Journal of Informetrics 19 (2025) 101632

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Integrating persistence process into the analysis of technology convergence using STERGM

Guancan Yang [a], Di Liu [a], Ling Chen [a], Kun Lu [b,*]

[a] School of Information Resource Management, Renmin University of China, No. 59 Zhongguancun Street, Haidian District, Beijing 100872, PR China
[b] School of Library and Information Studies, University of Oklahoma, 401 West Brooks St., Norman, OK 73019, United States

ARTICLE INFO

Keywords:
Technology convergence
STERGM
Formation process
Persistence process
Predictive performance

ABSTRACT

Understanding the dynamics of technology convergence is indispensable for both academic and industrial perspectives. Traditional analyses have mainly focused on the link formation process, overlooking the role that persistence process plays in shaping technology networks. This paper endeavors to fill this gap by incorporating the persistence process into the analysis of technology convergence using the *Separate Temporal Exponential Random Graph Model* (STERGM). Utilizing a decade-long dataset of breast cancer drug patents, we provide a comprehensive view of technology convergence mechanisms and their predictive capabilities. Our findings reveal significant differences in network effects between formation and persistence processes, indicating that focusing on only one may misrepresent the evolution of technology networks. The combined model achieves an F1 score of 69.54% in empirical forecasting, confirming its practical utility. Additionally, we introduce Intensification Networks to examine how existing ties strengthen or weaken over time, uncovering the critical role of intensification in the long-term evolution of technology convergence. By capturing both the formation of new ties and the intensification of existing ones, our model offers a more nuanced and forward-looking understanding of convergence dynamics, particularly in identifying potential areas for future technology convergence.

1. Introduction

Technology convergence integrates key technologies from various fields to address complex issues, acting as a major driver for future industrial and technological growth. This topic has attracted academics in technology management, policy-making, and competitive intelligence. As a catalyst for innovation, convergence fosters emerging and disruptive technologies. In innovation management, understanding convergence guides technology development and informs research investment. For science and technology policymakers, early detection allows adaptation of policy tools to match the current pace of industrial and technology development (Karvonen et al., 2010). Grasping the mechanisms underlying convergence is vital for its reliable prediction.

Previous research has examined methods for the early identification and prediction of converging technologies. A common approach constructs an indicator system to characterize these multidimensional technologies from a measurement perspective. In this context, scholars have developed a range of patent-based indicators, often using IPC (International Patent Classification) systems to assess the degree of convergence (No & Park, 2010). Some studies have further extended their methodologies to incorporate IPC-level

* Corresponding author.
E-mail address: kunlu@ou.edu (K. Lu).

https://doi.org/10.1016/j.joi.2024.101632
Received 13 October 2023; Received in revised form 11 November 2024; Accepted 12 December 2024
Available online 21 December 2024

networks and specialized concepts such as entropy and gravity as additional metrics for identifying technological convergence (Han & Sohn, 2016). Moreover, recent advancements in this field have introduced two types of indexes—fluctuation and continuity—to provide a more nuanced analysis of the dynamic patterns of technological convergence (Yun & Geum, 2019).

Network analysis is increasingly recognized as a pivotal approach to understanding technology convergence, notably through the adoption of the 'technology network' concept (Kogler, Rigby, & Tucker, 2013; Broekel, Balland, Burger, & Van Oort, 2014; Juhász, Broekel, & Boschma, 2020). These networks typically use patent classification codes as nodes and their co-occurrences as edges. These codes represent different technologies and the trends of co-occurrence relationships reflect their convergence. Link prediction, a common method in this context, helps forecast new connections within technology networks, potentially identifying future paths for technological convergence (Kim & Sohn, 2020; Oh et al., 2020). However, link prediction does not fully capture the complexity of technology convergence, especially in multi-technology contexts and evolving node dynamics (Kim & Lee, 2017; Lee et al., 2021; Wang & Hsu, 2023). While existing studies have acknowledged these gaps and proposed solutions, link dynamics remains an under-explored area.

Link dynamics, also known as tie dynamics, represent the temporal changes in network connections—specifically formation, persistence, and dissolution (Dahlander & McFarland, 2013; Krivitsky & Handcock, 2014). Formation refers to the initial establishment of a link between two technologies, marking the beginning of convergence (Youn et al., 2015; Broekel & Bednarz, 2018). Persistence denotes the ongoing interaction between technologies over time, reflecting sustained convergence and the stability of their relationship (Cantner & Graf, 2006; Copeland et al., 2023). Dissolution signifies the reduction or cessation of these connections due to factors such as market shifts, technological advancements, or strategic decisions (Koka et al., 2006; Ter Wal, 2014). While these concepts are well-established in network dynamics literature, they have been less explored in the context of technology convergence. Moreover, the intensification of relationships between technologies—that is, the strengthening of existing ties—is crucial for understanding technology evolution (Juhász et al., 2020).

Incorporating all facets of link dynamics offers a more nuanced understanding of technological evolution. This approach emphasizes long-term stability by considering link emergence, maturation, and decline. Relying exclusively on link formation overlooks key dynamics affecting the technology lifecycle. Powell et al. (2005) and Yan et al. (2016) highlight the need to explore the temporal evolution of technological alliances.

To deepen our understanding of technology convergence, we integrate persistence into the analysis using Separable Temporal Exponential Random Graph Models (STERGM) (Krivitsky & Goodreau, 2019). This allows us to assess whether the mechanisms of persistence mirror those of formation or differ significantly. As Krivitsky and Handcock (2014, p. 36) noted, "processes for forming and dissolving ties differ," suggesting that the dynamics of formation, persistence, and dissolution each require unique consideration in the broader context of technological evolution. This study aims to explore these dynamics and their implications for predicting technology convergence.

RQ1: What factors influence the dynamics of technology convergence in both formation and persistence? Are the mechanisms for formation and persistence the same or different?

This question seeks to identify the variables influencing both the formation and persistence of technology ties. It assesses whether the factors driving the formation of new ties differ from those sustaining existing ties. Understanding these mechanisms offers a more comprehensive view of technology convergence.

RQ2: Does integrating persistence mechanisms enhance the predictive power for technology convergence?

This question assesses whether models incorporating both formation and persistence processes have greater predictive power than those considering only formation. By incorporating persistence mechanisms, this research seeks to improve the accuracy and reliability of predicting technology convergence.

The remainder of this paper is organized as follows: the second section reviews pertinent studies; the third section outlines the research framework, methodologies, effect hypothesis selection, network terms, and model assessment; the fourth section presents an empirical analysis, which includes network data preprocessing, descriptive statistics, network and model construction, goodness-of-fit evaluation, and interpretation; and the final section offers conclusions and discusses the limitations of this study.

2. Literature review

2.1. Technology convergence

Several related concepts of technology convergence have been used in literature, and sometimes interchangeably. Technology convergence generally refers to the gradual merging of distinct technological domains to create new systems and innovations (Curran et al., 2010), often reshaping industries and creating new markets (Bröring et al., 2006; Kim et al., 2015). In contrast, technology fusion involves integrating technological elements within an existing domain rather than creating a new one (Curran & Leker, 2011). Another related concept is technological recombination, which entails combining existing technologies or knowledge components in novel ways to create innovations (Weitzman, 1998; Fleming, 2001). Computational models have been used to study these related concepts (Xiao et al., 2022, Cao et al., 2023), although this study focuses on technology convergence.

Technology convergence research has evolved into three main areas:

Exploration of technology convergence drivers: Early research focused on internal mechanisms and external drivers of technology convergence. Internally, it's viewed as a recombination and search process based on prior technological coupling, leading to the "technology recombination theory" (Fleming, 2001). Externally, factors like technology itself, market demands, and regulation influence convergence, giving rise to theories like the "co-evolution theory" and "convergence chain theory" (Hacklin et al., 2009).

However, much of this work remained theoretical or case-based due to the infancy of quantitative analysis tools (Youtie & Shapira, 2008; Jeong et al., 2012).

Advancements in technology convergence pattern recognition: With more sophisticated quantitative tools, researchers began examining and measuring technology convergence patterns. Methods such as association rules (Lee et al., 2015), information entropy (Cho & Kim, 2014), intensity map analysis (Geum et al., 2012), and evolution patterns (Choi et al., 2015) were introduced. Despite various methods, a unified framework was lacking, and progress in robust recognition models was limited. Research often focused on growth trends over time without fully exploring underlying dynamics and mechanisms (Sick & Bröring, 2022).

Emergence of technology convergence prediction models: More recently, with the rise of big data and artificial intelligence, researchers have turned to using these tools to enhance the predictive capabilities of technology convergence in dynamic competitive environments (Ibrahim, Elamer, & Ezat, 2021). Prediction models, especially those based on machine learning techniques like link prediction (San Kim & Sohn, 2020; Choi, Afifuddin, & Seo, 2022), have been employed to anticipate trends in technology convergence. However, it is important to note that link prediction and technology convergence prediction are not completely synonymous. Link prediction typically forecasts the linkage between two technology elements, while technology convergence often involves the recombination and linkage of multiple elements (Kim & Lee, 2017; Lee et al., 2021; Wang & Hsu, 2023). Additionally, node dynamics play a crucial role, as the structure of the network evolves over time, affecting the potential for technology convergence (Oh, et al., 2020; Yang, et al., 2024). Beyond the binary relationships, the intensification of relationships between technology elements also warrants attention, as highlighted by Juhász et al. (2020). Thus, while link prediction is a useful tool, further refinement is needed to fully capture the complexity of technology convergence in dynamic environments.

