Clustering and Blockmodeling Temporal Networks – Two Indirect Approaches

Vladimir Batagelj

Abstract Two approaches to clustering and blockmodeling of temporal networks are presented: the first is based on an adaptation of the clustering of symbolic data described by modal values and the second is based on clustering with relational constraints. Different options for describing a temporal block model are discussed.

Keywords: social networks, network analysis, blockmodeling, symbolic data analysis, clustering with relational constraints

1 Temporal Networks

Temporal networks described by *temporal quantities* (TQs) were introduced in the paper [2]. We get a *temporal network* $\mathcal{N}_\mathcal{T} = (\mathcal{V}, \mathcal{L}, \mathcal{T}, \mathcal{P}, \mathcal{W})$ by attaching the *time* $\mathcal{T}$ to an ordinary network, where $\mathcal{V}$ is the set of nodes, $\mathcal{L}$ is the set of links, $\mathcal{P}$ is the set of node properties, $\mathcal{W}$ is the set of link weights, and $\mathcal{T} = [T_{min}, T_{max})$ is a linearly ordered set of time points $t \in \mathcal{T}$ which are usually integers or reals.

In a temporal network nodes/links activity/presence, nodes properties, and links weights can change through time. These changes are described with TQs. A TQ is described by a sequence $a = [(s_r, f_r, v_r) : r = 1, 2, \ldots, k]$ where $[s_r, f_r)$ determines a time interval and v_r is the value of the TQ a on this interval. The set $T_a = \bigcup_r [s_r, f_r)$ is called the *activity set* of a. For $t \notin T_a$ its value is *undefined*, $a(t) = ⌘$.

Assuming that for every $x \in \mathbb{R} \cup \{⌘\} : x + ⌘ = ⌘ + x = x$ and $x \cdot ⌘ = ⌘ \cdot x = ⌘$ we can extend the addition and multiplication to TQs

Vladimir Batagelj (✉)
IMFM, Jadranska 19, 1000 Ljubljana, Slovenia & IAM UP, Muzejski trg 2, 6000 Koper, Slovenia & HSE, 11 Pokrovsky Bulvar, 101000 Moscow, Russian Federation,
e-mail: `vladimir.batagelj@fmf.uni-lj.si`

P. Brito et al. (eds.), *Classification and Data Science in the Digital Age*,
Studies in Classification, Data Analysis, and Knowledge Organization,
https://doi.org/10.1007/978-3-031-09034-9_8

$$(a+b)(t) = a(t) + b(t) \quad \text{and} \quad T_{a+b} = T_a \cup T_b$$
$$(a \cdot b)(t) = a(t) \cdot b(t) \quad \text{and} \quad T_{a \cdot b} = T_a \cap T_b$$

Let $T_V(v) \subseteq \mathcal{T}$, $T_V \in \mathcal{P}$, be the activity set for a node $v \in \mathcal{V}$ and $T_L(\ell) \subseteq \mathcal{T}$, $T_L \in \mathcal{W}$, the activity set for a link $\ell \in \mathcal{L}$. The following *consistency condition* must be fulfilled for activity sets: If a link $\ell(u, v)$ is active at the time point t then its end-nodes u and v should be active at the time point t : $T_L(\ell(u, v)) \subseteq T_V(u) \cap T_V(v)$.

In the following we will need

1. *Total*: $\text{total}(a) = \sum_i (f_i - s_i) \cdot v_i$
2. *Average*: $\text{average}(a) = \dfrac{\text{total}(a)}{|T_a|}$ where $|T_a| = \sum_i (f_i - s_i)$
3. *Maximum*: $\max(a) = \max_i v_i$

To support the computations with TQs we developed in Python the libraries TQ and Nets, see `https://github.com/bavla/TQ` .

2 Traditional (Generalized) Blockmodeling Scheme

A *blockmodel* (BM) [11] consists of structures obtained by identifying all units from the same cluster of the clustering / *partition* $\mathbf{C} = \{C_i\}$, $\pi(v) = i \Leftrightarrow v \in C_i$. Each pair of clusters (C_i, C_j) determines a block consisting of links linking C_i to C_j. For an exact definition of a blockmodel we have to be precise also about which blocks produce an arc in the *reduced graph* on classes and which do not, what is the *weight* of this arc, and in the case of generalized BM, of what *type*. The reduced graph can be represented by relational matrix, called also *image matrix*.

