0.8231</u>	0.8201	0.8248	0.8238	<u>0.8237</u>	0.8181	0.8209	0.7962	<u>0.8221</u>	0.8036	<u>0.8216</u>	0.7889
2020	0.7640	0.8357	<u>0.8612</u>	0.8392	<u>0.8667</u>	0.8421	0.8684	0.8327	<u>0.8379</u>	0.8355	<u>0.8453</u>	0.8398	<u>0.8603</u>
2021	0.5090	<u>0.5506</u>	0.5321	<u>0.5521</u>	0.5399	0.5539	0.5391	0.5483	<u>0.5503</u>	0.5499	<u>0.5508</u>	<u>0.5514</u>	0.5483

Table 2 shows the performances of 13 link predictors, where the first column displays the AUCs when only macroscopic variables $B_{g_i g_j}$ are used as link-prediction scores for node pairs (i, j) . From columns 2 to 13, two adjacent columns represent the AUCs of a two-hop link predictor f , and the AUCs of the synergistic link prediction that combines f and $B_{g_i g_j}$, respectively. For example, the column “WCN” shows the AUCs for the link-prediction scores s_{ij}^{WCN} , which are calculated by the WCN link predictor as

$$s_{ij}^{\text{WCN}} = \sum_{k \in |\Gamma(i) \cap \Gamma(j)|} (a_{i,k} + a_{j,k}), \quad (35)$$

(see Table 1). The column “S-WCN” represents the AUCs for the synergistic link prediction that combines s_{ij}^{WCN} and $B_{g_i g_j}$. The link-prediction score of S-WCN is calculated based on Equation (33) as

$$s_{ij}^{\text{S-WCN}}(t) = \sum_{B \leq B_{g_i g_j}(t)} \sum_{s \leq s_{ij}^{\text{WCN}}(t)} p(a_{ij}^{\text{train}} = 1 | B, s). \quad (36)$$

In Table 2, we first compared the performance of the 13 link predictors and marked the optimal AUC in bold font. It can be observed that synergistic link predictors obtained the best performance in the majority of years (14 out of 19). S-WRA, S-WAA, S-rWCN, and S-rWRA achieved the best AUC in 5, 4, 3, and 2 years, respectively. Exceptions occurred in 2004, 2009, 2011, 2019, and 2021. The best link predictor in 2004 and 2011 was $B_{g_i g_j}$. WRA achieved the best performance in 2009 and 2021, and the best link predictor in 2019 was WAA.

We then compared the performance of each two-hop link predictor with that of its synergistic counterpart and underlined the one with better performance. The results show that in most cases, the synergistic link predictor outperformed the corresponding two-hop link predictors. Exceptions occurred in a few years. In 2009 and 2019, all of the two-hop link predictors outperformed their synergistic counterparts. In 2012, WCN and WRA achieved better performance than their synergistic counterparts. In 2013, the performances of rWCN and rWAA exceeded the performances of their synergistic counterparts. In 2017, rWCN and rWRA performed better than their synergistic counterparts. In 2021, four algorithms outperformed their synergistic counterparts, namely, WCN, WAA, WRA, and rWRA.

Such results demonstrate the superiority of synergistic link predictors over the link predictors that consider only microscopic or macroscopic variables. Further, intriguing phenomena can be observed when we compare the link-prediction performance in Table 2 with the PID analysis in Figure 9.

- For 2004, the values of $\rho_{\text{unq}}(B | s)$ are significantly greater than zero, unlike any other years. Such an observation indicates that the macroscopic variable $B_{g_i g_j}$ provides much unique information for link prediction beyond s_{ij} . At the same time, the redundant information ρ_{red} approaches (Figure 9a,b) or even exceeds (Figure 9c–f) the synergistic information ρ_{syn} . Correspondingly, Table 2 shows that $B_{g_i g_j}$ outperformed any two-hop link predictor or synergistic link predictor and achieved the best performance.
- For 2009, Table 2 indicates that for each pair of two-hop and synergistic link predictors, the two-hop link predictors remarkably outperform the synergistic ones. Correspondingly, Figure 9 shows that the values of $\rho_{\text{unq}}(s | B)$ are very large and clearly exceed the values of ρ_{red} and ρ_{syn} . Meanwhile, the values of $\rho_{\text{unq}}(B | s)$ approach zero for all the six two-hop link predictors.
- The situation for 2011 is similar to that for 2004. $B_{g_i g_j}$ outperformed all the two-hop and synergistic link predictors in Table 2 and the redundant information ρ_{red} approaches (Figure 9a,b), and even exceeded (Figure 9c–f) the synergistic information ρ_{syn} in the PID results in Figure 9. In addition, it is shown that the values of $\rho_{\text{unq}}(s | B)$ are relatively small, and $\rho_{\text{unq}}(B | s)$ exceeds zero in Figure 9a,b.

- The results for 2012, 2013, and 2017 in Table 2 could be summarized in the same way; that is, there are some two-hop link predictors that outperformed their synergistic counterparts, but the best performance among all the 13 link predictors was still achieved by some kind of synergistic link predictors. Such phenomena could be partially explained by relatively large unique information $\rho_{\text{unq}}(s | B)$. Meanwhile, it is noteworthy that most of the differences in performance between two-hop link predictors and their synergistic counterparts are not significant for 2012 and 2017. This might be attributed to the poor performance of $B_{g_i g_j}$ (AUCs of 0.6698 and 0.6748, respectively), which resulted in incorporating inaccurate information, hence the unsatisfactory performance of some of the synergistic link predictors.
- The link-prediction performance of 2019 is similar to that of 2009 in Table 2, whereas their PID results in Figure 9 are noticeably different. The 2019 PID results show that although $\rho_{\text{unq}}(s | B)$ is relatively large, it does not exceed ρ_{syn} and ρ_{red} in most cases. Given the fact that the performance of $B_{g_i g_j}$ is remarkably weaker than that of the two-hop link predictors, we speculate that the main reason for the unsatisfactory performance of the synergistic link predictors in 2019 may be the inaccurate information brought by $B_{g_i g_j}$.
- The year of 2021 is a special case. The experiments of 2021 aimed to predict the new links that appear in 2022 based on the snapshot of 2021, but the data collected in 2022 are incomplete and only cover the first half of the year. Since the AUCs for 2021 in Table 2 are relatively small compared to the AUCs for the year before and after 2021, we suspect that the current experimental results in 2021 could not reflect the general characteristics of link prediction.
- For other years, all the synergistic link predictors outperformed the corresponding two-hop link predictors, and the best performance was also achieved by some synergistic link predictors. Moreover, in some representative years, such as 2014–2016, the synergistic link predictors achieved noticeable performance improvements compared to $B_{g_i g_j}$ and two-hop link predictors, and relatively large amounts of synergistic information ρ_{syn} can be observed in the PID results.

