How the ego perspective shapes the temporal motifs in human face to face interactions

Antonio Longa[1,2], Giulia Cencetti[1], Andrea Passerini[2], and Bruno Lepri[1]

[1]MobS Lab, Fondazione Bruno Kessler,Trento, Italy
[2]SML Lab, University of Trento, Italy

Temporal graphs are indispensable in modelling social interactions because a standard graph is not able to capture the related temporal dynamics. A decade ago, Milo *et al.* [1] had introduced the notion of *network motifs*, that are recurrent and statistically significant sub-graphs of a network. In the beginning, *motifs* were mainly used in biology and later researchers started to use this concept in other disciplines such as applications to temporal graphs. There are mainly two methods to tackle the problem of identifying the significance of such patterns in temporal graphs. An approach consists in focusing on static snapshots of temporal graphs. In contrast, other researchers derive static *motifs* from the aggregated graph and then, they validate the statistical significance of the identified pattern. An example of the latter technique is reported in Paranjape *et al.* [2]. In their work, the authors introduce a fashion for the extraction of *temporal motifs*, that consists of retrieve motifs from the aggregated graph. Then, they fetch the temporal edge, and finally, they count the instances occurring within δ time units.

Despite the goodness of the results, their approach assumes that such sub-graphs can only be found in directed graphs with non-overlapping times on edge's label. However, there are scenarios in which such requirements are not satisfied. An example of such circumstances is the datasets provided by SocioPatterns[1], representing face-to-face human intercommunication. The information was collected using a device that detects the interactions mentioned above every 20 seconds. In other terms, if a person P_i is talking with a person P_j, it is assumed that P_j is discussing with P_i. Such data allow us to design an undirected graph representing the social interaction among people. In particular, we define a graph $G_t = (V_t, E_t)$ for each instance of time t. The set of vertices V_t is the set of people involved in a conversation at time t. Whenever people interact with each other at instance t, an edge between them is created.

Another drawback of the method proposed in [2] is the hypothesis that a sub-graph that is not statistically relevant in the aggregated graph will not be significant in the temporal one. In this sense, we argue that their approach might not be ideal for the SocioPattern case study, because a *motif* that is statistically insignificant in the aggregated graph, may be vital in the temporal version.

In this extended abstract, we present a novel strategy to extract statistically significant sub-graphs in temporal networks by concentrating on the egocentricity of a node. We argue that by aggregating the temporal graphs, temporal-dependent information such as the length over time of the interactions, their frequency, periodicity and others are lost. For these reasons, for each node n contiguous static snapshots of graphs created as stated before we compose graphs as follows. First, we consider $G = \{G_1 \dots G_n\}$ as a set of time-consecutive graphs. Then, for each graph in G beginning from G_1, we focus on a specific node a, and we construct a new graph in which we add a and its neighbourhood at time t and similarly for instant $t+1$. Finally, we link the nodes that are common among sequential time-bins. An example of a three-time-bins structure is illustrated in Figure 1. Starting from the extracted structures, we implemented a null model in order to extract patterns and sub-graphs that are statistically significant. The way in which we build such structures is what allowed us to drastically diminish the number of significant patterns allowing us to enhance the data mining and information retrieval process.

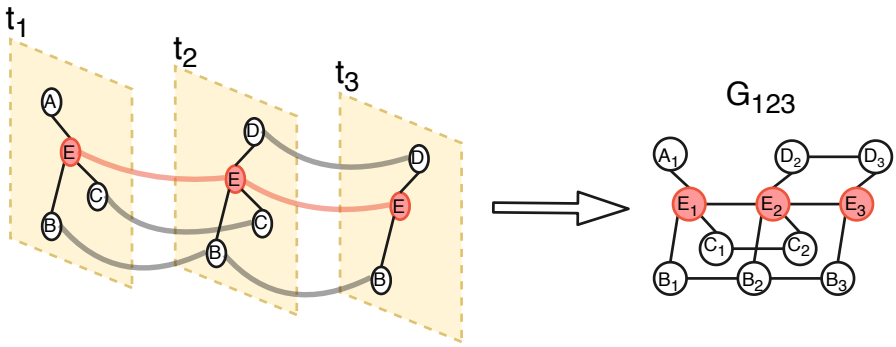

Figure 1: On the left, three snapshots of a temporal graph focusing on the ego node E. On the right, the structure created by merging three temporal layers from the point of view of the egocentric node E

To succeed in our aims, we applied an innovative technique that combines data mining and network science to identify the backbone of the complex scenario represented by social interactions. With our approach, we ultimately disentangled the bricks that constitute the essence of human interactions to build our knowledge of them.

References

[1] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[2] Ashwin Paranjape, Austin R Benson, and Jure Leskovec. Motifs in temporal networks. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 601–610, 2017.

[1]http://www.sociopatterns.org/