Working Paper: Force-Directed Layout of Non-Markovian Temporal Networks

Ingo Scholtes
ETH Zürich
Chair of Systems Design
Weinbergstrasse 56/58
CH-8092 Zürich
Switzerland
ischoltes@ethz.ch

May 6, 2014

1 Motivation

Complex systems occurring in various disciplines are increasingly being represented, analyzed and modeled in terms of networks. When exploring the structure and function of such complex systems, the visualization of network topologies has become a standard technique. Many - if not most - of the systems being considered are inherently dynamic and in recent years more and more data on their dynamics are becoming available in the form of so-called *dynamic*, *time-varying networks* or *temporal networks*. Different from static representations, in such temporal networks each edge representing an interaction between two nodes has an additional time-stamp indicating when the interaction occurs. The visual exploration of such temporal networks is typically based on a *time-aggregated representation* in which the underlying sequence of time-varying interactions are aggregated into a static network. Compared to *dynamic visualizations* which show a detailed picture of a system's evolution, time-aggregate representations of temporal networks provide a number of advantages: First and foremost, it is in general easier to recognize *spatial* or *structural* patterns rather than temporal correlations. Secondly, to date there are much more options for the visualization of static networks and due to their reduced dimensionality, their handling is generally less expensive from a computational point of view. Finally, a further important aspect is the relative ease with which figures of static networks can be included in scientific publications compared to dynamic representations. Because of these arguments, the study of time-aggregated networks is common and has proven to provide insights about the characteristics of complex systems. Nevertheless, it is clear that - when aggregating interaction sequences into static networks - important information, e.g. about temporal correlations or the ordering of interactions are lost.

This article proposes a simple mechanism that can be used to include temporal correlations present in longitudinal data on temporal networks in the layout of time-aggregated network structures. We particularly focus on *non-Markovian properties* of contact sequences in temporal networks, i.e. the fact that future contacts not only depend on the current contact, but also on past contacts. Recent studies have shown that these correlations are present in empirical temporal networks and that they can significantly affect the topology of causal interactions [1, 2, 3, 4]. Being the most widely used class of algorithms, we further focus on *force-directed layout algorithms* and demonstrate how the Fruchterman-Reingold algorithm can be extended in a way such that the resulting layout represents non-Markovian properties of temporal networks. A particular benefit of our approach is that it allows to create rather simple static visualizations of inherently dynamic systems, which capture the structure and patterns present both in the topology as well as in the ordering of interactions. We demonstrate our approach by visualising a set of examples for temporal networks with non-Markovian properties. The visualisations show that the resulting visualizations reveal important aspects like e.g. community structures of the dynamics which - although present in empirical data sets - a) are not represented in static, aggregate networks and b) cannot easily be seen even in dynamic visualizations of temporal networks.

The remainder of this paper is organized as follows: In section 2 we briefly introduce force-directed layout algorithms and discuss issues arising when using them for the layout of time-aggregated representations of temporal networks. We then provide a high-level introduction to non-Markovian characteristics in temporal networks. We further introduce how they can be represented by higher-order time-aggregated networks. Based on this higher-order representations, we then introduce a force-directed layout mechanism that allows to integrate non-Markovian properties - and thus the causal topology of temporal networks - into the layout of aggregate networks. In section 3 we demonstrate our approach using both empirical and synthetic temporal networks. Finally, in section 4 we summarize our contributions and provide an overview over future research opportunities in the field of temporal network layouts.

2 Force-Directed Layouts of Temporal Networks

In the spatial representation of networks, a particularly important aspect is the choice of the layout algorithm, i.e. a function which - depending on the network topology - positions nodes in such a way that communities, symmetries and other structures become visible. A particulary popular class of layout algorithms are *force-directed* approaches, which assign attractive forces to edges in such a way that pairs of connected nodes are pulled together, while an antagonistic repulsive force exists between all pairs of nodes. The concept of force-directed layouts was introduced in [5] and optimized later both in terms of layout quality as well as computational requirements [6, 7, 8]. Further improvements of this general approach include the use of multi-level techniques that combine a force-directed placement with clustering techniques [9], the optimization of clustering [10], a proper balancing of attractive and repulsive forces [11] or which introduce tunable trade-offs between speed and layout quality [12].