Based on the above observations, we could summarize the following three properties of link prediction on technology-convergence networks from the PID perspective. First, combining two different scales (i.e., microscopic and macroscopic) of the network structure commonly yields better link prediction results, mainly because the two scales jointly provide a large amount of synergistic information. Second, jointly considering two scales may not be superior to using a single scale when there is too much redundant or unique information and not enough synergistic information. Finally, when one scale provides much more inaccurate information than the other, jointly considering two scales may not be preferable to focusing just on one scale.

Furthermore, another intriguing observation is that in the early stage of the development of ITS (i.e., 2003 to 2008), a large amount of redundant information existed in the PID results shown in Figure 9. Meanwhile, the AUCs of $B_{g_i g_j}$ in Table 2 approach or exceed the AUCs of the two-hop link predictors in most cases. In the middle of the ITS technology's development (i.e., 2009 to 2013), remarkable unique information $\rho_{\text{unq}}(s | B)$ existed in the PID results in most instances, and the two-hop link predictors outperformed $B_{g_i g_j}$. Synergistic information progressively dominated the PID results in the late stages of ITS technology development (i.e., 2014 and after), for which the synergistic link predictors noticeably outperformed both $B_{g_i g_j}$ and two-hop link predictors in most circumstances.

The aforementioned findings may provide key insights into the generation mechanism of technology convergence in ITS. In the early stages of ITS development, technology convergence often occurs between IPC subclasses belonging to the same IPC class. In other words, technologies in similar domains tend to converge. In the middle stages, microscopic structures in technology-convergence networks become the primary driving forces of technology convergence. Finally, in the late stages, the dominant factor in the generation of

technology convergence is the synergistic effect between technologies' similarity and their microscopic structures in the technology-convergence network.

5. Conclusions

This study aimed to provide theoretical underpinnings for previous empirical research on the anticipation of technology convergence using link-prediction algorithms. Existing studies mainly focused on using different link-prediction algorithms to produce accurate predictions of technology convergence in a variety of technology domains, such as ICT, 3D printing, and electric vehicles. However, theoretical explanations for such a prediction paradigm are still relatively lacking: What properties and growth mechanisms of the technology-convergence network guarantee the predictability of future links? To answer that question, we analyzed the ITS-related patent applications between November 1995 and July 2022 and provided empirical evidence to demonstrate that technology-convergence networks are fractal networks with self-similar growth processes. Such findings provide a possible theoretical explanation for the success of link-prediction algorithms in the anticipation of technology convergence, which means that the topological structure of the future snapshot is similar to that of the present snapshot.

Further, the mechanism of how the fractality of snapshots contributes to link prediction was also explored. This study adopted the PID framework to the link-prediction problem and found that the scaled-down replica is attributable to its synergy with microscopic structures. A link-prediction method based on joint probability distributions was then proposed to utilize such synergistic information. Experiments on the ITS-related patent applications demonstrated that our proposed synergistic link-prediction algorithm outperforms benchmark methods in most cases. In addition, the PID results imply that the main driving force of technology convergence varies at different stages of technology development. The results of this study could help researchers better comprehend the generation mechanisms of technology convergence from a network science perspective, which in turn could also assist companies and policymakers in developing more accurate predictive algorithms and R&D policies for potential technology convergence opportunities.

The main limitation of this study is that it overlooks the higher-order interactions of technology areas in technology-convergence networks. For example, for three technology areas (i.e., IPC4 codes) that apply for the same patent, their interactions were modeled as three pairwise connections in this study. However, such a conventional pairwise modeling approach could lead to a loss of information, since the patent could only be created when the three technology areas were combined together. Therefore, future research on technology convergence could benefit from using higher-order networks [132–134], in which the links are “simplicial complexes” that can connect more than two nodes [135–137].

Moreover, the SLP approach proposed in this study is quite intuitive, lacking a careful design for the formula of the link prediction index. In particular, the sparsity of the links in real-world networks could cause serious issues for the estimation of probabilities, although we tried to tackle this issue, at least partially, via a CDF-like formula. In the future, we would like to extend this study from the perspective of link prediction. First, we would like to examine whether the synergy between community structures and local similarity measures is widespread in real-world networks. Second, a better-designed link/community-based strategy could be proposed to make greater use of the synergistic effect. Third, extensive comparisons with different link/community-based strategies could be conducted in a broad diversity of real-world networks.

Author Contributions: Conceptualization, Y.X.; methodology, Y.X.; software, Y.X.; validation, Y.X., K.C., X.R., B.C. and W.K.C.; formal analysis, Y.X.; investigation, Y.X.; resources, B.C. and W.K.C.; data curation, Y.X. and K.C.; writing—original draft preparation, Y.X.; writing—review and editing, K.C., X.R., B.C. and W.K.C.; visualization, Y.X. and K.C.; supervision, B.C. and W.K.C.; project administration, B.C. and W.K.C.; funding acquisition, B.C. and W.K.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the Project from Science and Technology Innovation Committee of Shenzhen (grant no. JCYJ20210324135011030), the Science and Technology Innovation Committee of Shenzhen-Platform and Carrier (International Science and Technology Information Center), the Project from Science and Technology Innovation Committee of Shenzhen (WDZC20200818121348001, JCYJ20210324115604012, JCYJ20190813173401651), the National Natural Science Foundation of China (grant no. 71971127), the High-end Foreign Expert Talent Introduction Plan (G2021032022L), the Guangdong Pearl River Plan (2019QN01X890), the Tsinghua University Spring Breeze Fund (grant no. 2021Z99CFW043), the Tsinghua University Fund (grant no. HW2020005, JC2021009, JC2022018), the Tsinghua-Toyota Joint Research Fund, and Hylink Digital Solutions Co., Ltd. (120500002).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study used patent data from the Derwent Innovation database, which can be found at <https://derwentinnovation.clarivate.com.cn/ui/zh/>, accessed on 1 July 2022.

