TEMPORAL NETWORK MOTIFS: MODELS, LIMITATIONS, EVALUATION

Penghang Liu[1], *Valerio Guarrasi*[2], *and Ahmet Erdem Sarıyüce*[1]
[1] *Department of Computer Science and Engineering*
University at Buffalo
{*penghang, erdem*}*@buffalo.edu*
[2] *Sapienza University of Rome*
guarrasi1995@mail.com

SIAM Workshop on Network Science 2020
July 9–10 · Toronto

Summary

Investigating the frequency and distribution of small subgraphs with a few nodes/edges, i.e., motifs, is an effective analysis method for stationary networks. Motif-driven analysis is also useful for temporal networks where the number of motifs is significantly larger due to the additional temporal information on edges. This variety makes it challenging to design a temporal motif model that can consider all aspects of temporality. In this work, we survey the existing temporal network motif models, discuss the advantages and limitations, and present a comparative evaluation. We argue how the different aspects of temporal networks are considered in each model. In the evaluation, we focus on the timing parameters and investigate the parameter space for temporal adjacency among events. We believe that our comparative survey and authentic evaluation will catalyze the research on temporal motifs.

Motivation

Temporality brings new challenges for network analysis. Motif-driven techniques, for instance, should consider the temporal information on edges which significantly increases the number motifs with respect to stationary networks. Order of the edges, inter-event time intervals, and durations are some of the aspects that need to be incorporated. Thus, it is beyond non-trivial to design a temporal network motif model that considers all those characteristics while being practical. There are several studies that propose temporal motif models. Those studies are introduced in various subfields of computer and network science, thus mostly unaware of each other. Consequently, there does not exist a unified approach that can address the limitations of those models while leveraging their novelty.

Temporal Motif Models

To the best of our knowledge, there are four models for temporal network motifs:

- **Kovanen et al.** [**2**] proposed the first model and introduced the notion of temporal adjacency to relate the events in a motif.
- **Song et al.** [**4**] introduced another model that aims streaming workloads where the motifs are found on-the-fly and the events can be partially ordered.
- **Hulovatyy et al.** [**1**] incorporated the induced subgraph idea to improve Kovanen et al.'s model and also discussed the events with durations.
- **Paranjape et al.** [**3**] proposed a practical model where the timing constraints are specified with respect to all the events in a given motif.

There are several aspects of temporal networks and motifs that are handled in a different way in each of those four models. Table 1 presents an overview. Here we discuss only two of those:

Motif as induced subgraph. Considering all the edges among a given set of nodes (rather than selecting a subset) has been shown to be more effective in motif-based analysis for stationary networks. Because the non-induced motifs become artificially recurrent and shadows the importance of larger induced structures. For instance, an induced square motif ($1 \to 2$, $2 \to 3$, $3 \to 4$, $1 \to 4$) implies that no diagonal edges exist (i.e., $1 \to 3$ and $2 \to 4$ do no exist) whereas a non-induced square motif has no such restriction (i.e., every 4-clique is also a square). In the temporal network context, the first model in [2] does not require a motif to be induced. However, [1] argues that the temporal motif must be induced. [3] also requires the motifs to be induced as long as the events *satisfy the timing constraints*. Song et al. [4], on the other hand,

Table 1: **Aspects of temporal motif models**

Article	**Kovanen et al.[2]**	**Song et al.[4]**	**Hulovatty et al.[1]**	**Paranjape et al.[3]**
Induced subgraph	✘	✘	✔	✔
Event durations	✘	✘	✔	✘
Partial ordering	✔	✔	✘	✘
Consecutive events	✔	✘	✘	✘
Directed edges	✔	✔	✘	✔
Node/Edge labels	✘	✔	✘	✘
Adjacent events in Δ_C	✔	✘	✔	✘
Entire motif in Δ_W	✘	✔	✘	✔

approach the issue from a different perspective and argues that non-induced temporal motifs are useful if the network is streaming.

Timing Constraints. Connectivity in the temporal dimension is a key feature for temporal motif models. Kovanen et al. proposed a model where each consecutive event pair should occur close in time and ensured this by defining an upper bound for the difference between the timestamps [2]. Formally, they define the temporal motif as a connected temporal subgraph such that for any pair of consecutive events that share a node, the time difference should be less than Δ_C. The same approach is also used by Hulovatyy et al. [1] for timing constraints. Note that both models require graph connectivity to consider consecutive events. On the other hand, [4] and [3] consider a window-based temporal connectivity where all the events in a temporal motif occur within a given time interval, denoted as Δ_W. One can consider to use both parameters to have a trade-off between the two extremes of Δ_W and Δ_C. Note that, depending on the number of events in the temporal motif, one of those two timing constraints can be useless for certain values of Δ_W and Δ_C. Given a motif with m events and $\frac{\Delta_C}{\Delta_W}$ ratio, we have the following:

$$\text{Constraints} = \begin{cases} \Delta_C & \text{if } 0 \le \dfrac{\Delta_C}{\Delta_W} \le \dfrac{1}{m-1} \\[2ex] \Delta_C, \Delta_W & \text{if } \dfrac{1}{m-1} < \dfrac{\Delta_C}{\Delta_W} < 1 \\[2ex] \Delta_W & \text{if } \dfrac{\Delta_C}{\Delta_W} \ge 1 \end{cases}$$

Experiments

We evaluate the four temporal motifs on various temporal and directed network datasets from several domains, including phone messages, emails, Facebook wall interactions, posts in Q/A websites, and call detail records (CDR). We study two sets of motifs in our evaluations; 1) Three-node, three-event motifs, 2) Square (four-node, four-event) motifs. Note that we only consider the motifs that grow as a single component, by adding one event at a time.

We observe that the temporal motif models that only consider the time window (Δ_W) amplify the repetitive motifs where interactions are in the form of repetitions, ping-pongs, and bursts. Enforcing the Δ_C constraint helps to find less motifs of this type since each consecutive event pairs should be close to each other. Comparing the only Δ_C and only Δ_W counts, we see that the number of repetitive motifs is reduced by nearly 40 percent, while the reduction in the other motif types is significantly less.

The models which only consider Δ_C also have defects. Since Δ_C does not bring any control in the motif timespan (difference between the last and first events); timespans follow a normal distribution where the mean is close to the Δ_C values. By bringing the time window constraint to the model, the distribution becomes uniform, which implies that the motifs with various timespans are equally likely to be discovered.

References

[1] Y. Hulovatyy, H. Chen, and T. Milenković. Exploring the structure and function of temporal networks with dynamic graphlets. *Bioinformatics*, 31(12):i171–i180, 2015.

[2] L. Kovanen, M. Karsai, K. Kaski, J. Kertész, and J. Saramäki. Temporal motifs in time-dependent networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(11):P11005, 2011.

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

[4] C. Song, T. Ge, C. Chen, and J. Wang. Event pattern matching over graph streams. *Proceedings of the VLDB Endowment*, 8(4):413–424, 2014.