2.2. *Formation and persistence processes*

Scholars studying technology convergence have increasingly recognized the intricate interdependencies inherent in the formation of convergent technologies. Rather than developing in isolation, technologies are understood to interact and occasionally merge with other technologies, thus sparking innovation (Fleming & Sorenson, 2001). Consequently, several studies have proposed that the extent of technology convergence can be measured by examining the number of technological classes cited outside a patent's core technological area (Verhoeven, Bakker, & Veugelers, 2016). This process is frequently depicted as a dynamic interaction between diverse technological components, giving rise to innovative technological configurations (Tang, et al., 2020). Concurrently, the concept of convergence chain is understood as the blurring of boundaries between previously distinct areas of science, technology, markets, or industries (Hacklin et al., 2009).

In the context of network dynamics, formation and persistence should be understood as two distinct processes, driven by different underlying mechanisms. Formation of technological links often results from short-term needs, such as rapid innovation, regulatory shifts, or competitive pressure. On the other hand, persistence refers to the long-term viability and sustained relevance of these links, often linked to a technology's market dominance or technological maturity. Persistence differentiates truly converging technologies from those in temporary flux. This distinction is crucial because convergence is not merely about the initial formation of links but about their ability to endure and contribute to long-term technological advancements (Statnet Development Team, 2021).

Tie persistence and dissolution serve as mutually exclusive yet complementary aspects, broadening our view beyond just tie formation. Formation alone does not confirm convergence; persistence is essential to verify that these technological interactions result in meaningful and sustained innovation. Persistence differentiates stable technologies from those in flux. Without sustained interaction, links are unlikely to drive true convergence or have lasting impacts on innovation. Industry models like Gartner's S-curve and the Hype curve also emphasize these dynamics, suggesting that technologies experience phases of hype, decline, and eventual stabilization (Fenn & Raskino, 2008).

2.3. *ERGM for analyzing technology networks*

Technology convergence networks capture the dynamic nature of technology convergence, highlighting complex interactions among technologies. Network modeling clarifies this by revealing new connections, dissolving old ones, and tracking network evolution over time. Exponential Random Graph Models (ERGMs) offers a way to understand complex system dependencies via simulation (Handcock et al., 2008). Unlike traditional methods, ERGM can handle mixed variables in complex networks and provides network-wide insights into technology convergence. The parameters derived from these models can also be predictive. In recent years, ERGM has received attention from the fields of scientometric and intelligence analysis (Peng, 2015; An & Ding, 2018; Zhang, et al., 2018).

In contrast, the STERGM is specifically designed to analyze dynamic changes within networks. It separates the network evolution into two distinct processes: the formation of new ties and the dissolution of existing ones. This bifurcation allows STERGM to dynamically model network changes over time, offering a more nuanced understanding of how relationships in technology networks form and dissolve (Carnegie et al., 2015). Empirical evidence, including comparisons by Fritz, Lebacher, & Kauermann (2020), suggests that STERGMs perform better than other models in contexts similar to ours, reinforcing our methodological choice. STERGM has been applied across various fields like international politics, social media, natural resource governance, business operations, and pedagogy, yielding insights beneficial to those areas (Dokuka & Valeeva, 2015; Angst & Hirschi, 2017; Lebacher, Thurner, & Kauermann, 2021; Xu, 2021). However, to our knowledge, there has been no study of the tie formation and dissolution mechanism in technology networks using the STERGM. This study aims to fill this gap.

3. Research design

3.1. Framework

In this study, we aim to elucidate the dynamic mechanisms underlying technology convergence, with a focus on predictive modeling. Our approach is structured into three main components, as illustrated in Fig. 1.

The first step involves data collection and preparation. We gather IPC pairs at the 4-digit level from a specific technological field and segment them based on their temporal occurrence to form distinct networks. Descriptive statistics are then calculated to reveal key patterns, ensuring the data is well-prepared for model building.

In the second step, we use STERGMs to construct models in two phases. We begin by selecting key mechanisms such as preferential attachment, triadic closure, network baseline effects, node attributes, and homophily. First, we model tie formation, followed by an extended model incorporating persistence. Sequentially, we compare models to assess how each mechanism influences formation and persistence. We then conduct a comparative analysis to observe parameter evolution between the two processes. At each stage, Goodness of Fit (GOF) and model evaluations ensure robustness.

The final step extends the model into predictive scenarios. We test its ability for premature prediction, forecasting future network dynamics. Additionally, we expand the STERGMs to focus on intensification networks, shifting from new tie formation to strengthening existing ties, demonstrating the model's ability to handle more complex scenarios over time.

3.2. STERGMs

ERGM models the probability distribution of a network Y based on specified network statistics $g(y)$ and associated parameters θ (Robins et al., 2007):

$$\Pr(Y = y) = \frac{exp(\theta \cdot g(y))}{Z(\theta)} \tag{1}$$

where $Z(\theta)$ is the normalizing constant ensuring a proper probability distribution.

STERGM extends this framework by modeling the network's evolution between discrete time points t and $t + 1$ as two separate processes: tie formation and tie dissolution. This separability allows us to independently examine factors influencing the creation of new ties and the persistence or dissolution of existing ties (Krivitsky & Handcock, 2014). The probability of tie formation is modeled as:

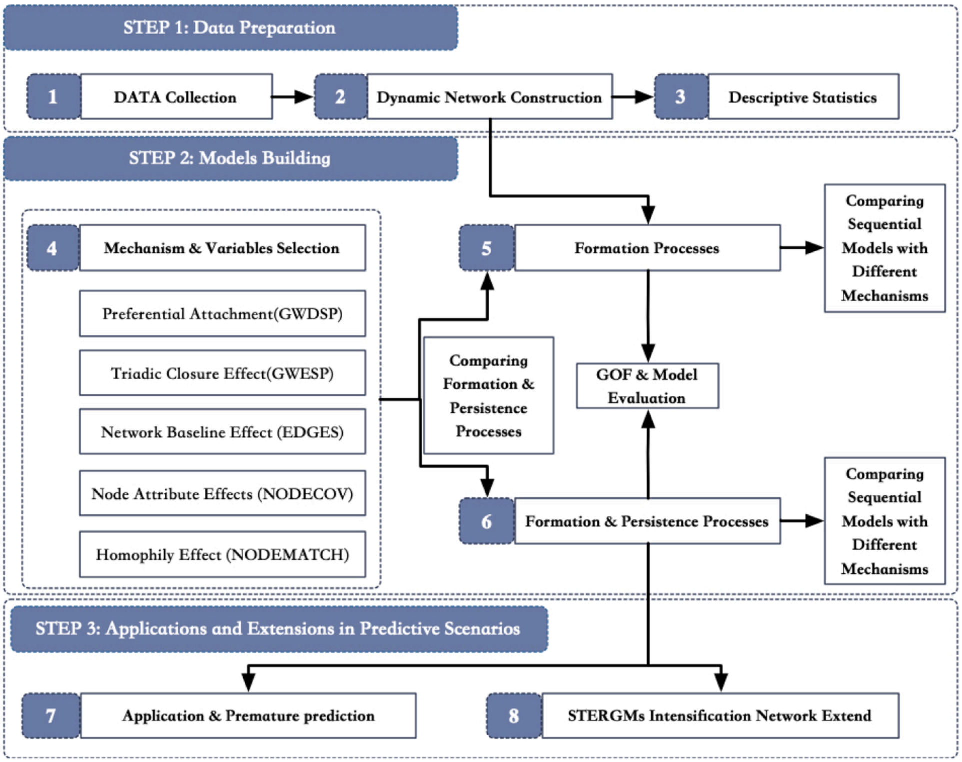


Fig. 1. Research Framework.

$$Pr(Y^{+} = y^{+} | Y^{t} = y^{t}; \theta^{+}) = \frac{exp(\theta^{+} \cdot g^{+}(y^{+}, y^{t}))}{Z^{+}(\theta^{+}, y^{t})} \tag{2}$$

Similarly, the probability of tie dissolution is:

$$Pr(Y^{-} = y^{-} | Y^{t} = y^{t}; \theta^{-}) = \frac{exp(\theta^{-} \cdot g^{-}(y^{-}, y^{t}))}{Z^{-}(\theta^{-}, y^{t})} \tag{3}$$

The network at time $t + 1$ is then given by:

$$Y^{t+1} = (Y^{t} \cup Y^{+}) \backslash Y^{-} \tag{4}$$

In these equations: Y^{+} represents the set of new ties formed between t and $t + 1$. Y^{-} represents the set of ties dissolved during the same period. A conceptual framework of tie formation and dissolution in STERGM can be found in Morris et al. (2015). In our context, tie formation (Y^{+}) corresponds to the emergence of new technological linkages—indicating areas where technologies are beginning to converge. Tie persistence (the complement of tie dissolution Y^{-}) reflects the continuation of existing technological relationships, highlighting sustained integration efforts.

Structural zero-completion. STERGM requires networks to have the same set of nodes across all time periods. However, the emergence of new nodes over time presents an analytical challenge (Xie et al., 2019). To address this, we use the maximum set of nodes over all time periods and augment each network by adding missing nodes for each specific time point. The issue is that these added nodes should not form links with existing nodes since they did not exist at that time.

To overcome this, we employ structural zeros to define zero entries in the network matrix, restricting links between nodes that logically cannot have links (Tsamardinos & Borboudakis, 2010). This approach allows the inclusion of variable node sets in dynamic network analysis while maintaining the network's integrity by preventing unrealistic links.

3.3. Rationale for using STERGM in technology convergence

STERGM's distinct advantage is that it captures both these aspects simultaneously, offering a comprehensive lens through which to examine the evolution of technology.

Capturing dynamic processes: Traditional models treat networks as static snapshots, failing to account for the evolving nature of technological connections. Unlike static ERGMs, STERGM incorporates both the formation and dissolution of links over time, capturing how technologies not only converge but also how connections persist or dissolve (Leifeld & Cranmer, 2019). This allows for a more realistic representation of technological evolution, considering the dynamic processes driven by market shifts, innovation cycles, and regulatory changes.