Acknowledgments: We gratefully acknowledge Ting Zeng, Cuizhong Guan, and their colleagues from Tsinghua University Library for their most valuable support in collecting patent data from Derwent Innovation database. Likewise, we thank the support from Causal Emergence Reading Group supported by the Save 2050 Programme jointly sponsored by Swarma Club and X-Order. In particular, we would like to thank Jiang Zhang, Zhang Zhang, Aobo Lü, and Ting Wang from Causal Emergence Reading Group for sharing their knowledge with us. In addition, we would like to thank Yaochun Jiang, Lu Yu, Xiaoying Lin, Qianqian Zhang, and their colleagues from China Truer IP, and Chao Li from the Telecommunication Examination Department of the China National Intellectual Property Administration, who provided valuable information for this study from the perspective of patent practice. Finally, we would like to thank the Party Building Pacesetter Program of Tsinghua University for supporting the Party Branch of Tsinghua-Berkeley Shenzhen Institute 2019 Ph.D. Students, which provides us with the opportunities and platforms to communicate with professionals and practitioners in the intellectual property sector.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Rescaled Networks Statistics as Measures of Self-Similar Growth

The definitions of the average degree and the five rescaled network statistics for each snapshot in $\mathcal{G}^{\text{hist}}$ are given below. Note that the snapshots in $\mathcal{G}^{\text{hist}}$ may not be fully connected, so only the largest connected component in each snapshot is considered in this study.

Average degree. The average degree \bar{d} of a snapshot is defined as

$$\bar{d} = \frac{\sum_{i=1}^N d_i}{N}, \quad (\text{A1})$$

where N is the number of nodes and d_i is the weighted degree of node i . Since the IPC co-occurrence network considered in this study is a weighted network and the weight of the edge (m, i) is denoted as $a_{m,i}$, the weighted degree d_i is given as

$$d_i = \sum_{m \in \Gamma(i)} a_{m,i}, \quad (\text{A2})$$

where $\Gamma(i)$ denotes the set of neighbors of node i . The weighted degree d_i is the sum of weights of all the associated links of node i , also known as the strength of node i .

Rescaled degree: The rescaled degree of node k is given as

$$d_k^{\text{res}} = \frac{d_k}{\bar{d}}, \quad (\text{A3})$$

where \bar{d} and d_k are defined in Equations (A1) and (A2), respectively.

Clustering coefficient over the rescaled degree: The clustering coefficient of node k is given as

$$C(k) = \frac{2}{d_k^{\text{res}}(d_k^{\text{res}} - 1)} \sum_{ij} \left(\tilde{a}_{k,i} \cdot \tilde{a}_{ij} \cdot \tilde{a}_{j,k} \right)^{\frac{1}{3}}, \quad (\text{A4})$$

where $\tilde{a}_{k,i}$, \tilde{a}_{ij} , and $\tilde{a}_{j,k}$ are the normalized weights of links (k, i) , (i, j) , and (j, k) , respectively, [138]. In this study, the normalization was performed by dividing the largest weights, i.e., $\tilde{a}_{ij} = a_{ij} / \max(a_{ij})$ [139]. Then, the clustering coefficient over the rescaled degree is the average of C_k of all the nodes with the same rescaled degree, whose formal definition is given as

$$\bar{C}(d^{\text{res}}) = \frac{\sum_{k \in \mathcal{N}(d^{\text{res}})} C(k)}{|\mathcal{N}(d^{\text{res}})|}, \quad (\text{A5})$$

where $\mathcal{N}(d^{\text{res}})$ denotes the set of nodes whose rescaled degrees are d^{res} [140].

Normalized average neighbor degree: For the node k , the average degree of its neighbors is given as

$$d_{\text{nn}}^{\text{res}}(k) = \frac{\sum_{i \in \Gamma(k)} d_i^{\text{res}}}{d_k^{\text{res}}}, \quad (\text{A6})$$

and the normalized average neighbor degree over the rescaled degree is given as

$$\bar{d}_{\text{nn}}^{\text{res}}(d^{\text{res}}) = \frac{\sum_{k \in \mathcal{N}(d^{\text{res}})} d_{\text{nn}}^{\text{res}}(k) \bar{d}}{|\mathcal{N}(d^{\text{res}})| \bar{d}^2}, \quad (\text{A7})$$

where $\bar{d} = \frac{\sum_{i=1}^N d_i}{N}$ and $\bar{d}^2 = \frac{\sum_{i=1}^N d_i^2}{N}$ [141,142].

Modularity: The modularity of the community partitions is given as

$$Q = \sum_{r=1}^{N_B} \left[\frac{\sum_{i \in \mathcal{N}(r)} \sum_{j \in \mathcal{N}(r)} a_{ij}}{m} - \gamma \left(\frac{\sum_{i \in \mathcal{N}(r)} d_i}{2m} \right)^2 \right], \quad (\text{A8})$$

where the summation is from r equal to one to all N_B communities, $\mathcal{N}(r)$ denotes the set of nodes in the r -th community, and $m = \frac{\sum_{i=1}^N d_i}{2}$ is the summation of all the edge weights [143]. The communities in each snapshot are detected based on the Louvain method [130].

AMI: For the snapshot at time t , we measure the AMI between its community partitions \mathbf{g}_t and the community partitions of the previous snapshot \mathbf{g}_{t-1} . The formal definition of AMI is given as

$$\text{AMI}(\mathbf{g}_t, \mathbf{g}_{t-1}) = \frac{I(\mathbf{g}_t, \mathbf{g}_{t-1}) - \mathbb{E}[I(\mathbf{g}_t, \mathbf{g}_{t-1})]}{\max\{H(\mathbf{g}_t), H(\mathbf{g}_{t-1})\} - \mathbb{E}[I(\mathbf{g}_t, \mathbf{g}_{t-1})]}, \quad (\text{A9})$$

where $I(\cdot)$, $H(\cdot)$, and $\mathbb{E}[\cdot]$ denote the mutual information, the entropy, and the expectation, respectively.

Appendix B. A Community-Structure-Based Approximation of the Box-Covering Algorithm

The detailed procedures of the community-structure-based box-covering algorithm are given in Algorithm A1. It is worth noting that minor adjustments were made, so some details are different from those in the original paper [38]. In addition, the parameters in Algorithm A1 were set specifically for this study. The resolution list was set as $\mathbf{R} = [12.25, 12, 11.75, \dots, 0.25]$, and ζ_{\min} and ζ_{\max} were set as 0.05 and 0.3, respectively.

Algorithm A1 Giudicianni's community-structure-based box-covering algorithm.**Input:**

A snapshot G_t^{hist} .

A list of η different resolutions $\mathbf{R} = [R_1, R_2, \dots, R_\eta]$;

Output:

A list of the number of boxes \mathbf{N}_B .

A list of the average size of boxes \mathbf{I}_B .