Compared to the layout of static networks, so far only a comparably small number of works have addressed layout mechanisms suitable for *temporal networks* which - in addition of the topology of interactions - also consider the timing or ordering of events. To the best of our knowledge, works addressing network dynamics have exclusively been focused on the optimisation of graph layouts for subsequent snapshots of graphs in such a way to minimize the cognitive effort to relate nodes and trace the evolution of network structures [13, 14, 15].

In this paper, we address a fundamentally different question. Focusing on the layout of *time-aggregated* - and thus *static* - representations of *temporal networks*, we consider how correlations present in the underlying *time-stamped interaction sequence* can be used to enhance the layout of an aggregate network. In particular, our proposed method addresses temporal networks with non-Markovian characteristics as measured by *betweenness preference* - a recently proposed correlation present in empirical temporal networks - into standard force-based layouts. In the following, we will introduce some basic notations that allow us to introduce our approach in more detail.

We define a temporal network G as a tuple $G^T = (V, E)$ that consists of a set of nodes V and a set of *time-stamped* edges $E^T \subseteq V \times V \times \mathbf{N}$, where each edge $e \in E$ represents an interaction (v, w, t) between nodes $v, w \in V$ occurring at a discrete time step t. Based on a temporal network G^T, a weighted time-aggregated representation $G^{(1)}$ can be defined. For this, we assume that an edge $(v, w) \in E^{(1)}$ exists in the static network $G^{(1)}$ iff v and w are connected at some time stamp t, i.e. if $\exists t : (v, w, t) \in E^T$. Furthermore, a weight of an edge in the time-aggregated network can be defined based on the activity of an edge, i.e. for a discrete notion of time, the weight $w^{(1)}(v, w)$ simply counts the number of time steps during which an edge v, v was active. A simple example for a temporal network G^T consisting of seven nodes a, b, c, d, e, f, g and six consecutive time-stamped edges is given below:

$$\begin{aligned}
e_1 &= (a, c, 1) \\
e_2 &= (c, d, 2) \\
e_3 &= (b, c, 3) \\
e_4 &= (c, e, 4) \\
e_5 &= (f, c, 5) \\
e_6 &= (c, g, 6)
\end{aligned}$$

Most popular force-based layout algorithms (like e.g. [6, 7, 8]) are based on the heuristics that the ideal distance between two nodes v and w should be proportional to the length of the shortest path $p(v, w)$ between nodes v and w in G_{Agg}. However, it is important to note that in temporal networks the length of a shortest path in the time-aggregated representation is not necessarily equivalent to the length (or even existence) of a *shortest time-respecting path* in the underlying sequence of edges (compare [16]). For static networks that are time-aggregated representations of temporal networks the distance between the spatial placement of nodes thus *does not* necessarily correspond to neither the fact that a path exists, nor to its length. This can easily be seen in Figure **??**, which shows a layout of the time-aggregated representation of the example temporal network as computed by the Fruchterman-Reingold algorithm. In the static representation, any pair of the nodes a, b, d, e, f, g is connected via a path of length 2 while a time-respecting path in the temporal sequence of interactions (see example above) only exist between the nodes a and d. For real-world temporal networks, it has been shown that the ordering of interactions can lead to effective interaction topologies which are - in fact - much more complex than one would expect from a weighted aggregate representation. For this reason, standard force-based layouts are *not* a good choice for the layout of time-aggregate representations of temporal networks.

2.1 Higher-Order Representations of Non-Markovian Temporal Networks

As argued above, an important aspect when studying the topology of temporal networks is the question which time-respecting paths exist compared to what is expected from their static, time-aggregated representations. A particularly simple way to study temporal networks is by looking at so-called *two-paths* $u- > v- > w$, the simplest possible building blocks of time-respecting paths consisting of an ordered sequence of two time-stamped interactions (u, v, t) and (v, w, t') for time stamps $t < t'$. A second-order time-aggregated representation of temporal networks which captures the statistics of these two-paths has been introduced in citeScholtes2013. We briefly repeat this definition in the following. We define a *second-order aggregate network* $G^{(2)} = \left(V^{(2)}, E^{(2)}\right)$ in which each node $e \in V^{(2)}$ represents an edge in the first-order aggregate network $G^{(1)}$. As edges $E^{(2)}$, we define all possible paths of length two in $G^{(1)}$, i.e. the set of all pairs (e_1, e_2) for edges $e_1 = (a, b)$ and $e_2 = (b, c)$ in $G^{(1)}$. Based on a temporal network G^T, weights $w^{(2)}(e_1, e_2)$ in the second-order network are defined based on the frequency of two-paths $(a, b; t-1) \rightarrow (b, c; t)$ in G^T, while proportionally correcting for multiple two-paths $(a', b; t-1) \rightarrow (b, c'; t)$ passing through node b at the same time. We define