Differentiating formation and persistence drivers: Technology convergence is driven by distinct factors at different stages. STERGM helps differentiate these by isolating the drivers behind link formation and persistence. For instance, link formation may be influenced by the total number of connections (**EDGES**) between IPC code pairs, while persistence can be shaped by preferential attachment (**GWDSP**), where highly connected technologies attract more links (Lusher et al., 2013). Triadic closure (**GWESP**) enhances network clustering as technologies with common partners are more likely to form and sustain connections (Hunter, 2007). node attributes (**NODECOV, NODEMATCH**) also influence stability, as similar technologies tend to form long-lasting synergies.

Modeling complex network structures: STERGM is especially suited for analyzing technology convergence as it models complex network structures, incorporating both node attributes (e.g., technology size) and structural dependencies (e.g., homophily and triadic closure) (Robins et al., 2007). Unlike traditional models that assume link independence, STERGM allows for interdependencies between links, capturing the reality that technological collaborations often depend on existing connections within the network (Goodreau et al., 2009).

3.4. Model assessment

In statistical modeling, assessing the fit of the model is critically important as it ensures that the model is not only mathematically sound but also accurately reflects the characteristics of the observed data. There are several methods to evaluate the fit of statistical models. In this study, we will initially employ the **Akaike Information Criterion (AIC)** or the **Bayesian Information Criterion (BIC)** as the standard for evaluating the optimal model during multiple rounds of iterative model validation. Lower AIC and BIC values indicate a better fit and are used to compare one model with another to determine the best-fitting model. In statistical network modeling, it is often argued that suitable network models should not only provide reasonable predictions for added edges but also be able to represent the topologies of the observed network. This approach is based on sampling from the distribution of networks under the estimated parameters and then comparing the network properties of these sampled networks with those of the observed networks (Handcock et al., 2008). Subsequently, we use **Graphical Goodness-of-fit Diagnostics** to confirm how well the optimal model performs in fitting the actual network. Finally, in the practical prediction phase, we utilize **Precision-Recall (PR)** metrics to assess the model's performance in prediction.

4. Empirical analysis

Our analysis focuses on breast cancer treatment patents to highlight the sector's role in technology convergence. This domain merges diverse fields—biotechnology, pharmacology, medical devices, and IT—to craft more accurate and individualized therapies, illustrating the potential of technology convergence. Innovations like the integration of targeted therapy with chemotherapy, gene sequencing for customized treatments, and the employment of artificial intelligence in imaging and surgery underscore this convergence. These advancements, as reviewed by Chopra et al. (2023), showcase the amalgamation of imaging, genetics, surgery, and computational science, significantly enhancing breast cancer care.

4.1. Data

The patent data of the breast cancer field were obtained from the Derwent Innovations Index which indexes patents from 52 patent-issuing authorities from 1963 to the present. We refer to the PubMed database for the search query of MeSH subject terms related to "breast cancer" to enhance the search's comprehensiveness.[1] The most recent ten years of patent data related to breast cancer drugs from 2011 to 2020 were selected for this study (downloaded on March 20, 2022), and 39784 raw records were acquired. The search query is shown in Table 1.

The IPC is a standardized system for organizing and categorizing technological inventions and innovations. It is used by patent offices around the world to classify patents and patent applications, and is an important tool for searching and accessing information on patented technologies. Despite the development of the Cooperative Patent Classification (CPC), our choice to employ the IPC system is motivated by its extensive historical coverage and universal recognition, ensuring a consistent and comprehensive long-term analysis of patent trends crucial for our study. A full IPC code contains 12 digits, which can be divided into 5 levels, namely, section, class, subclass, main group, and subgroup. Admittedly, IPC codes at the subclass level primarily offer a broad overview of a technological field, which may be insufficient for capturing the specific nuances of individual technologies. However, as our research focus is technology convergence—namely, the relationships or combinations between different technological areas—the utility of subclass-level IPC codes becomes salient. Intersections and combinations among various subclasses can offer an understanding of which technological domains are more likely to converge. The use of 4-digit IPC codes (sub-class level) to analyze technology convergence is commonly adopted by other researchers (Kim & Sohn, 2020; Feng, et al., 2020).

Furthermore, we also considered using the 1-digit IPC code at the section level as a broad categorization for patents. To specifically examine the homophily mechanism, we selected 5 categories according to the breast cancer data to be treated as internal features of IPC codes: Class A (Human Necessities), Class B (Operations and Transport), Class C (Chemistry and Metallurgy), Class G (Physics), and Class H (Electricity).

4.2. Dynamic networks building

We harvested the data year-by-year and established multiple two-mode networks based on the "patent-IPC code" affiliations. These two-mode networks were subsequently projected into one-mode networks, resulting in the creation of multiple IPC code co-occurrence networks. Following this, we combined all these networks in chronological order into a network series. The co-occurrence relationships between nodes are manifested as edges, and their frequencies are denoted as weights. We set the weight threshold greater than one to mitigate random noise in constructing the networks.

4.3. Characteristics of dynamic networks

Table 2 presents the descriptive statistics of the network series. The table includes metrics such as node size, edge size, clustering coefficient, density, and mean degree. Overall, these statistical indicators do not exhibit a consistent upward or downward trend over time but remain relatively stable. Given that the network constructed is a dynamic network, a structural zero-completion operation is required for all nodes prior to employing STERGM. For the decade spanning 2011–2020, this yielded a full network node size of 155. The term "**NSIZE**" in Table 3 refers to the count of nodes present in each actual time slice.

4.4. Variables selection

One of the key advantages of the ERGMs and variations is their ability to incorporate multiple types of interactions and dependencies between variables. By adding factors to the model in a stepwise fashion, researchers will be able to identify which factors have the greatest impact on the formation and persistence of these technologies. Variables used in this article, as well as the mechanisms, network configurations, and interpretations they represent in the process of technology formation and persistence, are outlined in Table 3.

4.5. Goodness-of-fit diagnostics

Goodness-of-fit tests, using Graphical Goodness-of-fit diagnostics, are presented in Fig. 2 to validate the STERGM's representation of the actual technology convergence network. Metrics like degree, shared partners, and geodesic distance were examined across 5,000

[1] Breast Neoplasms. (2022). National Library of Medicine Website. Retrieved 17 July 2022, from https://www.ncbi.nlm.nih.gov/mesh/68001943.

Table 1
Search query and results.

Search Query	Time Interval	Results
Breast Neoplasm (Topic) or Breast Tumor? (Topic) or Breast Cancer (Topic) or Mammary Cancer? (Topic) or Malignant Neoplasm of Breast (Topic) or Breast Malignant Neoplasm? (Topic) or Malignant Tumor of Breast (Topic) or Breast Malignant Tumor? (Topic) or Cancer of Breast (Topic) or Cancer of the Breast (Topic) or Human Mammary Carcinoma (Topic) or Human Mammary Neoplasm? (Topic) or Breast Carcinoma? (Topic)	2011–2020	39784

Table 2
Descriptive statistics of the network series.

Time Interval	Edge size	Node size	Clustering coefficient	Density	Mean degree
2011	385	74	0.417	0.142	10.405
2012	363	79	0.364	0.117	9.189
2013	379	80	0.368	0.119	9.475
2014	353	76	0.397	0.123	9.289
2015	344	66	0.420	0.160	10.424
2016	370	81	0.409	0.114	9.135
2017	396	79	0.405	0.128	10.025
2018	461	90	0.366	0.115	10.244
2019	373	72	0.423	0.145	10.361
2020	449	87	0.408	0.120	10.321

simulated networks, based on Model 10 parameters. The box plot and solid black line in Fig. 2 closely align, indicating a robust model fit. Overall, the diagnostics confirm that our STERGM accurately captures the network's observed topological and attribute features.

4.6. Model results

4.6.1. STERGMs excluding persistence processes

Table 4 presents the results of the STERGMs that do not incorporate persistence processes, based on a consolidated 10-year network series. The primary objective is to iteratively construct a variety of STERGMs to investigate the influence of different variables on the formation of the technology network. At this juncture, only the tie formation process ('**Form**') is taken into account. The Markov Chain Monte Carlo Maximum Likelihood Estimation (MCMC-MLE) technique is utilized within the Statnet toolkit for parameter estimation. A marked improvement in model fit is indicated by a significant reduction in the AIC, dropping from 13782.292 to 12749.804. In addressing the model's structure, it has been deemed essential to integrate measures of triadic closure into the baseline model to serve as a control for the transition from two-mode to one-mode projections. As such, GWESP will be added as a control variable. In fact, the inclusion of GWESP accounts for the most significant decrease from the null model's AIC, but as it serves as a control variable, we focus on the impact of adding other variables to the model beyond this control. The most significant enhancement in model performance is attributed to incorporating the main effects model, which factors in the square root of eigenvector centrality, yielding a significant reduction in the AIC from 13782.292 to 12959.980. In the final analysis, Model 6 shows that most factors are statistically significant for the formation of converging technologies. However, variables such as **NODEMATCH (A-A), NODEMATCH (G-G), NODECOV (NSIZE)**, and **ABSDIFF (NSIZE)** do not reach statistical significance, indicating their lesser influence in this context.

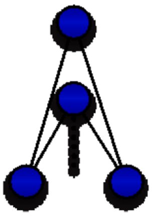
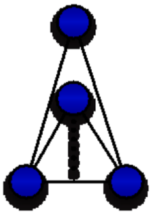
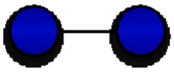
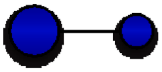
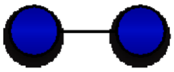
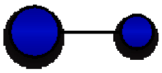
Model 1 serves as the baseline, incorporating **GWESP** as a control for triadic closure. The **EDGES** term is negative and statistically significant, indicating that as the number of edges increases, the likelihood of additional tie formation decreases—a common feature of sparse networks, where link formation becomes less likely as the network grows denser.