- 1: $\mathbf{N}_B \leftarrow []$.
- 2: $\mathbf{I}_B \leftarrow []$.
- 3: Obtain the largest connected component of G_t^{hist} as G_t^{LCC} .
- 4: Obtain the number of nodes in G_t^{LCC} as N^{LCC} .
- 5: **for** $i = 1$ to η **do**
- 6: Set $N_B(i) = 0$.
- 7: Set $I_B(i) = 0$.
- 8: Perform community detection on G_t^{LCC} using Louvian algorithm with setting the resolution as R_i .
- 9: **for** Each detected community c **do**
- 10: $N_B(i) = N_B(i) + 1$.
- 11: Obtain the diameter of the subgraph c as l_c .
- 12: $I_B(i) = I_B(i) + l_c$.
- 13: **end for**
- 14: **if** $\zeta_{\min} \cdot N^{\text{LCC}} < N_B(i) < \zeta_{\max} \cdot N^{\text{LCC}}$ **then**
- 15: $\mathbf{N}_B \leftarrow \mathbf{N}_B \cup N_B(i)$.
- 16: $\mathbf{I}_B \leftarrow \mathbf{I}_B \cup \frac{I_B(i)}{N_B(i)}$.
- 17: **end if**
- 18: **end for**
- 19: **return** $\mathbf{N}_B, \mathbf{I}_B$.

References

1. Adner, R.; Levinthal, D.A. The emergence of emerging technologies. *Calif. Manag. Rev.* **2002**, *45*, 50–66. [\[CrossRef\]](#)
2. Rotolo, D.; Hicks, D.; Martin, B.R. What is an emerging technology? *Res. Policy* **2015**, *44*, 1827–1843. [\[CrossRef\]](#)
3. Hacklin, F.; Raurich, V.; Marxt, C. How incremental innovation becomes disruptive: The case of technology convergence. In Proceedings of the 2004 IEEE International Engineering Management Conference (IEEE Cat. No.04CH37574), Singapore, 18–21 October 2004; pp. 32–36.
4. Kim, K.; Jung, S.; Hwang, J. Technology convergence capability and firm innovation in the manufacturing sector: An approach based on patent network analysis. *R D Manag.* **2019**, *49*, 595–606. [\[CrossRef\]](#)
5. Tang, Y.; Lou, X.; Chen, Z.; Zhang, C. A study on dynamic patterns of technology convergence with IPC Co-Occurrence-based analysis: The case of 3D printing. *Sustainability* **2020**, *12*, 2655. [\[CrossRef\]](#)
6. Song, C.H.; Elvers, D.; Leker, J. Anticipation of converging technology areas—A refined approach for the identification of attractive fields of innovation. *Technol. Forecast. Soc. Chang.* **2017**, *116*, 98–115. [\[CrossRef\]](#)
7. Lee, C.; Hong, S.; Kim, J. Anticipating multi-technology convergence: A machine learning approach using patent information. *Scientometrics* **2021**, *126*, 1867–1896. [\[CrossRef\]](#)
8. Sick, N.; Bröring, S. Exploring the research landscape of convergence from a TIM perspective: A review and research agenda. *Technol. Forecast. Soc. Chang.* **2022**, *175*, 121321. [\[CrossRef\]](#)
9. Lee, W.S.; Han, E.J.; Sohn, S.Y. Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents. *Technol. Forecast. Soc. Chang.* **2015**, *100*, 317–329. [\[CrossRef\]](#)
10. Caviggioli, F. Technology fusion: Identification and analysis of the drivers of technology convergence using patent data. *Technovation* **2016**, *55–56*, 22–32. [\[CrossRef\]](#)
11. Oh, S.; Choi, J.; Ko, N.; Yoon, J. Predicting product development directions for new product planning using patent classification-based link prediction. *Scientometrics* **2020**, *125*, 1833–1876. [\[CrossRef\]](#)
12. Kwon, O.J.; An, Y.; Kim, M.; Lee, C. Anticipating technology-driven industry convergence: Evidence from large-scale patent analysis. *Technol. Anal. Strateg. Manag.* **2020**, *32*, 363–378. [\[CrossRef\]](#)

