

Triad transition probabilities characterize complex networks

Katarzyna Musial, Krzysztof Juszczyzyn, and Marcin Budka

A data-driven approach to analysing and characterizing complex networks outperforms simulations based on traditional network models.

Complex networked systems are present in every aspect of our lives. Life itself is made possible by the intricate biological interactions within gene regulatory networks and food webs. Technological networks such as the Internet have changed the way the world seeks information and does business. And technology-enabled social networks have drastically altered how people meet and communicate.

In general, complex networks feature large numbers of highly interconnected units exhibiting time-dependent behaviour. A relatively simple interactivity between neighbouring units may often compound into emergent collective behaviour of surprising sophistication. The need to analyse complex systems and to predict changes in them is crucial: from assessing the potential effects of human activity on food webs, to defining telecommunications service offerings according to expected user behaviour. Yet the scale, complexity and dynamics of today's technology-based complex networks have proven resistant to traditional network analysis methods. Predicting structural changes in such networks remains a challenge.

Self-organization, synchronization and other emergent phenomena of networked structures have been widely studied for many years, mainly via simulations and theoretical analysis. Unfortunately, this work often yields unsatisfactory real-world results. Most models addressing the growth of complex networks strive to reproduce certain global characteristics of those networks: e.g. node degree distribution,¹ clustering coefficient² and network diameter.³ At the same time, models developed specifically for social networks naturally focus on features typical of those networks.⁴⁻⁸ But both approaches tend to ignore the local structure of complex systems.

Inspired by the modern-day 'data explosion',⁹ we propose using data-driven techniques to infer dynamic structure from local characteristics. We start from the observation that the local topologies of social networks differ significantly from those of standard network models (see Figure 1). Indeed, when local topology is expressed in terms of network motifs (e.g. subgraphs

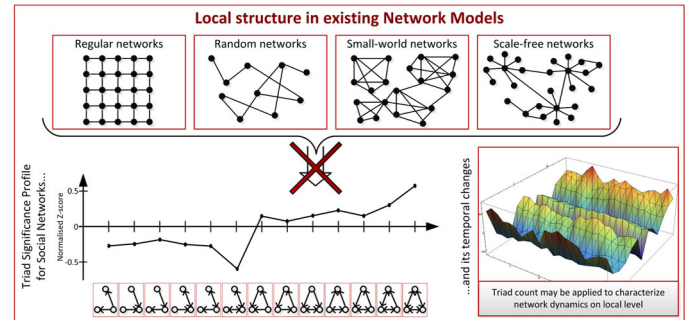


Figure 1. The triad as a tool for analysing network structure. Within complex social networks, standard network models are inadequate to describe even static local topologies, let alone their evolution over time.

containing three to seven nodes), self-evolving networks show visibly biased frequency distributions that contrast sharply with those of artificially generated networks. We then propose an approach toward quantifying network evolution schemes which, though rooted in graph analysis, uses inherent network dynamics, as revealed through observations across recorded history.

The smallest non-trivial subgraph of a network is a triad: a directed network of exactly three nodes. One technique for quantifying the connections within triads, called a triad census, involves slicing the history of the network into time windows, determining the configurations of multiple triads in each period and tallying the results. Since the three nodes in a triad form three pairs, and since each pair is able to connect in either, neither or both directions, there are 64 ($= 4^3$) possible distinct triad configurations. (If no distinction is made among the three nodes, this number decreases to 16.) We propose an enhanced procedure called the Triad Transition Method,¹⁰⁻¹² in which the configurations of a large number of triads in the network are accumulated across successive time windows. Given any two possible triad configurations, the triad census is used to estimate the probability that a triad in the first configuration will transition to the second in the next time window. The resulting Triad Transition Matrix stores these probabilities for every possible before-and-after configuration pair.

