Spatial-Temporal Information Processing Pathways and Mechanisms Identified in AI Citation Networks

Keywords: citation networks, Artificial Intelligence (AI), topological data analysis (TDA), information cascades, alluvial diagrams

Extended Abstract

Introduction: *Self-organization*, *synergies*, and *emergence* are three distinct features of complex systems. Generally, self-organizations of a complex system first occur at multiple length scales, and result in persistent synergetic interactions, which subsequently generate emergence phenomena *per se*. This process can happen causally, or anti-causally, where the emergence phenomena trickle down to affect the initial synergies and self-organizations reciprocally, forming a closed loop. In equilibrium, detailed balances ensure the dynamics of systems become stable, and we refer to these states or structures as *stable states* or *stable regimes.* However, to elucidate these processes, one needs to resort to multi-scaling in both spatial and temporal dimensions on the three stages, respectively. To this end, we propose a spatial-temporal framework to understand how AI science organizes and evolves, where characterizing such dynamics is critical for future human interaction with AI.

Methods: TDA is a powerful tool that can be used to study spatial data, where it has been adapted to study stock markets (*1-4*). In TDA, we start with a filtration procedure, where we vary a filtration parameter ϵ to investigate self-organization at different spatial scales. Owing to emergence phenomena equilibrating with the system, they can form structures that persist in a wide range of scales, which are associated with emergent structures. In this work, we apply TDA to citation networks, to identify such persistent emergent structures. By comparing the persistent emergent structures in successive years, we can characterize their structural and topological dynamics.

For citation networks, alluvial diagram (*5-7*) is a crucial technique to elucidate the evolution of research disciplines or topics, in terms of how they merge, split, continue, grow, or diminish. In this work, we collected bibliographic records from Web of Science using keywords such as "artificial intelligence", "neural networks", which consists of more than 1.5 million records dating from 1970 to 2023. We construct alluvial diagrams from 1991 to 2005, and track all the information pathways of the top 20 topic clusters, in terms of their information cascades such as continuation, splitting, and merging in different years.

Results: In our preliminary study, we applied TDA to obtain bibliographic coupling networks (BCNs) at different weights (common references) w and found that at $w = 1$, we normally have core-shell-like structures, only in some years they fragmented into many pieces. By checking the year when this happens, we identify the AI breakthrough years as 1993 and 1996. In Figure 1(A)-(B), we found an AI + medicine cluster in 1993, which originated from two clusters in 1992, split into two clusters in 1994. By analyzing the bibliographical contents in each cluster, we identified that the topic of AI + medicine cluster is using expert systems to perform medical applications such as diagnostic reasoning, medical knowledge acquisition, etc (see Figure 1(C)). In 1994, the field split into two related but separate topics: one is radiology and imaging, and the other is on computer science topics such as problem-solving, decision planning, etc. In 1996, four clusters merged into a single large cluster in 1996, before it split into four clusters in 1997. We found that the cluster in 1996 contains a mixture of several different topics; one cluster in 1997 focuses on using AI for medical applications, such as

computer-aided diagnosis and identification, whereas the rest three are more into research-oriented topics such as binary trees, graphs, learning theory, learning algorithms, etc.
Outlook: We plan to obtain BCNs, alluvial diagram, and topics of clusters from year 2006 to 2023. Our second aim is to identify shorter cascades during normal periods, one longer cascade associated with breakthrough years, and to perform a detailed analysis to inform the research topics involved in these cascades. Our work does not involve any ethical considerations.

References

1. P. T.-W. Yen, S. A. Cheong, Using topological data analysis (TDA) and persistent homology to analyze the stock markets in Singapore and Taiwan. *Frontiers in Physics* **9**, 572216 (2021).
2. P. T.-W. Yen, S. A. Cheong, Scale-Dependent Inverse Temperature Features Associated With Crashes in the US and Japanese Stock Markets. *Complexity* **2025**, 9451788 (2025).
3. P. T.-W. Yen, K. Xia, S. A. Cheong, Understanding changes in the topology and geometry of financial market correlations during a market crash. *Entropy* **23**, 1211 (2021).
4. P. T.-W. Yen, K. Xia, S. A. Cheong, Laplacian spectra of persistent structures in Taiwan, Singapore, and US stock markets. *Entropy* **25**, 846 (2023).
5. D. Chavalarias, J.-P. Cointet, Phylomemetic patterns in science evolution—the rise and fall of scientific fields. *PloS one* **8**, e54847 (2013).
6. M. Herrera, D. C. Roberts, N. Gulbahce, Mapping the evolution of scientific fields. *PloS one* **5**, e10355 (2010).
7. M. Rosvall, C. T. Bergstrom, Mapping change in large networks. *PloS one* **5**, e8694 (2010).