$$w^{(2)}\left(e_1, e_2\right) := \sum_t \frac{\delta_{(a,b;t-1)}\delta_{(b,c;t)}}{\sum_{a',c' \in V} \delta_{(a',b;t-1)}\delta_{(b,c';t)}}, \quad (1)$$

where $\delta_{(a,b;t)} = 1$ if edge $(a, b; t)$ exists in the temporal network G^T and $\delta_{(a,b;t)} = 0$ otherwise. Note that this definition considers two-paths consisting of directly consecutive edge activations. It is simple to generalize weights to capture two-paths $(a, b; t') - (b, c, t)$ for $1 \leq t - t' \leq \epsilon$.

Figure 1 shows an example of two temporal networks (middle) that give rise to the same first-order aggregate network $G^{(1)}$ (left). Two second-order aggregate networks $G^{(2)}$ and $\hat{G}^{(2)}$ corresponding to the two temporal networks are shown on the right. In the following, we propose a simple extension to force-based layout algorithms, which takes into account attractive forces both based on the first and the second-order aggregate network. For the sake of simplicity, we focus on the algorithm proposed in [7]. In particular, we propose to add a cumulative attractive force between pairs of nodes acting according to time-respecting paths of length two as represented in the second-order aggregate network. As a simple example, consider the pair d and c in Figure 1. In the second-order network $\hat{G}^{(2)}$ an attractive force exists between d and c, which is mediated via the link between the second-order nodes $d-b$ and $b-c$. In the second-order network $G^{(2)}$ this force is absent, since no

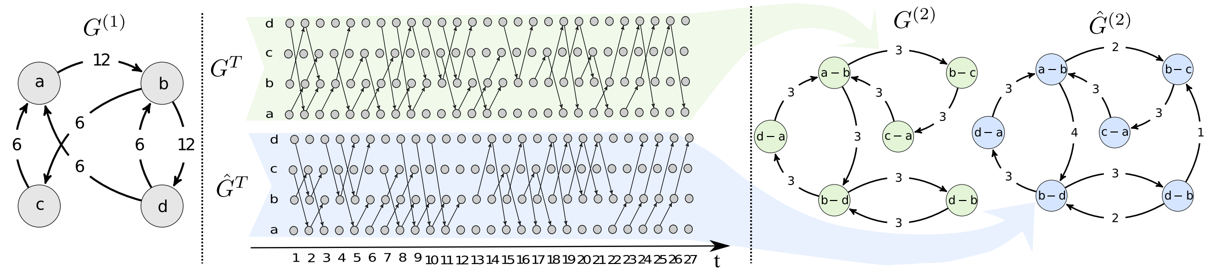


Figure 1: Two example temporal networks (middle) giving rise to the same aggregate network (middle). The two second-order aggregate representations corresponding to the temporal networks are shown on the right.

time-respecting path of length two exists that interconnects d and c via b (or any other node). The full adjusted Fruchterman-Reingold algorithm is shown in Algorithm 1 and 2. In addition, we propose a scaling parameter γ, which allows to balance the attractive forces resulting from the first- and the second-order network and which thus allows to balance the influence of either the topological or the temporal dimension.

3 Illustration

A simple demonstration of the layout for a synthetic temporal network is shown in Figure 2. The temporal network consists of a correlated interaction sequence which has been created in such a way that time-respecting paths of length two are over-represented (compared to what is expected from the first-order aggregate network) within three communities, each consisting of five nodes. The aggregate topology of this temporal network is the fully connected graph. Figure 1 (a) shows a standard Fruchterman-Reingold layout which ignores the temporal dimension of the underlying dynamic network. Figure (b) shows our proposed modified force-based layout, which combines forces from the first- and the second-order network. The three (temporal) communities are clearly reflected in the positioning of nodes. For Figures 2 (c) and (d) we have reshuffled the timestamps of the original interaction sequence and repeated the layout. Since the temporal correlations are destroyed by the reshuffling, the modified layout shows no differences from a standard Fruchterman-Reingold layout.