13. Feng, S.; An, H.; Li, H.; Qi, Y.; Wang, Z.; Guan, Q.; Li, Y.; Qi, Y. The technology convergence of electric vehicles: Exploring promising and potential technology convergence relationships and topics. *J. Clean. Prod.* **2020**, *260*, 120992. [[CrossRef](#)]
14. Lee, J.; Ko, N.; Yoon, J.; Son, C. An approach for discovering firm-specific technology opportunities: Application of link prediction to F-term networks. *Technol. Forecast. Soc. Chang.* **2021**, *168*, 120746. [[CrossRef](#)]
15. Aaldering, L.J.; Leker, J.; Song, C.H. Competition or collaboration?—Analysis of technological knowledge ecosystem within the field of alternative powertrain systems: A patent-based approach. *J. Clean. Prod.* **2019**, *212*, 362–371. [[CrossRef](#)]
16. Hong, S.; Lee, C. Effective indexes and classification algorithms for supervised link prediction approach to anticipating technology convergence: A comparative study. *IEEE Trans. Eng. Manag.* **2021**, 1–12. [[CrossRef](#)]
17. Kim, T.S.; Sohn, S.Y. Machine-learning-based deep semantic analysis approach for forecasting new technology convergence. *Technol. Forecast. Soc. Chang.* **2020**, *157*, 120095. [[CrossRef](#)]
18. Choi, S.; Afifuddin, M.; Seo, W. A Supervised/ Learning-Based Approach to Anticipating Potential Technology Convergence. *IEEE Access* **2022**, *10*, 19284–19300. [[CrossRef](#)]
19. Zhang Jinzhu, H.Y.; Zhang Jinzhu, H.Y. Prediction and evaluation of technology convergence based on multi-features fusion. *Data Anal. Knowl. Discov.* **2022**, *6*, 33–44.
20. Lü, L.; Zhou, T. Link prediction in complex networks: A survey. *Phys. A Stat. Mech. Appl.* **2011**, *390*, 1150–1170. [[CrossRef](#)]
21. Liu, Z.; Zhang, Q.M.; Lü, L.; Zhou, T. Link prediction in complex networks: A local naïve Bayes model. *EPL* **2011**, *96*, 48007. [[CrossRef](#)]
22. Wang, W.Q.; Zhang, Q.M.; Zhou, T. Evaluating network models: A likelihood analysis. *EPL* **2012**, *98*, 28004. [[CrossRef](#)]
23. Chen, D.B.; Gao, H.; Lü, L.; Zhou, T. Identifying influential nodes in large-scale directed networks: The role of clustering. *PLoS ONE* **2013**, *8*, e77455. [[CrossRef](#)] [[PubMed](#)]
24. Zhang, Q.M.; Xu, X.K.; Zhu, Y.X.; Zhou, T. Measuring multiple evolution mechanisms of complex networks. *Sci. Rep.* **2015**, *5*. [[CrossRef](#)] [[PubMed](#)]
25. Pech, R.; Hao, D.; Lee, Y.L.; Yuan, Y.; Zhou, T. Link prediction via linear optimization. *Phys. A Stat. Mech. Appl.* **2019**, *528*, 121319. [[CrossRef](#)]
26. Song, C.; Havlin, S.; Makse, H.A. Origins of fractality in the growth of complex networks. *Nat. Phys.* **2006**, *2*, 275–281. [[CrossRef](#)]
27. Xi, L.; Wang, L.; Wang, S.; Yu, Z.; Wang, Q. Fractality and scale-free effect of a class of self-similar networks. *Phys. A Stat. Mech. Appl.* **2017**, *478*, 31–40. [[CrossRef](#)]
28. Zheng, M.; Garcia-Pérez, G.; Boguñá, M.; Serrano, M.Á. Scaling up real networks by geometric branching growth. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2018994118. [[CrossRef](#)]
29. Song, C.; Havlin, S.; Makse, H.A. Self-similarity of complex networks. *Nature* **2005**, *433*, 392–395. [[CrossRef](#)]
30. Wen, T.; Cheong, K.H. The fractal dimension of complex networks: A review. *Inf. Fusion* **2021**, *73*, 87–102. [[CrossRef](#)]
31. Gao, L.; Peng, J.; Tang, C. Optimizing the first-passage process on a class of fractal scale-free trees. *Fractal Fract.* **2021**, *5*, 184. [[CrossRef](#)]
32. Babič, M.; Marinkovic, D.; Bonfanti, M.; Calì, M. Complexity modeling of steel-laser-hardened surface microstructures. *Appl. Sci.* **2022**, *12*, 2458. [[CrossRef](#)]
33. Garcia-Pérez, G.; Boguñá, M.; Serrano, M.A. Multiscale unfolding of real networks by geometric renormalization. *Nat. Phys.* **2018**, *14*, 583–589. [[CrossRef](#)]
34. Zheng, M.; Allard, A.; Hagmann, P.; Alemán-Gómez, Y.; Serrano, M.A. Geometric renormalization unravels self-similarity of the multiscale human connectome. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 20244–20253. [[CrossRef](#)] [[PubMed](#)]
35. Ortiz, E.; García-Pérez, G.; Serrano, M.A. Geometric detection of hierarchical backbones in real networks. *Phys. Rev. Res.* **2020**, *2*, 033519. [[CrossRef](#)]
36. Chen, D.; Su, H.; Wang, X.; Pan, G.J.; Chen, G. Finite-size scaling of geometric renormalization flows in complex networks. *Phys. Rev. E* **2021**, *104*, 034304. [[CrossRef](#)]
37. Chen, D.; Su, H.; Zeng, Z. Geometric renormalization reveals the self-similarity of weighted networks. *IEEE Trans. Comput. Soc. Syst.* **2022**, 1–9. [[CrossRef](#)]
38. Giudicianni, C.; Di Nardo, A.; Greco, R.; Scala, A. A community-structure-based method for estimating the fractal dimension, and its application to water networks for the assessment of vulnerability to disasters. *Water Resour. Manag.* **2021**, *35*, 1197–1210. [[CrossRef](#)]
39. Cannistraci, C.V.; Alanis-Lobato, G.; Ravasi, T. From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. *Sci. Rep.* **2013**, *3*, 1613. [[CrossRef](#)]
40. Yan, B.; Gregory, S. Finding missing edges in networks based on their community structure. *Phys. Rev. E* **2012**, *85*, 056112. [[CrossRef](#)]
41. Ding, J.; Jiao, L.; Wu, J.; Liu, F. Prediction of missing links based on community relevance and ruler inference. *Knowl.-Based Syst.* **2016**, *98*, 200–215. [[CrossRef](#)]
42. Gao, F.; Musial, K.; Gabrys, B. A community bridge boosting social network link prediction model. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017, Sydney, Australia, 31 July–3 August 2017; pp. 683–689.
43. Jalili, M.; Orouskhani, Y.; Asgari, M.; Alipourfard, N.; Perc, M. Link prediction in multiplex online social networks. *R. Soc. Open Sci.* **2017**, *4*, 160863. [[CrossRef](#)] [[PubMed](#)]