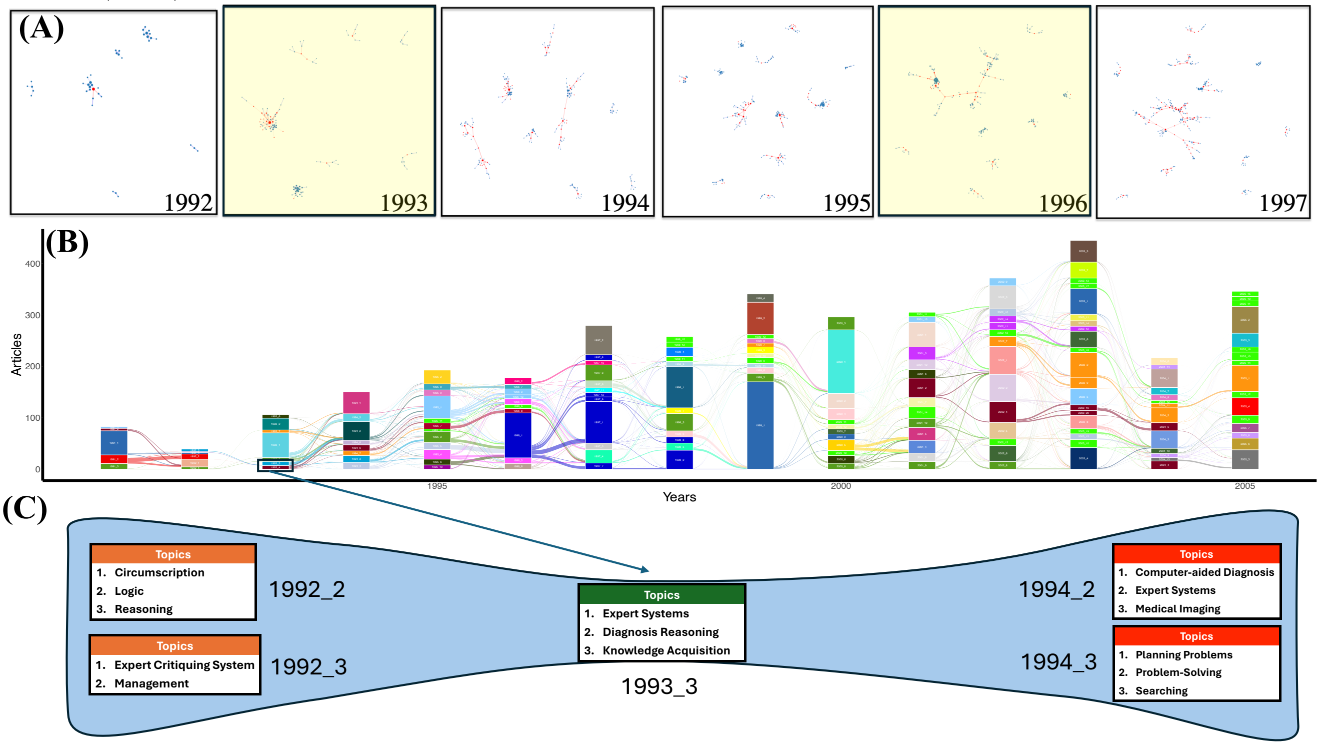

Figure 1. **Bibliographic citation networks (BCNs), alluvial diagrams, and topic of clusters.** (A) BCNs of AI papers from 1992 to 1997. BCNs with white backgrounds correpsond to normal years; BCNs with yellow backgrounds are with breakthroughs. (B) Alluvial diagram showing how AI science evolved throughout the years. (C) Topic of clusters before and after breakthrough years 1993 of AI science.