In Models 2 and 3, node-level attributes **NODECOV (Eigenvector Centrality)** and **ABSDIFF (Eigenvector Centrality)** are introduced. **NODECOV** shows a significant positive effect, highlighting those technologies with higher centrality are more likely to form new ties. Interestingly, **ABSDIFF** also has a positive effect, indicating that links can form even between nodes with substantial differences in centrality, challenging the expectation that similar centrality drives new connections. Both **NODECOV (NSIZE)** and **ABSDIFF (NSIZE)** are non-significant, suggesting that node size has little influence on the formation of ties.

Model 4 introduces homophily effects, with significant positive results for **NODEMATCH (A-A)** and **NODEMATCH (B-B)**, suggesting that technologies within the same IPC sections are more likely to form ties. In Model 5, **NODEMATCH (C-C)** is significant, indicating that homophily in section C also drives new connections. However, in Model 6, the significance of **NODEMATCH (C-C)** diminishes when **GWESP** and **GWDSP** are added, suggesting that homophily in section C is less influential when more comprehensive network effects are considered.

Finally, Model 6 incorporates both **GWESP** and **GWDSP** to capture triadic closure and preferential attachment. Both variables are positive and statistically significant, showing that IPC codes with higher centrality tend to form new links with similarly central technologies, reinforcing clustering and centrality dynamics in the network. The α value for these terms was determined through multiple tests, and log (4) was found to optimize the model's performance, as reflected in the lowest AIC.

Table 3
Main Variables and Their Roles in the STERGM Models.

Variables	Mechanism	Configurations	Meaning	References
EDGES	***Network Baseline Effect***		Represents the overall number of connections in the network, calculated as the sum of all connections between IPC code pairs.	Lusher et al., 2013
GWDSP	***Degree-Based Preferential Attachment***		Quantifies an IPC code's popularity within the network, calculated as the total number of edges connected to that specific technology, reflecting the "rich-get-richer" dynamic.	Jones & Handcock, 2003; Abbasi et al., 2012
GWESP	***Triadic Closure Effect***		Measures the strength of connections between nodes, calculated as the number of common neighbors (shared partners), reflecting network clustering or cohesiveness.	Lu & Li, 2023; Losacker, 2022
NODECOV (Eigenvector Centrality)	***Node Attribute Main Effect (Continuous)***		Measures node importance based on its connections to other well-connected nodes, reflecting a node's overall influence in the network.	Li et al., 2022
ABSDIFF (Eigenvector Centrality)	***Node Attribute Main Effect (Continuous)***		Quantifies the attribute-based difference between nodes, calculated as the absolute difference in eigenvector centrality, indicating variations in node influence within the network.	Ma et al., 2022
NODECOV (NSIZE)	***Node Attribute Main Effect (Continuous)***		Reflects the size of the node, often measured by the total number of patents or technology outputs associated with the node (IPC code), indicating its activity level.	Li et al., 2022
ABSDIFF (NSIZE)	***Node Attribute Main Effect (Continuous)***		Quantifies the difference in node size between two IPC codes, measured by the absolute value of their size difference, highlighting disparities in technology output.	Ma et al., 2022
NODEMATCH (IPC code at section level)	***Homophily Effect Based on IPC Section Level***		Measures the likelihood of forming connections between technologies with the same IPC section-level classification, reflecting the principle of homophily (technology similarity).	Chakraborty et al., 2020; Peng, 2015

4.6.2. STERGMs including persistence processes

To gain a comprehensive understanding of the role of persistence in the evolution of technology convergence networks, we chose to fix the tie formation process ('**Form**') and step wisely introduce factors associated with the network persistence process ('**Persist**'). As depicted in Table 5, each model considers both '**Form**' and '**Persist**' processes. Building on Model 6, Models 7 through 12 introduce variables related to network persistence, aiming to evaluate its influence on the evolution of technology convergence while accounting for formation effects. **Model 7** uses **GWESP** as a control for triadic closure and focuses on the **EDGES** feature to establish a baseline for how other variables impact the model.

NODECOV (Eigenvector Centrality) remains significant in the formation process, indicating that high-centrality nodes are crucial for generating new ties. **ABSDIFF (Eigenvector Centrality)** similarly suggests that smaller centrality differences facilitate link formation. However, in the persistence process (Models 8–12), **NODECOV** weakens slightly, and **ABSDIFF** reverses its effect, with larger centrality differences hindering the sustainability of ties. This shift highlights that, while centrality is critical for both formation and persistence, disparities in node importance negatively impact long-term network stability.

Both **NODECOV (NSIZE)** and **ABSDIFF (NSIZE)** are non-significant in the formation process, indicating that node size does not significantly affect the creation of ties. However, in persistence, **ABSDIFF (NSIZE)** shows a significant negative impact in Models 11 and 12, suggesting that larger size differences between technologies reduce the likelihood of sustaining ties.

NODEMATCH variables also display shifts between formation and persistence. Homophily in categories A, B, and C plays a significant role in tie formation in Model 7, but persistence dynamics evolve as additional variables are introduced. **NODEMATCH (C-C)** retains significance across models, showing a robust homophily effect even with the inclusion of more complex variables like **preferential attachment**. In contrast, **NODEMATCH (A-A)** and **NODEMATCH (B-B)** fluctuate, indicating that homophily effects in these

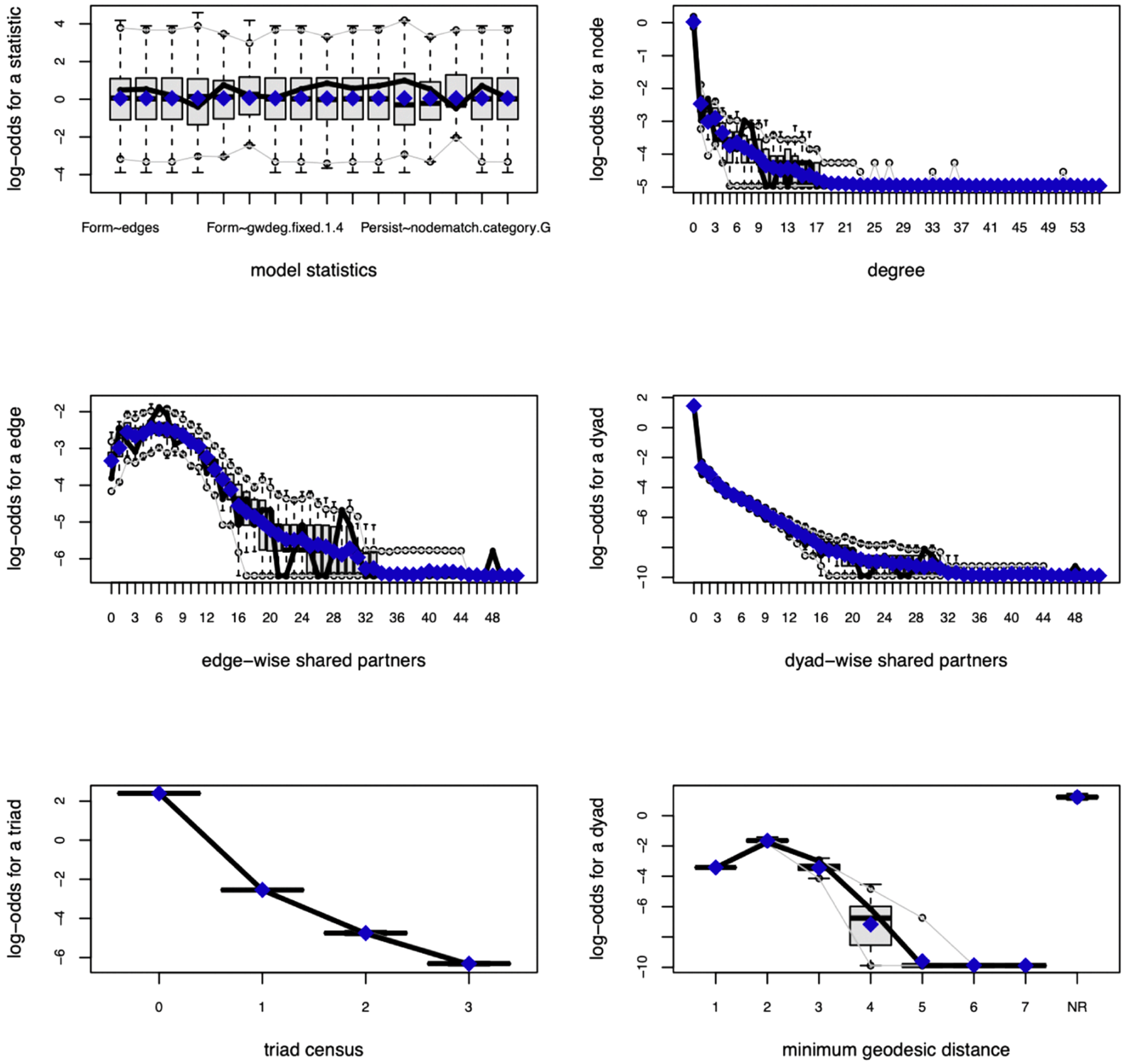


Fig. 2. The Graphical Goodness-of-fit diagnostics for STERGM with persistence process.

classes are more sensitive to other factors and less stable in persistence.

Finally, **GWDSP** remains significant in the persistence process, underscoring the importance of preferential attachment in maintaining existing links. This suggests that technologies with common neighbors continue to maintain their relationships over time, highlighting the key role of preferential attachment in both forming and sustaining ties.

4.6.3. *Parameter comparison between formation and persistence models*

In this section, we focus on the differences between the formation and persistence processes in Model 12, examining how various factors uniquely drive the creation of new ties and the maintenance of existing ones.