Continued on next page

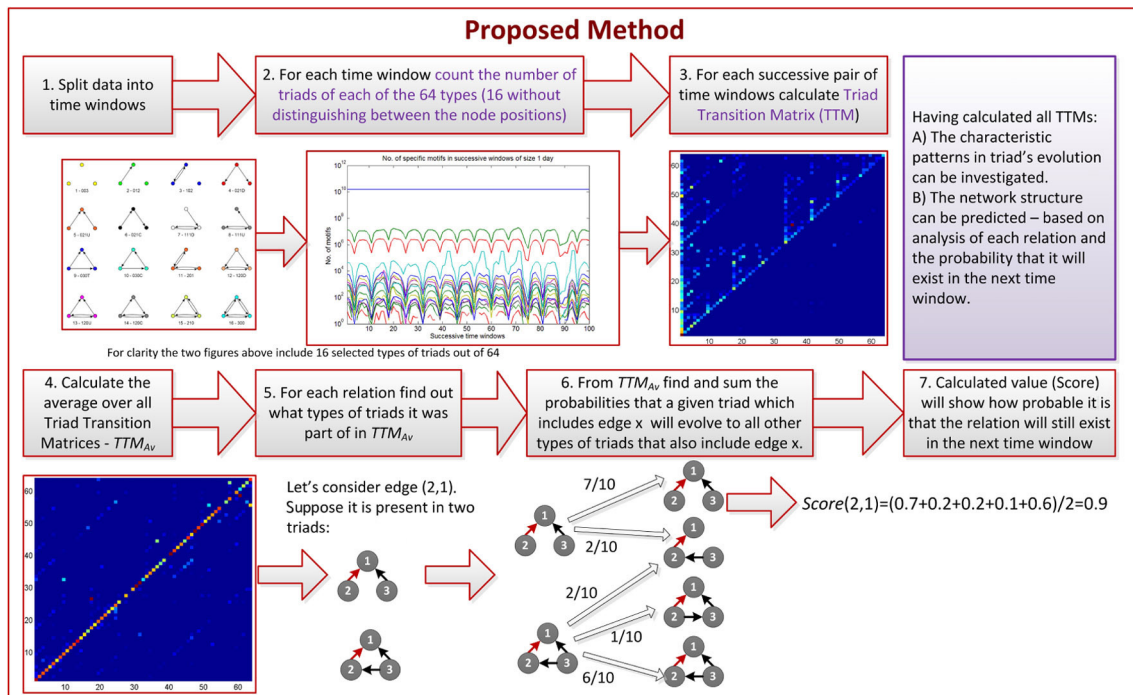


Figure 2. Flowchart and example of a predictive procedure that relies on Triad Transition Matrices.

Figure 2 presents the Triad Transition Method, broken down into individual steps. Although its statistical analysis of topological connection patterns assumes no a priori knowledge about the nature of relations among nodes, we find that its ability to predict link dynamics outperforms existing methods, especially in the case of sparse networks analysed across short timescales. Moreover, to the extent that the computed triad transition probabilities capture local connection pattern dynamics, the approach provides a useful tool for classifying distinct dynamic networks (social, biological etc.) according to the local evolutionary trajectories of a collection of nodes.

In real-world complex networks, global structure arises from local interaction patterns in a non-obvious way. The Triad Transition Method represents one of the first attempts to use historical data about local topology to predict global dynamics. As such, it may be seen as an initial step toward not only understanding and classifying these systems, but also predicting their evolution. Our future work on predictive methods will focus on considering the context of the analysed system and on non-structural network features. This will enable us to tighten the connection between machine learning, autonomy, predictive analytics and complex networked systems research.

Author Information

Katarzyna Musial
PAIS
King's College London
London, UK

Katarzyna Musial is a lecturer in computer science in the Department of Informatics. Her research is interdisciplinary, involving theoretical foundations as well as applications of complex networked systems. Her recent work has focused on the dynamics and evolution of networked systems.

Krzysztof Juszczyszyn
Wroclaw University of Technology
Wroclaw, Poland

Krzysztof Juszczyszyn is an assistant professor at the Institute of Computer Science. His research concentrates on dynamic network models applied to social networks based on communication technologies, local topology analysis and Semantic Web systems.

Marcin Budka

Bournemouth University
Bournemouth, UK

Marcin Budka is a lecturer in computational intelligence at the School of Design, Engineering and Computing. His research interests encompass machine learning, data mining, predictive modelling and computational intelligence. The evolution of complex systems is one relatively little-explored application of his expertise.

References

1. A.-L. Barabasi, **Linked: How Everything Is Connected to Everything Else and What It Means for Science, Business and Everyday Life**, Plume, 2003.
2. D. Watts, **Small Worlds: The Dynamics of Networks between Order and Randomness**, Princeton University Press, 2002.
3. B. Bollobas, **Random Graphs**, Academic, London, 1985.
4. R. Kumar, J. Novak, and A. Tomkins, *Structure and evolution of online social networks*, **12th ACM SIGKDD Int'l Conf. Knowl. Discovery Data Mining**, pp. 611–617, ACM Press, 2006.
5. J. Lescovec, L. Backstrom, R. Kumar, and A. Tomkins, *Microscopic evolution of social networks*, **14th ACM SIGKDD Int'l Conf. Knowl. Discovery Data Mining**, pp. 462–470, 2008.
6. B. Bringmann, M. Berlinger, F. Bonch, and A. Gionis, *Learning and predicting the evolution of social networks*, **IEEE Intell. Syst.** 25 (4), pp. 26–35, 2010.
7. D. Braha and Y. Bar-Yam, *From centrality to temporary fame: dynamic centrality in complex networks*, **Complexity** 12, pp. 59–63, 2006.
8. D. Liben-Nowell and J. Kleinberg, *The link-prediction problem for social networks*, **J. Am. Soc. Inf. Sci. Technol.** 58 (7), pp. 1019–1031, 2007.
9. T. Hey, S. Tansley, and K. Tolle, **The Fourth Paradigm: Data-Intensive Scientific Discovery**, Microsoft Press, Munich, 2009.
10. K. Juszczyszyn, M. Budka, and K. Musial, *The dynamic structural patterns of social networks based on triad transitions*, **2011 Int'l Conf. Adv. Soc. Netw. Anal. Mining (ASONAM)**, pp. 581–586, 2011.
11. K. Juszczyszyn, K. Musial, and M. Budka, *Link prediction based on subgraph evolution in dynamic social networks*, **3rd IEEE Int'l Conf. Social Comp. (SocialCom 2011)**, pp. 27–34, 2011.
12. K. Juszczyszyn, K. Musial, and M. Budka, *On analysis of complex network dynamics—changes in local topology*, **5th Wkshp. Soc. Netw. Mining Anal., co-located w/17th ACM SIGKDD Int'l Conf. Knowl. Discovery Data Mining (SNA-KDD)**, pp. 57–66, 2011.