Algorithm 1: Skeleton of Fruchterman-Reingold Algorithm

`/* parameters: area multiplicator A, width W, height H, iterations I */`

$a = W \cdot H$;

$maxDist = \frac{\sqrt{A \cdot a}}{10}$;

$k = \frac{\sqrt{A \cdot a}}{1+n}$;

for $v \in V$ **do**

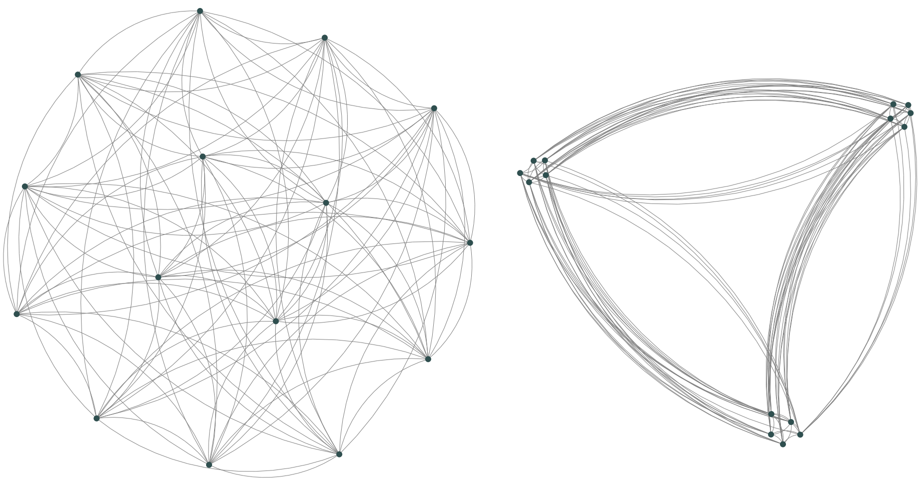
- `/* Assign random initial position */`
- $pos[v] = RandomPosition(W, H)$;

for $i \in \{1, \ldots, I\}$ **do**

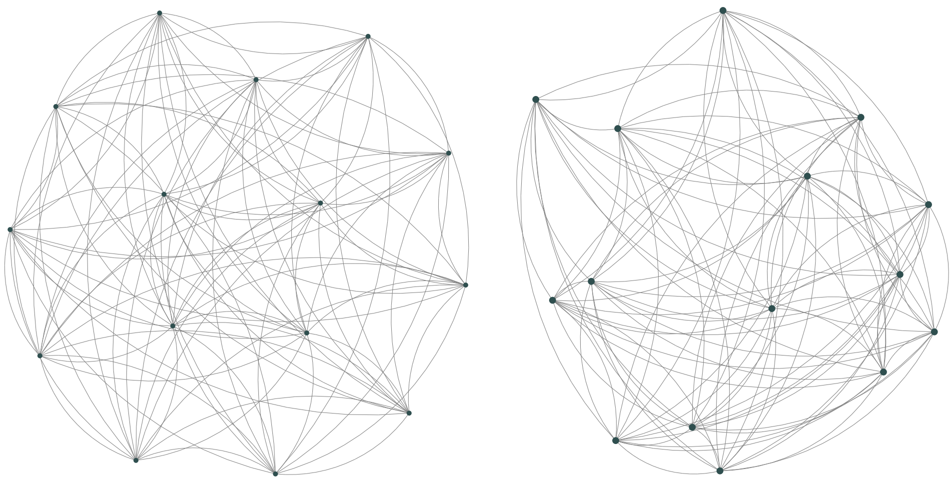
- `/* Initialize per-node displacement vectors */`
- **for** $v \in V$ **do**
  - $disp[v] = (0, 0)$;
- `/* Apply force-based node displacements */`
- **for** $v \in V$ **do**
  - **for** $w \in V : v \neq w$ **do**
    - `/* Difference vector Δ */`
    - $\Delta = pos[v] - pos[w]$;
    - `/* Repulsive force */`
    - $disp[v] = disp[v] + \frac{\Delta}{len(\Delta)} \cdot \frac{k^2}{len(\Delta)}$;
    - `/* Attractive force */`
    - $disp[v] = disp[v] - \frac{\Delta}{len(\Delta)} \cdot attractiveForce(v, w)$;
    - $disp[w] = disp[w] + \frac{\Delta}{len(\Delta)} \cdot attractiveForce(v, w)$;
- `/* Apply node displacements */`
- **for** $v \in V$ **do**
  - $pos[v] = pos[v] + \frac{disp[v]}{len(disp[v])} \cdot min(len(disp[v]), maxDist)$;
  - $disp[v] = (0, 0)$;