First, **NODECOV (Eigenvector Centrality)** is significant and positive in both processes, but its effect is stronger in formation (3.328 for formation vs. 1.174 for persistence). This suggests that highly central technologies play a more critical role in forming new ties than in maintaining them. Conversely, **ABSDIFF (Eigenvector Centrality)** negatively impacts formation but positively influences persistence, indicating that greater differences in centrality hinder the creation of ties while supporting their long-term maintenance. **NODECOV (NSIZE)** is non-significant for tie formation but has a significant positive effect on persistence (0.002, $p < 0.01$), showing that larger technologies are more likely to sustain existing connections.

Second, the **NODEMATCH** homophily effects vary between formation and persistence. Class C has a strong homophily effect in persistence (0.661), indicating that technologies within the same IPC section are likely to maintain ties over time. While Class C is also positive in formation, its influence is less pronounced. Class A shows a strong homophily effect in tie formation but loses significance in persistence, implying that while ties form easily within Class A, they are less likely to persist over time.

Table 4
Model fit of STERGMs for the technology convergence formation process.

Model terms	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
EDGES	**Form**	**Form**	**Form**	**Form**	**Form**	**Form**
	-6.378*** (0.067)	**-9.426***** (0.174)	**-6.340***** (0.069)	**-6.400***** (0.072)	**-9.457***** (0.179)	**-8.860***** (0.163)
GWESP	**0.910***** (0.018)	**0.621***** (0.023)	**0.907***** (0.019)	**0.901***** (0.019)	**0.620***** (0.022)	**0.830***** (0.028)
NODECOV (Eigenvector Centrality)		**5.551***** (0.247)			**5.553***** (0.241)	**3.326***** (0.242)
ABSDIFF (Eigenvector Centrality)		**4.151***** (0.231)			**4.132***** (0.230)	**0.906**** (0.287)
NODECOV (NSIZE)			-0.000 (0.000)		-0.000 (0.000)	-0.001 (0.001)
ABSDIFF (NSIZE)			0.000 (0.000)		0.000 (0.001)	0.001 (0.001)
NODEMATCH (A-A)				**0.206**** (0.073)	**0.198**** (0.075)	0.147 (0.077)
NODEMATCH (B-B)				**0.307*** (0.132)	0.154 (0.148)	**0.281*** (0.140)
NODEMATCH (C-C)				0.054 (0.155)	**0.376*** (0.160)	**0.347*** (0.157)
NODEMATCH (G-G)				-1.278 (0.720)	0.346 (0.737)	0.040 (0.717)
GWDSP						**0.045***** (0.002)
AIC	13782.292	12959.980	13752.486	13771.959	12961.589	12749.804
BIC	13801.461	12998.318	13790.824	13829.465	13057.433	12855.233

* Indicates significance at the 90% level;
** indicates significance at the 95% level;
*** indicates significance at the 99% level.

Finally, both **GWDSP** and **GWESP** are consistently significant across formation and persistence. **GWDSP** shows that highly connected technologies continue to attract new ties, while **GWESP** emphasizes that clustered ties are more likely to be maintained. These results highlight the importance of preferential attachment and transitivity in both forming and sustaining ties, underscoring their role in shaping the long-term structure of technology networks.

4.7. Model comparison

Moreover, Fig. 3 showcases a comparative analysis of the PR curves for different STERGM configurations, contrasting models that solely focus on the formation process against those that integrate both formation and persistence processes. These specific STERGMs were formulated using data from 2011-2019 and were evaluated against the 2020 data, serving as the observation network. The evaluation metrics reflect a notable enhancement in model fit when employing STERGMs that incorporate both formation and persistence processes (PR AUC 0.742), as opposed to models focusing solely on formation processes (PR AUC 0.665). The respective PR curves are displayed in Fig. 3: the left panel delineates the formation-only process, whereas the right panel encapsulates both the formation and persistence processes. This distinction underscores the superior fitting performance achieved by the integrated STERGMs, showcasing a more comprehensive understanding of network dynamics.

The STERGM model, developed with 2011-2019 data, forecasted the probability of technology convergence between various category pairs. For evaluation, pairs with a predicted probability greater than 0.5 were retained, totaling 316 predicted convergences. The 2020 data documented 449 instances of technology convergence. To gauge the model's accuracy, a confusion matrix (Table 6) was generated:

The STERGM model's performance metrics, when evaluated against the actual 2020 data, indicated an accuracy rate of approximately 98.05%. Its recall (sensitivity) was around 59.24%, signifying its ability to capture a significant fraction of actual technology convergences. The precision rate was approximately 84.18%, indicating that a considerable percentage of its predictions were accurate. Furthermore, with an F1-score of around 69.54%, the model exhibited a balanced performance between precision and recall, highlighting the robust predictive accuracy of the STERGM methodology.

4.8. Discussion on premature predictions

Predicting technology convergence has always been a challenging endeavor. STERGM emerges as a powerful tool, excelling in capturing and interpreting key mechanisms within networks, especially for promising technology pairs. However, every model has its limitations. For disruptive and burst technology convergences, STERGM might not provide precise predictions. An example is the breast cancer domain in 2020, where several high-probability predictions did not materialize. These **"premature predictions"** met initial conditions for convergence but had not yet converged, rather than indicating flaws in the model.

Table 5
Model fit of STERGMs for the technology convergence including persistence.

Model terms	Model 7		Model 8		Model 9		Model 10		Model 11		Model 12	
	Form	Persist	Form	Persist	Form	Persist	Form	Persist	Form	Persist	Form	Persist
EDGES	**-8.847*****	**-1.551*****	**-8.852*****	**-2.976*****	**-8.857*****	**-1.614*****	**-8.852*****	**-1.729*****	**-8.854*****	**-3.219*****	**-8.855*****	**-2.852*****
	(0.164)	(0.082)	(0.158)	(0.246)	(0.160)	(0.084)	(0.164)	(0.091)	(0.163)	(0.257)	(0.164)	(0.257)
GWESP	**0.829*****	**0.613*****	**0.831*****	**0.508*****	**0.831*****	**0.619*****	**0.830*****	**0.623*****	**0.829*****	**0.525*****	**0.829*****	**0.651*****
	(0.028)	(0.022)	(0.027)	(0.028)	(0.028)	(0.022)	(0.027)	(0.022)	(0.027)	(0.029)	(0.027)	(0.038)
NODECOV (Eigenvector Centrality)	**3.316*****		**3.303*****	**2.519*****	**3.314*****		**3.315*****		**3.325*****	**2.578*****	**3.328*****	**1.174****
	(0.242)		(0.239)	(0.301)	(0.244)		(0.249)		(0.247)	(0.314)	(0.246)	(0.379)
ABSDIFF (Eigenvector Centrality)	**0.908****		**0.904****	0.399	**0.901****		**0.898****		**0.904****	0.440	**0.912****	**-1.712*****
	(0.283)		(0.288)	(0.336)	(0.279)		(0.287)		(0.284)	(0.337)	(0.285)	(0.504)
NODECOV (NSIZE)	-0.001		-0.001		-0.001	**0.003****	-0.001		-0.001	**0.002****	-0.001	**0.002***
	(0.001)		(0.001)		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
ABSDIFF (NSIZE)	0.001		0.001		0.001	**-0.003****	0.001		0.001	**-0.002****	0.001	**-0.002***
	(0.001)		(0.001)		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)
NODEMATCH (A-A)	**0.148***		0.149		**0.148***		0.149	**0.394*****	0.148	**0.260****	**0.150***	**0.204***
	(0.075)		(0.078)		(0.075)		(0.077)	(0.091)	(0.077)	(0.098)	(0.076)	(0.097)
NODEMATCH (B-B)	**0.278***		**0.285***		0.276		**0.283***	0.194	**0.283***	0.019	0.279	0.097
	(0.139)		(0.145)		(0.144)		(0.142)	(0.165)	(0.142)	(0.165)	(0.144)	(0.170)
NODEMATCH (C-C)	**0.347***		**0.346***		**0.350***		**0.351***	**0.621****	**0.343***	**0.661*****	**0.347***	**0.661*****
	(0.160)		(0.155)		(0.162)		(0.159)	(0.191)	(0.157)	(0.192)	(0.158)	(0.194)
GWDSP	**0.045*****		**0.045*****		**0.045*****		**0.045*****		**0.045*****		**0.045*****	**0.042*****
	(0.002)		(0.002)		(0.002)		(0.002)		(0.002)		(0.002)	(0.007)
AIC	11243.350		11128.880		11217.825		11198.990		11124.721		11085.455	
BIC	11358.363		11263.062		11352.007		11342.757		11306.826		11277.144	

* Indicates significance at the 90% level;
** indicates significance at the 95% level;
*** indicates significance at the 99% level.

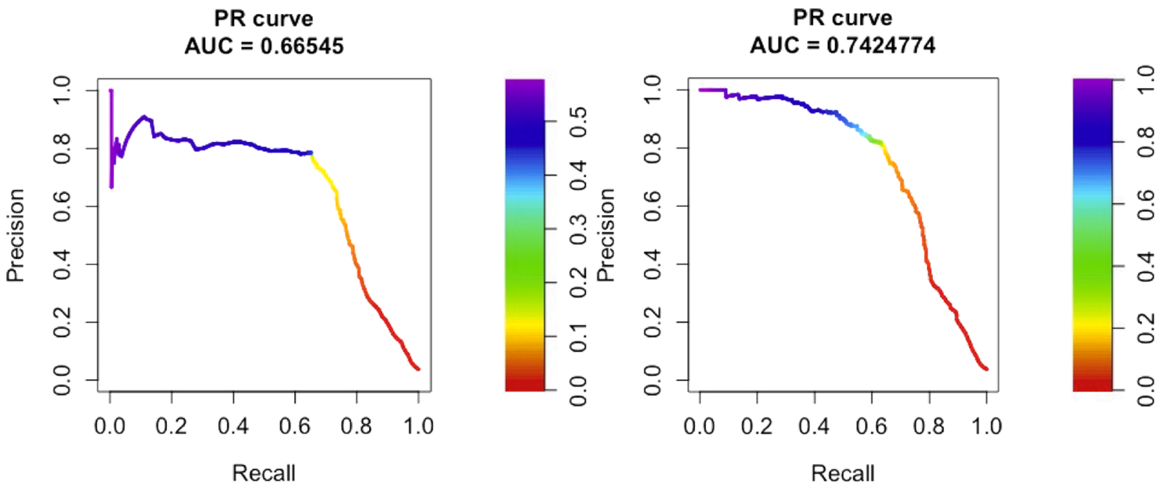


Fig. 3. Comparative analysis of the PR curves for different STERGM configurations.

