Internship 2024

Subject: **Causality-aware temporal graphs with connectivity constraints**

Advisors: Binh-Minh Bui-Xuan (buixuan@lip6.fr) and Marcelo Dias de Amorim (amorim@lip6.fr)

Hosting laboratory: LIP6/CNRS - Sorbonne Université

Context

Temporal graphs, also known as dynamic or time-varying graphs [1], have gained significant interest due to their ability to model the evolving nature of complex systems. These graphs capture the relationships between entities that change over time, making them invaluable for analyzing phenomena such as social networks, transportation systems, and biological processes. By incorporating temporal information, researchers can gain deeper insights into the dynamics of these systems, identify patterns and trends, and make more accurate predictions. This has led to a wide range of applications in fields like machine learning, data mining, and network analysis [2,3].

One of the inherent characteristics of temporal graphs is the *causality effect* [4], which we explain using the example depicted in Figure 1. In the figure, we show a simple temporal graph with two undirected edges $\{a, b\}$ and $\{b, c\}$ occurring at times t_1 and t_2, respectively. As $\{a, b\}$ occurs before $\{b, c\}$, a can reach c using b as an intermediate step (a time-respecting path). On the other hand, c cannot reach a.

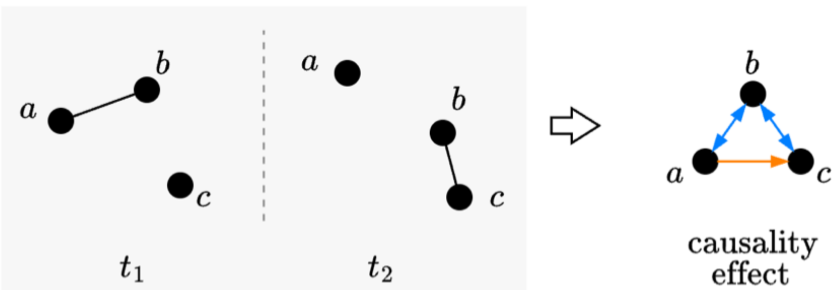

Figure 1. Illustration of a causality-aware temporal graph with three nodes. Since there is an undirected link between $\{a, b\}$ and $\{b, c\}$ at instants t_1 and t_2, respectively, the link between $\{a, c\}$ is directed, $a \rightarrow c$, as a can reach c via b, but not the other way around.

Investigating the dynamics of a causality-aware temporal graph is critical to understanding the nature of the network and the interactions among the participants. It also allows identifying hidden patterns and helps designers build efficient application-level solutions.

Problem statement

In this work, we are interested in mobile communication networks where devices can communicate directly and wirelessly with each other without having to rely on a fixed infrastructure (such as cellular networks, Wi-Fi, and satellites, to cite a few). In such a context, it is paramount to consider the inherent characteristics of wireless channels so the graph representation is representative of reality. Although temporal graphs find massive interest in practice, traditional modeling techniques, while striving to be general, overlook fundamental features in some specific wireless communication scenarios.

For instance, according to the very theoretic unit disk (/interval) graph model [5], there always is an undirected edge $\{a, b\}$ whenever a and b are within communication range (w.r.t. a distance as the crow flies, for example). Here, as long as $d(a, b) \leq \theta$ for some fixed $\theta \in \mathbb{R}$, there will be a link $\{a, b\}$ operating at maximum bandwidth, no matter if a and b are next to each other or on the border of their communication range θ. This unfortunately does not relate well real life communications such as wireless [6].

A more realistic model appears under the terminology of generalized random geometric graphs [7], using a probabilistic connection function: an undirected edge $\{a, b\}$ now exists with Poisson-like probability $P_{a,b} \approx e^{-\left(\frac{d(a,b)}{\theta}\right)^d}$, where θ still represents the communication range, and the new decay factor d models how the wireless signal weakens with distance. While very useful in capturing communication bandwidth in wireless networks, such a model lacks a temporal dimension for dynamic situations and practical tractability.

Moved by the need to capture network data collected from real-life situations, both wireless and dynamic, the problem we will consider in this internship is to mix theoretical and practical aspects of mobile networks and come up with a rich graph representation and analysis that go beyond the traditional generic representation of temporal graphs.

To begin with, we propose a blended simple framework unifying both temporal and geometric aspects in dynamic networks. A node not only has a position a (resp. b), but it now has a velocity $v_a = \frac{da}{dt}$, which can be discretized if need be. Rather than using stochastic existence of edges, we use edge-capacity at time t to model the decay factor — an undirected edge $\{a, b\}$ always exists between a and b. However, the capacity of transmission from a to b at time t is defined as $C_{a,b} \approx e^{-\left(\frac{d(a,b)}{\theta}\right)^d}$ with the same parameters as above.

Note from the above that the capacity of transmission is defined as a function depending of the position of a and b, of the form $C_{a,b} \approx f(a, b)$. In particular, we intend to use empirical values from previous experiments. As an extension, the capacity of transmission can be defined as a function of their derivatives instead, that is $C_{a,b} \approx f(v_a, v_b) = f(a^{'}, b^{'})$. This helps paving the way towards unifying computer science with control theory, by injecting temporal graphs into partial differential equations of the type $ab = f(a^{'}, b^{'})$ [8].

Internship goals

Although we expect the intern to come up with her or his own ideas, we have set some steps that we believe are sufficiently challenging and promising:

1. Problem formalization.
2. Study of metrics in temporal graphs.
3. Edge characterization with empirical measures of wireless links.
4. Revisit of graph metrics to include the characteristics investigated in the previous step.
5. Analysis of a real network containing thousands of nodes. Most likely, we will consider a city-wide vehicular dataset.
6. If possible, design and implementation of a visualization dashboard to illustrate a time-varying graph.
7. Scientific paper as an outcome of the internship.

Pre-requisites

1. Solid background in theoretical computer science.
2. Basic knowledge in networking.
3. Experience in data analysis.
4. Excellent writing skills.
5. English language.

References

1. Arnaud Casteigts, Paola Flocchini, Walter Quattrociocchi, and Nicola Santoro, "Time-varying graphs and dynamic networks", *Int. J. Parallel Emergent Distributed Syst.*, vol. 27, n. 5, pp 387-408, Apr. 2012.
2. Binh-Minh Bui-Xuan, Hugo Hourcade, and Cédric Miachon, "Computing small temporal modules in time logarithmic in history length", *Social Network Analysis and Mining*, vol. 12, art. n. 19, Dec. 2022.
3. Julian Dibbelt, Thomas Pajor, Ben Strasser, and Dorothea Wagner, "Connection scan algorithm", *ACM Journal of Experimental Algorithmics*, vol. 23, art. n. 1.7, pp. 1-56 Oct. 2018.
4. David Kempe, Jon Kleinberg, and Amit Kumar, "Connectivity and inference problems for temporal networks," *Journal of Computer and System Sciences*, vol. 64, n. 4, pp 820-842, Jun. 2002.
5. H.N. de Ridder et al., "Information system on graph classes and their inclusions". http://graphclasses.org.
6. Tomas Lagos Jenschke, Marcelo Dias de Amorim, and Serge Fdida, "Nearby Connections strategies: Features, usage, and empirical performance evaluation", *Internet of Things*, vol. 23, Oct. 2023.
7. Mathew D. Penrose, "Random geometric graphs", Oxford Univ. Press, 2003.

8. Jacques-Louis Lions, ″Optimal control of systems governed by partial differential equations“, Springer-Verlag, 1971.