44. Li, L.; Fang, S.; Bai, S.; Xu, S.; Cheng, J.; Chen, X. Effective link prediction based on community relationship strength. *IEEE Access* **2019**, *7*, 43233–43248. [[CrossRef](#)]
45. Bai, S.; Fang, S.; Li, L.; Liu, R.; Chen, X. Enhancing link prediction by exploring community membership of nodes. *Int. J. Mod. Phys. B* **2019**, *33*, 1950382. [[CrossRef](#)]
46. Javari, A.; Norouzitallab, M.; Jalili, M. Who will accept my request? Predicting response of link initiation in two-way relation networks. *arXiv* **2020**, arXiv:2012.11172.
47. Saxena, A.; Fletcher, G.; Pechenizkiy, M. HM-EIICT: Fairness-aware link prediction in complex networks using community information. *J. Comb. Optim.* **2022**, *44*, 2853–2870. [[CrossRef](#)]
48. Rosenberg, N. Technological change in the machine tool industry, 1840–1910. *J. Econ. Hist.* **1963**, *23*, 414–443. [[CrossRef](#)]
49. Kodama, F. Technology fusion and the new R&D. *Harvard Business Review* 1992. Available online: <https://hbr.org/1992/07/technology-fusion-and-the-new-rd> (accessed on 18 September 2022).
50. Bruce, M.; Leverick, F.; Littler, D.; Wilson, D. Success factors for collaborative product development: A study of suppliers of information and communication technology. *R D Manag.* **1995**, *25*, 33–44. [[CrossRef](#)]
51. Duysters, G.; Hagedoorn, J. Technological convergence in the IT industry: The role of strategic technology alliances and technological competencies. *Int. J. Econ. Bus.* **1998**, *5*, 355–368. [[CrossRef](#)]
52. Hacklin, F. *Management of Convergence in Innovation: Strategies and Capabilities for Value Creation beyond Blurring Industry Boundaries*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007.
53. Pennings, J.M.; Puranam, P. Market convergence & firm strategy: New directions for theory and research. In *Proceeding of the ECIS Conference, The Future of Innovation Studies*, Eindhoven, The Netherlands, 20–23 September 2001; Volume 20.
54. Curran, C.S. *The Anticipation of Converging Industries*; Springer: London, UK, 2013.
55. Karvonen, M.; Kässi, T. Patent citations as a tool for analysing the early stages of convergence. *Technol. Forecast. Soc. Chang.* **2013**, *80*, 1094–1107. [[CrossRef](#)]
56. Ko, N.; Yoon, J.; Seo, W. Analyzing interdisciplinarity of technology fusion using knowledge flows of patents. *Expert Syst. Appl.* **2014**, *41*, 1955–1963. [[CrossRef](#)]
57. Hussain, A.; Jeon, J.; Rehman, M. Technological convergence assessment of the smart factory using patent data and network analysis. *Sustainability* **2022**, *14*, 1668. [[CrossRef](#)]
58. Feng, L.; Wang, Q.; Wang, J.; Lin, K.Y. A review of technological forecasting from the perspective of complex systems. *Entropy* **2022**, *24*, 787. [[CrossRef](#)] [[PubMed](#)]
59. No, H.J.; Park, Y. Trajectory patterns of technology fusion: Trend analysis and taxonomical grouping in nanobiotechnology. *Technol. Forecast. Soc. Chang.* **2010**, *77*, 63–75. [[CrossRef](#)]
60. Nemet, G.F.; Johnson, E. Do important inventions benefit from knowledge originating in other technological domains? *Res. Policy* **2012**, *41*, 190–200. [[CrossRef](#)]
61. Yun, J.; Geum, Y. Analysing the dynamics of technological convergence using a co-classification approach: A case of healthcare services. *Technol. Anal. Strateg. Manag.* **2019**, *31*, 1412–1429. [[CrossRef](#)]
62. Lee, J.; Sohn, S.Y. Recommendation system for technology convergence opportunities based on self-supervised representation learning. *Scientometrics* **2021**, *126*, 1–25. [[CrossRef](#)]
63. Jee, J.; Shin, H.; Kim, C.; Lee, S. Six different approaches to defining and identifying promising technology through patent analysis. *Technol. Anal. Strateg. Manag.* **2022**, *34*, 961–973. [[CrossRef](#)]
64. Mazlumi, S.H.H.; Kermani, M.A.M. Investigation the structure of the Internet of things (IoT) patent network using social network analysis. *IEEE Internet Things J.* **2022**, *9*, 13458–13469. [[CrossRef](#)]
65. Wang, P.; Chan, W.K.V. A multilayer community detection algorithm based on aggregation in social Internet of things. In *Proceedings of the 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)*, Xi'an, China, 15–17 April 2022; pp. 610–614.
66. Ancherbak, S.; Kuruoglu, E.E.; Vingron, M. Time-dependent gene network modelling by sequential monte carlo. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2016**, *13*, 1183–1193. [[CrossRef](#)]
67. Kovács, I.A.; Luck, K.; Spirohn, K.; Wang, Y.; Pollis, C.; Schlabach, S.; Bian, W.; Kim, D.K.; Kishore, N.; Hao, T.; Calderwood, M.A.; Vidal, M.; Barabási, A.L. Network-based prediction of protein interactions. *Nat. Commun.* **2019**, *10*, 1240. [[CrossRef](#)]
68. Ye, Q.; Hsieh, C.Y.; Yang, Z.; Kang, Y.; Chen, J.; Cao, D.; He, S.; Hou, T. A unified drug–target interaction prediction framework based on knowledge graph and recommendation system. *Nat. Commun.* **2021**, *12*, 6775. [[CrossRef](#)] [[PubMed](#)]
69. Xiu, Y.; Ren, X.; Zhang, T.; Chen, Y.; Jiang, L.; Li, D.; Wang, X.; Zhao, L.; Chan, W.K. Time labeled visibility graph for privacy-preserved physiological time series classification. In *Proceedings of the 2022 7th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, Chengdu, China, 22–24 April 2022; pp. 280–284.
70. Chan, W.K.V.; Hsu, C. When human networks collide: The degree distributions of hyper-networks. *IIE Trans.* **2015**, *47*, 929–942. [[CrossRef](#)]
71. Babič, M.; Purković, D. A new systemic taxonomy of cyber criminal activity. *Politehnika* **2020**, *4*, 17–27. [[CrossRef](#)]
72. Xiu, Y.; Wang, G.; Chan, W.K.V. Crash diagnosis and price rebound prediction in NYSE composite index based on visibility graph and time-evolving stock correlation network. *Entropy* **2021**, *23*, 1612. [[CrossRef](#)]
73. Babič, M.; Marinković, D.; Kovačić, M.; Šter, B.; Cali, M. A new method of quantifying the complexity of fractal networks. *Fractal Fract.* **2022**, *6*, 282. [[CrossRef](#)]