Table 6
Confusion matrix for STERGM model predictions vs. actual outcomes.

		Predicted Outcomes	
		Predicted: Yes	Predicted: No
Actual Outcomes	**Actual: Yes**	266	183
	Actual: No	50	11436

We analyzed the predictive outcomes of the STERGM model with a focus on those premature predictions in the 2020 data. The Table 7 lists a subset of predictions, including IPC code pairs, predicted convergence probabilities, actual convergence in 2020, prediction categorization, and patent occurrences up to 2020 and projected until 2023. Examining patent activities in the DERWENT database, the C07F-C12N pairing had a high convergence probability of 0.946 but no occurrences surpassed the 2020 threshold. However, patent counts increased from 30 in 2020 to 132 by 2023. Similar trends are seen in pairings like C07H-C40B and C07K-C08G, highlighting advancements in these sectors.

Although expected convergences did not occur in 2020, the robust uptrend trend in patents indicates their future prominence. These patterns underscore the importance of sustained monitoring, pointing towards the potential of these combinations to spearhead notable technological breakthroughs in the forthcoming years.

To illustrate the significance and potential of these technology convergences, let us delve deeper into some of the notable IPC pairs highlighted by the STERGM model:

C07F-C12N: involving acyclic, carbocyclic, or heterocyclic compounds combined with microorganisms or enzymes, presents a promising frontier for novel therapeutic agents. Key developments include: 2-Arylchromone Derivatives: Leveraged from microbial fermentation, these have shown potent antitumor effects against triple-negative breast cancer cells, with further research on their action mechanisms and potential combination therapies underway (Li et al., 2022). Cyclopropane Fatty Acid Derivatives: Engineered lipase technology has yielded compounds effective against HER2-positive breast cancer cells, with current studies focusing on structural optimization and delivery methods for clinical use (Xu, et al., 2022). These advances, along with ongoing optimization and in vivo studies, underscore the potential for clinical translation, highlighting the dynamic interplay of biochemistry and biotechnology in developing cancer therapeutics.

C07J-C12N: The intersection of organic chemistry and biotechnology is reshaping breast cancer treatment by enhancing precision medicine. This integration enables targeted therapies against specific markers like HER2, boosts immunotherapies for more effective cancer cell destruction, and advances personalized cell therapy tailored to individual patient needs. The promise of this approach is evident in ongoing clinical trials, such as the phase 1 study of SAR442168 for advanced solid tumors (ClinicalTrials.gov, 2023),

Table 7
Premature predictions using the STERGM model and historical patent counts.

IPC Category Pairings	Predicted Convergence Probability	Presence in 2020 (Beyond Threshold)	Total Patents up to 2020	Total Patents up to 2023
C07F-C12N	0.946	0	30	132
C07H-C40B	0.802	0	14	146
C07K-C08G	0.776	0	13	72
C12M-C12P	0.758	0	13	74

indicating a leap forward in treatment efficacy and safety. Supported by MarketsandMarkets (2023), the C07J-C12N integration is seen as a significant innovation in combating breast cancer.

4.9. Discussion on link intensification

Traditional approaches to technology convergence often rely on binary network analysis, which focuses solely on whether ties exist between technologies. While useful for mapping connections, this approach overlooks the evolving strength and intensity of these ties. STERGM is effective in analyzing tie formation and persistence, but its binary framework fails to capture deeper technological dynamics, such as how relationships between technologies intensify over time. To address this limitation, we expanded our approach after completing our binary network analysis using STERGM to examine the concept of intensification networks.

In intensification networks, links are classified based on whether they have intensified (coded as 1) or not (coded as 0) compared to the previous period. This allowed us to evaluate whether the intensification of existing ties—beyond their binary presence—offered new insights into technology convergence dynamics. By creating an intensification network, we could assess whether technologies with existing ties continued to strengthen over time.

Table 8 presents a year-wise breakdown comparing the binary and intensification networks. The table tracks key processes such as formation, persistence, dissolution, intensified edges, and non-intensified edges. These metrics illustrate the differences between a binary network's ability to capture whether ties exist and an intensification network's ability to capture how strong or weak those ties become over time. For example, in the binary network, the formation process shows new technological connections emerging, while the persistence and dissolution metrics indicate whether those ties remain or disappear. However, in the intensification network, we shift focus to whether those ties become stronger (intensified edges) or remain constant (non-intensified edges). As observed, the number of intensified edges steadily grows over the years, reflecting the increasing strength of technological ties in the breast cancer patent networks studied. From 73 intensified edges in 2012 to 145 in 2020, we see a consistent pattern of deepening technological relationships.

In this study, the results, summarized in Table 9, highlight several key insights into the dynamics of technology convergence—insights that would have been overlooked had we focused solely on the binary network.

The comparison between the two network types reveals crucial distinctions, especially regarding centrality, homophily, and node size effects. In particular, the joint impact of **NODECOV (Eigenvector Centrality)** and **ABSDIFF (Eigenvector Centrality)** offers contrasting insights. In the intensification network, the influence of centrality on tie formation is significantly stronger (3.999) compared to the binary network (3.328), showing that centrality plays a more prominent role when accounting for the strength of connections. Meanwhile, the persistence phase sees a negative effect for **ABSDIFF** in both networks, indicating that larger disparities in node importance tend to weaken the likelihood of sustained ties, especially in the intensification context. Next, the homophily effects captured by **NODEMATCH (A-A, B-B, C-C)** are more pronounced in the intensification network, particularly during the formation process. For instance, **NODEMATCH (C-C)** in the intensification network has a much higher coefficient (0.811) than in the binary network (0.347). This suggests that technologies within the same category are more likely to deepen their ties rather than simply establish new ones, highlighting the greater relevance of homophily in the intensification framework.

In conclusion, while both the binary and intensification networks reveal similar overarching trends of technology convergence, the intensification network offers deeper insights into the dynamics of centrality and homophily.

5. Conclusions and limitations

5.1. Conclusions

This study analyzes technology convergence using STERGM, focusing on the persistence process. To answer the research questions:

RQ1: What factors influence the dynamics of technology convergence in both formation and persistence? Are the mechanisms for formation and persistence the same or different?

In summary, the dynamics of technology convergence are influenced by factors such as node size, homophily, preferential

Table 8
Comparison between binary and intensification network metrics.

Year	Binary network			Intensification Network	
	Formation	Persistence	Dissolution	Intensified Edges	Non-Intensified Edges
2011	385	0	0	0	385
2012	105	258	127	73	290
2013	126	253	110	70	309
2014	112	241	138	106	247
2015	94	250	103	49	295
2016	113	257	87	132	238
2017	119	277	93	128	268
2018	170	291	105	125	336
2019	86	287	174	112	261
2020	156	293	80	145	204

Table 9
Comparative STERGMs for binary and intensification networks.

Model terms	Binary network		Intensification Network	
	Form	Persist	Form	Persist
EDGES	**-8.855***** (0.164)	**-2.852***** (0.257)	**-9.664***** (0.303)	**-3.510***** (0.487)
GWESP	**0.829***** (0.027)	**0.651***** (0.038)	**0.823***** (0.048)	**0.728***** (0.066)
NODECOV (Eigenvector Centrality)	**3.328***** (0.246)	**1.174**** (0.379)	**3.999***** (0.471)	1.259 (0.726)
ABSDIFF (Eigenvector Centrality)	**0.912**** (0.285)	**-1.712***** (0.504)	**1.863***** (0.517)	**-2.010*** (0.931)
NODECOV(NSIZE)	-0.001 (0.001)	**0.002*** (0.001)	-0.001 (0.001)	0.001 (0.001)
ABSDIFF(NSIZE)	0.001 (0.001)	**-0.002*** (0.001)	0.001 (0.001)	-0.001 (0.001)
NODEMATCH (A-A)	**0.150*** (0.076)	**0.204*** (0.097)	**0.303**** (0.117)	0.108 (0.143)
NODEMATCH (B-B)	0.279 (0.144)	0.097 (0.170)	**0.476*** (0.224)	-0.054 (0.241)
NODEMATCH (C-C)	**0.347*** (0.158)	**0.661***** (0.194)	**0.811***** (0.208)	**0.820**** (0.287)
GWDSP	**0.045***** (0.002)	**0.042***** (0.007)	**0.045***** (0.005)	**0.058***** (0.013)
AIC	11085.455		4573.748	
BIC	11277.144		4749.219	

attachment, centrality, and network structural features like triadic closure. The mechanisms for tie formation and persistence differ, with certain factors like node size and centrality being more important for persistence, while homophily plays a role primarily in formation.

Our model reveals complex interdependencies between variables influencing technology convergence. For instance, **NODEMATCH** (homophily) is significant when **GWDSP** (network cohesion) is excluded but loses significance when GWDSP is included. This aligns with Barabási and Albert's (1999) scale-free network theory, where nodes with more connections preferentially attract new links, and contrasts with Breschi and Lissoni (2009), who emphasized the role of homophily in innovation networks. Additionally, some scholars have observed that homophily and the preferential attachment rule can have reverse effects on network formation, notably, homophily can amplify the effects of basic preferential attachment (Kim & Altmann, 2017), which partially aligns with our findings. Additionally, technologies with high centrality (**NODECOV** - Eigenvector Centrality) are more likely to establish persistent ties, emphasizing their importance within the network. A negative **ABSDIFF** (difference in centrality) suggests that technologies with similar levels of centrality are more likely to sustain their connections. This nuanced relationship contrasts with Park and Yoon (2014), who found that central nodes are active in both forming and maintaining ties. There are studies examining the impact of centrality on the formation of ties (Lu & Li, 2023). However, research on how centrality influences the maintenance of ties is relatively scarce. Our results suggest that centrality is more critical for persistence, highlighting the role of influential technologies in network stability.