Algorithm 2: Computation of forces based on first- and second order aggregate network

```
function attractiveForce(v,w);
Δ = pos[v] − pos[w];
/* attractive force based on first-order aggregate network     */
f_1 = w^1(v, w);
/* attractive force based on second-order aggregate network    */
f_2 = 0;
for x ∈ V do
    n_1 = (v, x);
    n_1 = (x, w);
    f_2 = f_2 + w^2(n_1, n_2);
return force1 + γforce2;
```



(a) Fruchterman-Reingold Layout (b) Layout using Betweenness Preference



(c) Shuffled edge sequence, Fruchterman-Reingold (d) Shuffled edge sequence, Layout using Betweenness Preference

Figure 2: A fully connected aggregate network built from a synthetic, highly correlated temporal network

4 Conclusion

We have presented a simple extension of force-based layouts to non-Markovian temporal networks. Our approach is based on the idea of combining attractive forces from the time-aggregated network, with forces that are computed based on a higher-order time-aggregated representation which captures the ordering of interactions in temporal networks. While we have focused on an extension that combines forces from the *first-* and the *second-order time-aggregated networks*, our approach can be generalised to higher orders as well. The resulting layouts capture structures and patterns present both in the network topology, as well as in the temporal ordering of interactions. The approach introduced in this paper thus provides interesting perspectives for the static visualisation of time-varying complex systems.

References

[1] René Pfitzner, Ingo Scholtes, Antonios Garas, Claudio J. Tessone, and Frank Schweitzer. Betweenness preference: Quantifying correlations in the topological dynamics of temporal networks. *Phys. Rev. Lett.*, 110:198701, May 2013.

[2] M. Rosvall, A. V. Esquivel, A. Lancichinetti, J. D. West, and R. Lambiotte. Networks with Memory. *ArXiv e-prints*, May 2013.

[3] Ingo Scholtes, Nicolas Wider, René Pfitzner, Antonios Garas, Claudio J. Tessone, and Frank Schweitzer. Slow-down vs. speed-up of diffusion in non-markovian temporal networks. *arXiv preprint*, 2013.

[4] R. Lambiotte, V. Salnikov, and M. Rosvall. Effect of Memory on the Dynamics of Random Walks on Networks. *ArXiv e-prints*, January 2014.

[5] P. A. Eades. A heuristic for graph drawing. In *Congressus Numerantium*, volume 42, pages 149–160, 1984.

[6] T. Kamada and S. Kawai. An algorithm for drawing general undirected graphs. *Inf. Process. Lett.*, 31(1):7–15, April 1989.

[7] Thomas M. J. Fruchterman and Edward M. Reingold. Graph drawing by force-directed placement. *Softw. Pract. Exper.*, 21(11):1129–1164, November 1991.

[8] Yifan Hu. Efficient and high quality force-directed graph drawing. *Mathematica Journal*, 10(1), 2006.

[9] Chris Walshaw. A multilevel algorithm for force-directed graph drawing. In *Proceedings of the 8th International Symposium on Graph Drawing*, GD '00, pages 171–182, London, UK, UK, 2001. Springer-Verlag.

[10] Andreas Noack. Energy models for graph clustering. *J. Graph Algorithms Appl.*, 11(2):453–480, 2007.

[11] Markus M. Geipel. Self-organization applied to dynamic network layout. *International Journal of Modern Physics C*, 18(10):1537– 1549, April 2007.

[12] M. Jacomy, S. Heymann, and T. Venturini. Forceatlas2: A continuous graph layout algorithm for handy network visualization. available online at `http://www.medialab.sciences-po.fr/publications/Jacomy_Heymann_Venturini-Force_Atlas2.pdf`, 2012.

[13] James Moody, Daniel McFarland, and Skye Bender-deMoll. Dynamic network visualization. *American Journal of Sociology*, 110(4):1206–1241, 2005.

[14] G. Kumar and M. Garland. Visual exploration of complex time-varying graphs. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):805 –812, September/October 2006.

[15] Ulrik Brandes, Natalie Indlekofer, and Martin Mader. Visualization methods for longitudinal social networks and stochastic actor-oriented modeling. *Social Networks*, 34(3):291 – 308, 2012.

[16] Petter Holme and Jari SaramÃďki. Temporal networks. *Physics Reports*, 519(3):97 – 125, 2012.