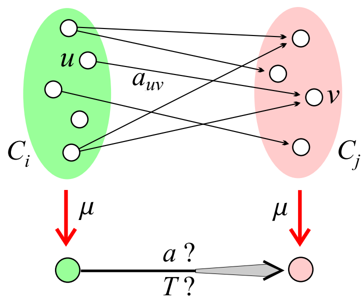


Fig. 1 Blockmodel.

To develop a BM method we specify a criterion function $P(\mu)$ measuring the "error" of the BM μ. We can introduce additional knowledge by constraining the partitions to a set Φ of feasible partitions. We are searching for a partition $\pi^* \in \Phi$ such that the corresponding BM μ^* minimizes the criterion function $P(\mu)$.

3 BM of Temporal Networks

For an early attempt of temporal network BM see [2, 5]. To the traditional BM scheme we add the time dimension. We assume that the network is described using temporal quantities [2] for nodes/links activity/presence, and some nodes properties and links weights. Then also the BM partition π is described for each node v with a temporal quantity $\pi(v,t)$: $\pi(v,t) = i$ means that in time t node v belongs to cluster i. The structure and activity of clusters $C_i(t) = \{v : \pi(v,t) = i\}$ can change through time, but they preserve their identity.

For the BM μ the clusters are mapped into BM nodes $\mu : C_i \longrightarrow [i]$. To determine the BM we still have to specify how the links from C_i to C_j are represented in the BM – in general, for the model arc $([i],[j])$, we have to specify two TQs: its *weight* a_{ij} and, in the case of generalized BM, its *type* τ_{ij}. The weight can be an object of a different type than the weights of the block links in the original temporal network.

We assume that in a temporal network $\mathcal{N} = (\mathcal{V}, \mathcal{L}, \mathcal{T}, \mathcal{P}, \mathcal{W})$ the links weight is described by a TQ $w \in \mathcal{W}$. In the following years we intend to develop BM methods case by case.

1. constant partition – nodes stay in the same cluster all the time:
   a. indirect approach based on clustering of TQs: $p(v) = \sum_{u \in N(v)} w(v,u)$, hierarchical clustering and leaders;
   b. indirect approach by conversion to the *clustering with relation constraint* (CRC);
   c. direct approach by (local) optimization of the criterion function P over Φ
2. dynamic partition – nodes can move between clusters through time. The details are still to be elaborated.

In this paper, we present approaches for cases 1.a and 1.b.

In the literature there exist other approaches to BM of temporal networks. A recent overview is available in the book [12].

3.1 Adapted Symbolic Clustering Methods

In [8] we adapted traditional leaders [13, 10] and agglomerative hierarchical [14, 1] clustering methods for clustering of modal-valued symbolic data. They can be almost directly applied for clustering units described by variables that have for their values temporal quantities.

For a unit X_i, each variable V_j is described with a size h_{ij} and a temporal quantity $\mathbf{x}_{ij}$, $X_{ij} = (h_{ij}, \mathbf{x}_{ij})$. In our algorithms we use *normalized* values of temporal variables $V' = (h, \mathbf{p})$ where

$$\mathbf{p} = [(s_r, f_r, p_r) : r = 1, 2, \ldots, k] \qquad \text{and} \qquad p_r = \frac{v_r}{h}$$

In the case, when $h = \text{total}(\mathbf{x})$, the normalized TQ $\mathbf{p}$ is essentially a probability distribution.

Both methods create cluster representatives that are represented in the same way.

3.2 Clustering of Temporal Network and CRC

To use the CRC in the construction of a nodes partition we have to define a dissimilarity measure $d(u, v)$ (or a similarity $s(u, v)$) between nodes. An obvious solution is $s(u, v) = f(w(u, v))$, for example

1. *Total activity*: $s_1(u, v) = \text{total}(w(u, v))$
2. *Average activity*: $s_2(u, v) = \text{average}(w(u, v))$
3. *Maximal activity*: $s_3(u, v) = \max(w(u, v))$

We can transform a similarity $s(u, v)$ into a dissimilarity by $d(u, v) = \frac{1}{s(u,v)}$ or $d(u, v) = S - s(u, v)$ where $S > \max_{u,v} s(u, v)$. In this way, we transformed the temporal network partitioning problem into a clustering with relational constraints problem [6, 360–369]. It can be efficiently solved also for large sparse networks.