74. Liu, N.; Fang, J.; Sun, J.; Li, S. Epidemic dynamics of a fractional-order SIR weighted network model and its targeted immunity control. *Fractal Fract.* **2022**, *6*, 232. [[CrossRef](#)]
75. Gallos, L.K.; Song, C.; Makse, H.A. A review of fractality and self-similarity in complex networks. *Phys. A Stat. Mech. Appl.* **2007**, *386*, 686–691. [[CrossRef](#)]
76. Song, C.; Gallos, L.K.; Havlin, S.; Makse, H.A. How to calculate the fractal dimension of a complex network: The box covering algorithm. *J. Stat. Mech. Theory Exp.* **2007**, *2007*, P03006. [[CrossRef](#)]
77. Sun, Y.; Zhao, Y. Overlapping-box-covering method for the fractal dimension of complex networks. *Phys. Rev. E* **2014**, *89*, 042809. [[CrossRef](#)]
78. Liao, H.; Wu, X.; Wang, B.H.; Wu, X.; Zhou, M. Solving the speed and accuracy of box-covering problem in complex networks. *Phys. A: Stat. Mech. Appl.* **2019**, *523*, 954–963. [[CrossRef](#)]
79. Kovács, P.T.; Nagy, M.; Molontay, R. Comparative analysis of box-covering algorithms for fractal networks. *Appl. Netw. Sci.* **2021**, *6*, 73. [[CrossRef](#)]
80. Wei, D.J.; Liu, Q.; Zhang, H.X.; Hu, Y.; Deng, Y.; Mahadevan, S. Box-covering algorithm for fractal dimension of weighted networks. *Sci. Rep.* **2013**, *3*, 3049. [[CrossRef](#)] [[PubMed](#)]
81. Wu, H.; Kuang, L.; Wang, F.; Rao, Q.; Gong, M.; Li, Y. A multiobjective box-covering algorithm for fractal modularity on complex networks. *Appl. Soft Comput.* **2017**, *61*, 294–313. [[CrossRef](#)]
82. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
83. Williams, P.L.; Beer, R.D. Nonnegative decomposition of multivariate information. *arXiv* **2010**, arXiv:1004.2515.
84. Wibral, M.; Priesemann, V.; Kay, J.W.; Lizier, J.T.; Phillips, W.A. Partial information decomposition as a unified approach to the specification of neural goal functions. *Brain Cogn.* **2017**, *112*, 25–38. [[CrossRef](#)]
85. Colenbier, N.; Van de Steen, F.; Uddin, L.Q.; Poldrack, R.A.; Calhoun, V.D.; Marinazzo, D. Disambiguating the role of blood flow and global signal with partial information decomposition. *NeuroImage* **2020**, *213*, 116699. [[CrossRef](#)]
86. Sherrill, S.P.; Timme, N.M.; Beggs, J.M.; Newman, E.L. Partial information decomposition reveals that synergistic neural integration is greater downstream of recurrent information flow in organotypic cortical cultures. *PLoS Comput. Biol.* **2021**, *17*, e1009196. [[CrossRef](#)]
87. Tax, T.M.S.; Mediano, P.A.M.; Shanahan, M. The partial information decomposition of generative neural network models. *Entropy* **2017**, *19*, 474. [[CrossRef](#)]
88. Takimoto, M. An approximation scheme for multivariate information based on partial information decomposition. *arXiv* **2020**, arXiv:2009.00791.
89. Wollstadt, P.; Schmitt, S.; Wibral, M. A rigorous information-theoretic definition of redundancy and relevancy in feature selection based on (partial) information decomposition. *arXiv* **2021**, arXiv:2105.04187.
90. Tokui, S.; Sato, I. Disentanglement analysis with partial information decomposition. *arXiv* **2022**, arXiv:2108.13753.
91. Rosas, F.E.; Mediano, P.A.M.; Jensen, H.J.; Seth, A.K.; Barrett, A.B.; Carhart-Harris, R.L.; Bor, D. Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data. *PLoS Comput. Biol.* **2020**, *16*, e1008289. [[CrossRef](#)] [[PubMed](#)]
92. Zhang, J. Neural information squeezer for causal emergence. *arXiv* **2022**, arXiv:2201.10154.
93. Mediano, P.A.M.; Rosas, F.E.; Luppi, A.I.; Jensen, H.J.; Seth, A.K.; Barrett, A.B.; Carhart-Harris, R.L.; Bor, D. Greater than the parts: A review of the information decomposition approach to causal emergence. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2022**, *380*, 20210246. [[CrossRef](#)]
94. Lorrain, F.; White, H.C. Structural equivalence of individuals in social networks. *J. Math. Sociol.* **1971**, *1*, 49–80. [[CrossRef](#)]
95. Jaccard, P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull. Torrey Bot. Club* **1901**, *37*, 547–579.
96. Ou, Q.; Jin, Y.D.; Zhou, T.; Wang, B.H.; Yin, B.Q. Power-law strength-degree correlation from resource-allocation dynamics on weighted networks. *Phys. Rev. E* **2007**, *75*, 021102. [[CrossRef](#)]
97. Adamic, L.A.; Adar, E. Friends and neighbors on the web. *Soc. Netw.* **2003**, *25*, 211–230. [[CrossRef](#)]
98. Clauset, A.; Moore, C.; Newman, M.E.J. Hierarchical structure and the prediction of missing links in networks. *Nature* **2008**, *453*, 98–101. [[CrossRef](#)]
99. Guimerà, R.; Sales-Pardo, M. Missing and spurious interactions and the reconstruction of complex networks. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 22073–22078. [[CrossRef](#)] [[PubMed](#)]
100. Vallès-Català, T.; Peixoto, T.P.; Sales-Pardo, M.; Guimerà, R. Consistencies and inconsistencies between model selection and link prediction in networks. *Phys. Rev. E* **2018**, *97*, 062316. [[CrossRef](#)] [[PubMed](#)]
101. Tang, J.; Qu, M.; Wang, M.; Zhang, M.; Yan, J.; Mei, Q. LINE: Large-scale information network embedding. In Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18–22 May 2015; pp. 1067–1077.
102. Grover, A.; Leskovec, J. Node2vec: Scalable feature learning for networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 855–864.
103. Cao, R.M.; Liu, S.Y.; Xu, X.K. Network embedding for link prediction: The pitfall and improvement. *Chaos Interdiscip. J. Nonlinear Sci.* **2019**, *29*, 103102. [[CrossRef](#)]
104. He, Y.L.; Liu, J.N.; Hu, Y.x.; Wang, X.z. OWA operator based link prediction ensemble for social network. *Expert Syst. Appl.* **2015**, *42*, 21–50. [[CrossRef](#)]