Unlike traditional models that focus on a single static process, we separate the factors influencing tie formation from those affecting persistence. For example, the **NSIZE** variable, reflecting node size, is insignificant during the formation phase but becomes significant during the persistence phase. This suggests that while larger technologies may not initiate new connections, they are vital for maintaining existing relationships, highlighting their stable role within the innovation ecosystem. Our study findings echo the conclusions of Juhasz & Lengyel (2018), yet also present distinctions. They identified that the formation of new relationships is predominantly influenced by opportunities and network positioning, whereas the persistence of these relationships is deeply affected by the perceived value of the connection. Our results also contrast with studies like Yoon and Park (2004) and Sun and Grimes (2017), which found that larger nodes are more likely to form new ties due to their attractiveness and resource availability. Our findings indicate that node size is more crucial for tie persistence than formation, possibly because established technologies focus on sustaining relationships rather than expanding them.

RQ2: Does integrating persistence mechanisms enhance the predictive power for technology convergence?

Our findings confirm that integrating persistence mechanisms into the predictive model significantly enhances its accuracy and effectiveness. This is demonstrated by the improved **Model Performance**, where the STERGM model incorporating both formation and persistence processes achieves a higher Precision-Recall AUC (PR AUC) of 0.742, compared to 0.665 for models considering only formation. The **Prediction Accuracy** is also superior, with the combined model yielding an F1 score of 69.54% for predicting technology convergence in 2020. Additionally, the analysis of **Type I Errors** and **premature predictions** reveals that our model effectively identifies technology pairs poised for future convergence, even if they have not yet converged within the expected timeframe. These results collectively verify that accounting for both the formation of new ties and the maintenance of existing ones provides a more comprehensive and reliable prediction of technology convergence trends.

5.2. Limitations and implications

Our framework for studying technology convergence is scalable and could be applied to other fields. However, parameter estimates from our empirical study in the breast cancer treatment may not transfer directly and requiring contextual reassessment to affirm their relevance and accuracy.

In our methodology, we transformed two-mode Patent-IPC networks into one-mode IPC co-occurrence networks, thereby assuming that all pairs of IPC codes within a patent are related. This simplification can sometimes be unrealistic. Although STERGM can handle two-mode data directly, as demonstrated by Broekel and Bednarz (2018), applying it to our study would require analyzing links between IPCs and patents, which does not directly address technology convergence. While two-path structures via common patents could be considered, they offer an indirect approach to our research questions. Additionally, the separation assumption inherent in STERGM simplifies the analysis but may fail to capture dynamics where technologies exclude or replace each other, potentially oversimplifying complex interdependencies within technology networks. Future research should explore other factors influencing technology convergence and investigate dynamic models that address the complexities of valued networks and relax the separation assumption.

This study offers significant practical and theoretical contributions by distinguishing between formation and persistence mechanisms in technology convergence. This differentiation enhances the identification and monitoring of disruptive and emerging technologies during their early stages, improving early scanning processes and enabling stakeholders to better anticipate and respond to evolving technological trends. Additionally, our findings show that separating tie formation from persistence improves prediction performance. This suggests that future research should develop more sophisticated predictive models that integrate both mechanisms. For instance, incorporating these processes into graph neural network models could significantly enhance the accuracy and applicability of technology convergence predictions in real-world contexts.

CRediT authorship contribution statement

Guancan Yang: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Di Liu:** Writing – original draft, Software, Formal analysis. **Ling Chen:** Formal analysis, Conceptualization. **Kun Lu:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization.

Acknowledgments

This research received financial support from the National Natural Science Foundation of China under grants 72274205, and Education Ministry's Key Research Base for Humanities and Social Sciences Major Projects under the grants 22JJD870001.

References

Abbasi, A., Hossain, L., & Leydesdorff, L. (2012). Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics, 6*(3), 403–412.

An, W., & Ding, Y. (2018). The landscape of causal inference: Perspective from citation network analysis. *The American Statistician, 72*(3), 265–277.

Angst, M., & Hirschi, C. (2017). Network dynamics in natural resource governance: A case study of Swiss landscape management. *Policy Studies Journal, 45*(2), 315–336.

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286*(5439), 509–512.

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography, 9*(4), 439–468.

Bröring, S., Martin Cloutier, L., & Leker, J. (2006). The front end of innovation in an era of industry convergence: evidence from nutraceuticals and functional foods. *R&D Management, 36*(5), 487–498.

Broekel, T., Balland, P. A., Burger, M., & Van Oort, F. (2014). Modeling knowledge networks in economic geography: A discussion of four methods. *The Annals of Regional Science, 53*(2), 423–452.

Broekel, T., & Bednarz, M. (2018). Disentangling link formation and dissolution in spatial networks: an application of a two-mode STERGM to a project-based R&D network in the German biotechnology industry. *Networks and Spatial Economics, 18*, 677–704.

Cantner, U., & Graf, H. (2006). The network of innovators in Jena: An application of social network analysis. *Research Policy, 35*(4), 463–480.

Carnegie, N. B., Krivitsky, P. N., Hunter, D. R., & Goodreau, S. M. (2015). An approximation method for improving dynamic network model fitting. *Journal of Computational and Graphical Statistics, 24*(2), 502–519.

Cao, X., Chen, X., Huang, L., Deng, L., Cai, Y., & Ren, H. (2023). Detecting technological recombination using semantic analysis and dynamic network analysis. *Scientometrics*, 1–32.

Chakraborty, M., Byshkin, M., & Crestani, F. (2020). Patent citation network analysis: A perspective from descriptive statistics and ERGMs. *Plos one, 15*(12), Article e0241797.

Cho, Y., & Kim, M. (2014). Entropy and gravity concepts as new methodological indexes to investigate technological convergence: Patent network-based approach. *PloS one, 9*(6), e98009.

Choi, J. Y., Jeong, S., & Kim, K. (2015). A study on diffusion pattern of technology convergence: Patent analysis for Korea. *Sustainability, 7*(9), 11546–11569.

Choi, S., Afifuddin, M., & Seo, W. (2022). A supervised learning-based approach to anticipating potential technology convergence. *IEEE Access, 10*, 19284–19300.

Chopra, S., Khosla, M., & Vidya, R. (2023). Innovations and challenges in breast cancer care: A review. *Medicina, 59*(5), 957.

ClinicalTrials.gov. (2023, November 14). A phase 1 study of SAR442168 in patients with advanced solid tumors. Retrieved from https://classic.clinicaltrials.gov/ct2/show/NCT06135714.

Copeland, M., Kamis, C., & West, J. S. (2023). To make and keep friends: The role of health status in adolescent network tie formation and persistence. *Social Networks, 74*, 216–223.

Curran, C. S., Bröring, S., & Leker, J. (2010). Anticipating converging industries using publicly available data. *Technological Forecasting and Social Change, 77*(3), 385–395.

Curran, C. S., & Leker, J. (2011). Patent indicators for monitoring convergence–examples from NFF and ICT. *Technological forecasting and social change, 78*(2), 256–273.

Dahlander, L., & McFarland, D. A. (2013). Ties that last: Tie formation and persistence in research collaborations over time. *Administrative Science Quarterly, 58*(1), 69–110. https://doi.org/10.1177/0001839212474272

Dokuka, S., & Valeeva, D. (2015). Statistical models for analysis of social network dynamics in educational studies. *Educational Studies Moscow*, (1), 201–213.

Feng, S., An, H., Li, H., Qi, Y., Wang, Z., Guan, Q., Li, Y., & Qi, Y. (2020). The technology convergence of electric vehicles: Exploring promising and potential technology convergence relationships and topics. *Journal of Cleaner Production, 260*, Article 120992.

Fenn, J., & Raskino, M. (2008). *Mastering the hype cycle: How to choose the right innovation at the right time*. Harvard Business Press.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science, 47*(1), 117–132.

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research Policy, 30*(7), 1019–1039.

Fritz, C., Lebacher, M., & Kauermann, G. (2020). Tempus volat, hora fugit: A survey of tie-oriented dynamic network models in discrete and continuous time. *Statistica Neerlandica, 74*(3), 275–299.

Geum, Y., Kim, C., Lee, S., & Kim, M. S. (2012). Technological convergence of IT and BT: Evidence from patent analysis. *Etri Journal, 34*(3), 439–449.

Goodreau, S. M., Kitts, J. A., & Morris, M. (2009). Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography, 46*(1), 103–125.

Hacklin, F., Marxt, C., & Fahrni, F. (2009). Coevolutionary cycles of convergence: An extrapolation from the ICT industry. *Technological Forecasting and Social Change, 76*(6), 723–736.

Han, E. J., & Sohn, S. Y. (2016). Technological convergence in standards for information and communication technologies. *Technological Forecasting and Social Change, 106*, 1–10.

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). statnet: Software tools for the representation, visualization, analysis and simulation of network data. *Journal of Statistical Software, 24*(1), 1548.

Hunter, D. R. (2007). Curved exponential family models for social networks. *Social Networks, 29*(2), 216–230.

Ibrahim, A. E. A., Elamer, A. A., & Ezat, A. N. (2021). The convergence of big data and accounting: Innovative research opportunities. *Technological Forecasting and Social Change, 173*, Article 121171.