105. Li, K.; Tu, L.; Chai, L. Ensemble-model-based link prediction of complex networks. *Comput. Netw.* **2020**, *166*, 106978. [[CrossRef](#)]
106. Ghasemian, A.; Hosseinmardi, H.; Galstyan, A.; Airolidi, E.M.; Clauset, A. Stacking models for nearly optimal link prediction in complex networks. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 23393–23400. [[CrossRef](#)]
107. Zhang, M.; Chen, Y. Link prediction based on graph neural networks. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, QC, Canada, 3–8 December 2018; pp. 5171–5181.
108. Cai, L.; Ji, S. A multi-scale approach for graph link prediction. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 3308–3315. [[CrossRef](#)]
109. Cai, L.; Li, J.; Wang, J.; Ji, S. Line graph neural networks for link prediction. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 5103–5113. [[CrossRef](#)]
110. Kumar, A.; Singh, S.S.; Singh, K.; Biswas, B. Link prediction techniques, applications, and performance: A survey. *Phys. A Stat. Mech. Appl.* **2020**, *553*, 124289. [[CrossRef](#)]
111. Zhou, T. Progresses and challenges in link prediction. *iScience* **2021**, *24*, 103217. [[CrossRef](#)]
112. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [[CrossRef](#)]
113. de Sá, H.R.; Prudêncio, R.B.C. Supervised link prediction in weighted networks. In Proceedings of the 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011; pp. 2281–2288.
114. Martinčić-Ipšić, S.; Močibob, E.; Perc, M. Link prediction on Twitter. *PLoS ONE* **2017**, *12*, e0181079. [[CrossRef](#)] [[PubMed](#)]
115. Wei, Y.M.; Kang, J.N.; Yu, B.Y.; Liao, H.; Du, Y.F. A dynamic forward-citation full path model for technology monitoring: An empirical study from shale gas industry. *Appl. Energy* **2017**, *205*, 769–780. [[CrossRef](#)]
116. Kang, J.N.; Wei, Y.M.; Liu, L.c.; Wang, J.W. Observing technology reserves of carbon capture and storage via patent data: Paving the way for carbon neutral. *Technol. Forecast. Soc. Chang.* **2021**, *171*, 120933. [[CrossRef](#)]
117. Kim, J.S.; Goh, K.I.; Kahng, B.; Kim, D. Fractality and self-similarity in scale-free networks. *New J. Phys.* **2007**, *9*, 177–177. [[CrossRef](#)]
118. Murata, T.; Moriyasu, S. Link prediction of social networks based on weighted proximity measures. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'07), Fremont, CA, USA, 2–5 November 2007; pp. 85–88.
119. Zhao, J.; Miao, L.; Yang, J.; Fang, H.; Zhang, Q.M.; Nie, M.; Holme, P.; Zhou, T. Prediction of links and weights in networks by reliable routes. *Sci. Rep.* **2015**, *5*, 12261. [[CrossRef](#)]
120. Kraskov, A.; Stögbauer, H.; Grassberger, P. Estimating mutual information. *Phys. Rev. E* **2004**, *69*, 066138. [[CrossRef](#)]
121. Ross, B.C. Mutual information between discrete and continuous data sets. *PLoS ONE* **2014**, *9*, e87357. [[CrossRef](#)]
122. Lin, Y.; Wang, P.; Ma, M. Intelligent transportation system(ITS): Concept, challenge and opportunity. In Proceedings of the 2017 IEEE 3rd International Conference on Big Data Security on Cloud (Bigdatasecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), Beijing, China, 26–28 May 2017; pp. 167–172.
123. Pan, M.; Zhou, Y.; Zhou, D.K. Comparing the innovation strategies of Chinese and European wind turbine firms through a patent lens. *Environ. Innov. Soc. Transit.* **2019**, *30*, 6–18. [[CrossRef](#)]
124. Wang, B.; Liu, Y.; Zhou, Y.; Wen, Z. Emerging nanogenerator technology in China: A review and forecast using integrating bibliometrics, patent analysis and technology roadmapping methods. *Nano Energy* **2018**, *46*, 322–330. [[CrossRef](#)]
125. Li, X.; Xie, Q.; Jiang, J.; Zhou, Y.; Huang, L. Identifying and monitoring the development trends of emerging technologies using patent analysis and Twitter data mining: The case of perovskite solar cell technology. *Technol. Forecast. Soc. Chang.* **2019**, *146*, 687–705. [[CrossRef](#)]
126. Xiong, Z.; Sheng, H.; Rong, W.; Cooper, D.E. Intelligent transportation systems for smart cities: A progress review. *Sci. China Inf. Sci.* **2012**, *55*, 2908–2914. [[CrossRef](#)]
127. Bishop, R. The potential for Vehicle-highway automation in the United States. In Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robot and Systems. Innovative Robotics for Real-World Applications. IROS '97, Grenoble, France, 11 September 1997; Vol. 3, pp. 1215–1216.
128. Tanaka, T.; Yamamoto, T.; Fukuda, S. Onboard system devices for a vehicle information and communication system. *Fujitsu Ten Tech. J.* **1995**, *7*, 26–34.
129. Auer, A.; Feese, S.; Lockwood, S.; Hamilton, B.A. *History of Intelligent Transportation Systems*; Technical Report FHWA-JPO-16-329; Department of Transportation, Intelligent Transportation Systems Joint Program Office: Washington, DC, USA, 2016.
130. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [[CrossRef](#)]
131. Lü, L.; Zhou, T. *Link Prediction*; Higher Education Press: Beijing, China, 2013.
132. Parastesh, F.; Mehrabbeik, M.; Rajagopal, K.; Jafari, S.; Perc, M. Synchronization in Hindmarsh–Rose neurons subject to higher-order interactions. *Chaos Interdiscip. J. Nonlinear Sci.* **2022**, *32*, 013125. [[CrossRef](#)]
133. Kumar, A.; Chowdhary, S.; Capraro, V.; Perc, M. Evolution of honesty in higher-order social networks. *Phys. Rev. E* **2021**, *104*, 054308. [[CrossRef](#)]
134. Alvarez-Rodriguez, U.; Battiston, F.; de Arruda, G.F.; Moreno, Y.; Perc, M.; Latora, V. Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.* **2021**, *5*, 586–595. [[CrossRef](#)]
135. Majhi, S.; Perc, M.; Ghosh, D. Dynamics on higher-order networks: A review. *J. R. Soc. Interface* **2022**, *19*, 20220043. [[CrossRef](#)]
136. Fan, J.; Yin, Q.; Xia, C.; Perc, M. Epidemics on multilayer simplicial complexes. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **2022**, *478*, 20220059. [[CrossRef](#)]

137. Sun, Q.; Wang, Z.; Zhao, D.; Xia, C.; Perc, M. Diffusion of resources and their impact on epidemic spreading in multilayer networks with simplicial complexes. *Chaos Solitons Fractals* **2022**, *164*, 112734. [[CrossRef](#)]
138. Ángeles Serrano, M.; Boguñá, M. Tuning clustering in random networks with arbitrary degree distributions. *Phys. Rev. E* **2005**, *72*, 036133. [[CrossRef](#)] [[PubMed](#)]
139. Onnela, J.P.; Saramäki, J.; Kertész, J.; Kaski, K. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E* **2005**, *71*, 065103. [[CrossRef](#)] [[PubMed](#)]
140. Pusch, A.; Weber, S.; Porto, M. Generating random networks with given degree-degree correlations and degree-dependent clustering. *Phys. Rev. E* **2008**, *77*, 017101. [[CrossRef](#)] [[PubMed](#)]
141. Barrat, A.; Barthélemy, M.; Pastor-Satorras, R.; Vespignani, A. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3747–3752. [[CrossRef](#)]
142. Guimerà, R.; Sales-Pardo, M.; Amaral, L.A.N. Classes of complex networks defined by role-to-role connectivity profiles. *Nat. Phys.* **2007**, *3*, 63–69. [[CrossRef](#)]
143. Clauset, A.; Newman, M.E.J.; Moore, C. Finding community structure in very large networks. *Phys. Rev. E* **2004**, *70*, 066111. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.