Jeong, S., Lee, S., Kim, J., Oh, S., & Kwak, K. (2012). Organizational strategy for technology convergence. *International Journal of Economics and Management Engineering, 6*(8), 1989–1995.

Jones, J. H., & Handcock, M. S. (2003). An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*(1520), 1123–1128.

Juhász, S., Broekel, T., & Boschma, R. (2020). Explaining dynamics of relatedness: The role of co-location, complexity, and collaboration. *Papers in Regional Science, 100*(1), 1–19.

Juhász, S., & Lengyel, B. (2018). Creation and persistence of ties in cluster knowledge networks. *Journal of Economic Geography, 18*(6), 1203–1226.

Karvonen, M., Kassi, T., & Kapoor, R. (2010). Technological innovation strategies in converging industries. *International Journal of Business Innovation and Research, 4*(5), 391–410.

Kim, J., & Lee, S. (2017). Forecasting and identifying multi-technology convergence based on patent data: The case of IT and BT industries in 2020. *Scientometrics, 111*(1), 47–65.

Kim, K., & Altmann, J. (2017). Effect of homophily on network formation. *Communications in Nonlinear Science and Numerical Simulation, 44*, 482–494.

Kim, N., Lee, H., Kim, W., Lee, H., & Suh, J. H. (2015). Dynamic patterns of industry convergence: Evidence from a large amount of unstructured data. *Research Policy, 44*(9), 1734–1748.

Kim, T. S., & Sohn, S. Y. (2020). Machine-learning-based deep semantic analysis approach for forecasting new technology convergence. *Technological Forecasting and Social Change, 157*, Article 120095.

Kogler, D. F., Rigby, D. L., & Tucker, I. (2013). Mapping Knowledge Space and Technological Relatedness in US Cities. *European Planning Studies, 21*(9), 1374–1391.

Koka, B. R., Madhavan, R., & Prescott, J. E. (2006). The evolution of interfirm networks: Environmental effects on patterns of network change. *Academy of Management Review, 31*(3), 721–737.

Krivitsky, P.N., & Goodreau, S.M. (2019, June). STERGM-Separable Temporal ERGMs for modeling discrete relational dynamics with statnet. Technical Report. http://cran.nexr.com/web/packages/tergm/vignettes/STERGM.pdf.

Krivitsky, P. N., & Handcock, M. S. (2014). A separable model for dynamic networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 76*(1), 29–46.

Lebacher, M., Thurner, P. W., & Kauermann, G. (2021). A dynamic separable network model with actor heterogeneity: An application to global weapons transfers. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 184*(1), 201–226.

Lee, W. S., Han, E. J., & Sohn, S. Y. (2015). Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents. *Technological Forecasting and Social Change, 100*, 317–329.

Lee, C., Hong, S., & Kim, J. (2021). Anticipating multi-technology convergence: A machine learning approach using patent information. *Scientometrics, 126*, 1867–1896.

Leifeld, P., & Cranmer, S. J. (2019). A theoretical and empirical comparison of the temporal exponential random graph model and the stochastic actor-oriented model. *Network Science, 7*(1), 20–41.

Li, Y., Chen, Y., Wang, X., et al. (2022). Microbial synthesis of 2-arylchromones, potent activators of ahr signaling with antitumor activity in triple-negative breast cancer cells. *ACS Chemical Biology, 17*(10), 2441–2452.

Li, Z., Wang, Y., & Deng, Z. (2022). Research on evolution characteristics and factors of Nordic green patent citation network. *Sustainability, 14*(13), 7743.

Losacker, S. (2022). License to green': Regional patent licensing networks and green technology diffusion in China. *Technological Forecasting and Social Change, 175*, Article 121336.

Lu, C., & Li, B. (2023). The influence factors of innovation networking formation based on ERGM: Evidence from the smart medical industry. *Journal of Digital Economy, 2*, 64–80.

Lusher, D., Koskinen, J., & Robins, G. (Eds.). (2013). *Exponential random graph models for social networks: Theory, methods, and applications*. Cambridge University Press.

Ma, D., Li, Y., Zhu, K., Huang, H., & Cai, Z. (2022). Who innovates with whom and why? A comparative analysis of the global research networks supporting climate change mitigation. *Energy Research & Social Science, 88*, Article 102523.

MarketsandMarkets. (2023). Breast Cancer Drugs Market - global outlook 2027. Retrieved from https://www.marketsandmarkets.com/Market-Reports/breast-cancer-market-201.html.

Morris, M., Handcock, M. S., Butts, C. T., Hunter, D. R., Goodreau, S. M., de-Moll, S. B., & Krivitsky, P. N. (2015). *Temporal Exponential Random Graph Models (TERGMs) for dynamic network modeling in statnet*. Statnet Development Team. Retrieved from http://statnet.org/Workshops/tergm_tutorial.html.

No, H. J., & Park, Y. (2010). Trajectory patterns of technology fusion: Trend analysis and taxonomical grouping in nanobiotechnology. *Technological Forecasting and Social Change, 77*(1), 63–75.

Oh, S., Choi, J., Ko, N., & Yoon, J. (2020). Predicting product development directions for new product planning using patent classification-based link prediction. *Scientometrics, 125*, 1833–1876.

Park, H., & Yoon, J. (2014). Assessing coreness and intermediarity of technology sectors using patent co-classification analysis: The case of ICT convergence. *Technological Forecasting and Social Change, 83*, 193–206.

Peng, T. Q. (2015). Assortative mixing, preferential attachment, and triadic closure: A longitudinal study of tie-generative mechanisms in journal citation networks. *Journal of Informetrics, 9*(2), 250–262.

Powell, W. W., White, D. R., Koput, K. W., & Owen-Smith, J. (2005). Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *American Journal of Sociology, 110*(4), 1132–1205.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks, 29*(2), 173–191.
San Kim, T., & Sohn, S. Y. (2020). Machine-learning-based deep semantic analysis approach for forecasting new technology convergence. *Technological Forecasting and Social Change, 157*, Article 120095.
Sick, N., & Bröring, S. (2022). Exploring the research landscape of convergence from a TIM perspective: A review and research agenda. *Technological Forecasting and Social Change, 175*, Article 121321.
Statnet Development Team. (2021, July). Temporal Exponential Random Graph Models (TERGMs) for dynamic network modeling in statnet. https://statnet.org/Workshops/tergm/tergm_tutorial.html.
Sun, Y., & Grimes, S. (2017). The actors and relations in evolving networks: The determinants of inter-regional technology transaction in China. *Technological Forecasting and Social Change, 125*, 125–136.
Tang, Y., Lou, X., Chen, Z., & Zhang, C. (2020). A study on dynamic patterns of technology convergence with IPC co-occurrence-based analysis: The case of 3D printing. *Sustainability, 12*(7), 2655.
Ter Wal, A. L. J. (2014). The dynamics of the inventor network in German biotechnology: Geographic proximity versus triadic closure. *Journal of Economic Geography, 14*(3), 589–620.
Tsamardinos, I., & Borboudakis, G. (2010). Permutation testing improves Bayesian network learning. In J. Fürnkranz, T. Johansen, M. Numao, & D. Entschev (Eds.), *Proceedings of the Joint European conference on machine learning and knowledge discovery in databases (ECML PKDD 2010)* (pp. 322–337). Berlin, Heidelberg: Springer. Vol. 6301.
Verhoeven, D., Bakker, J., & Veugelers, R. (2016). Measuring technological novelty with patent-based indicators. *Research Policy, 45*(3), 707–723.
Wang, J., & Hsu, T. Y. (2023). Early discovery of emerging multi-technology convergence for analyzing technology opportunities from patent data: The case of smart health. *Scientometrics*, 1–30.
Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics, 113*(2), 331–360.
Xie, J., Bi, Y., Sha, Z., Wang, M., Fu, Y., Contractor, N., Lin, G., & Chen, W. (2019). Data-driven dynamic network modeling for analyzing the evolution of product competitions. *Journal of Mechanical Design, 142*(3), 1–41.
Xiao, T., Makhija, M., & Karim, S. (2022). A knowledge recombination perspective of innovation: review and new research directions. *Journal of Management, 48*(6), 1724–1777.
Xu, Y. (2021). Evolution of audience duplication networks among social networking sites: Exploring the influences of preferential attachment, audience size, and niche width. *New Media & Society, 24*(5), 709–730.
Xu, J., Zhang, Y., Liu, X., et al. (2022). Engineered lipase enables production of cyclopropane fatty acid derivatives as potential inhibitors of HER2-positive breast cancer cells. *Journal of Medicinal Chemistry, 65*(22), 15444–15456.
Yan, Y., Zhang, J., & Guan, J. (2016). The dynamics of technological partners: A social network perspective. *Technology Analysis & Strategic Management, 28*(9), 1032–1046.
Yang, G., Xing, J., Xu, S., & Zhao, Y. (2024). A framework armed with node dynamics for predicting technology convergence. *Journal of Informetrics, 18*(4), Article 101583.
Youn, H., Strumsky, D., Bettencourt, L. M. A., & Lobo, J. (2015). Invention as a combinatorial process: evidence from US patents. *Journal of The Royal Society Interface, 12*(106), Article 20150272.
Yoon, B., & Park, Y. (2004). A text-mining-based patent network: Analytical tool for high-technology trend. *The Journal of High Technology Management Research, 15*(1), 37–50.
Youtie, J., & Shapira, P. (2008). Building an innovation hub: A case study of the transformation of university roles in regional technological and economic development. *Research Policy, 37*(8), 1188–1204.
Yun, J., & Geum, Y. (2019). Analysing the dynamics of technological convergence using a co-classification approach: A case of healthcare services. *Technology analysis & strategic management, 31*(12), 1412–1429.
Zhang, C., Bu, Y., Ding, Y., & Xu, J. (2018). Understanding scientific collaboration: Homophily, transitivity, and preferential attachment. *Journal of the Association for Information Science and Technology, 69*(1), 72–86.