3.3 Block Model

Having the partition π, to produce a BM we have to specify the values on its links. There are different options for model links weights $a(([i], [j]))$.

1. *Temporal quantities*: $a(([i], [j])) = \text{activity}(C_i, C_j) = \sum_{u \in C_i, v \in C_j} w(u, v)$, for $i \neq j$, and $a(([i], [i])) = \frac{1}{2}\text{activity}(C_i, C_i)$.
2. *Total intensities:* $a_t(([i], [j])) = \text{total}(a(([i], [j])))$.
3. *Geometric average intensities:* $a_g(([i], [j])) = \dfrac{a_t(([i], [j]))}{\sqrt{|C_i| \cdot |C_j|}}$.

4 Example: September 11th Reuters Terror News

The *Reuters Terror News* network was obtained from the CRA (Centering Resonance Analysis) networks produced by Steve Corman and Kevin Dooley at Arizona State University. The network is based on all the stories released during 66 consecutive days by the news agency Reuters concerning the September 11 attack on the U.S., beginning at 9:00 AM EST 9/11/01.

The nodes, $n = 13332$, of this network are important words (terms). For a given day, there is an edge between two words iff they appear in the same utterance (for details see the paper [9]). The network has $m = 243447$ edges. The weight of an edge is its daily frequency. There are no loops in the network. The network Terror News is undirected – so will be also its BM.

The Reuters Terror News network was used as a case network for the Viszards visualization session on the Sunbelt XXII International Sunbelt Social Network Conference, New Orleans, USA, 13-17. February 2002. It is available at `http://vlado.fmf.uni-lj.si/pub/networks/data/CRA/terror.htm`.

We transformed the Pajek version of the network into NetsJSON format used in libraries TQ and Nets. For a temporal description of each node/word for clustering we took its activity (sum of all TQs on edges adjacent to a given node v)

$$\mathrm{act}(v) = \sum_{u \in N(v)} w(v : u).$$

Our leaders' and hierarchical clustering methods are compatible – they are based on the same clustering error criterion function. Usually, the leaders' method is used to reduce a large clustering problem to up to some hundred units. With hierarchical clustering of the leaders of the obtained clusters, we afterward determine the "right" number of clusters and their representatives.

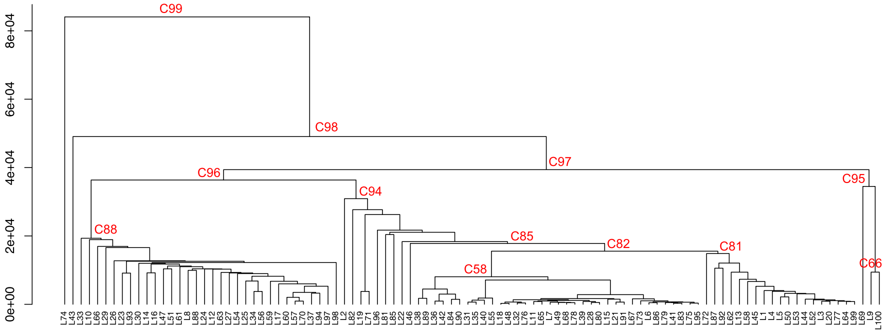


Fig. 2 Hierarchical clustering of 100 leaders in Terror News.

To cluster all 13332 words (nodes) in Terror News we used the adapted leaders' method searching for 100 clusters. We continued with the hierarchical clustering of the obtained 100 leaders. The result is presented in the dendrogram in Figure 2.

Fig. 3 Word clouds for clusters $C58$ and $C81$.

To get an insight into the content of a selected cluster we draw the corresponding word cloud based on the cluster's leader. In Figure 3 the word clouds for clusters $C58$ and $C81$ ($|C58| = 1396$, $|C81| = 2226$) are presented.

We can also compare the activities of pairs of clusters by considering the overlap of p-components (probability distributions) of their leaders. In Figure 4, we compare cluster $C58$ with cluster $C81$, and cluster $L96$ with cluster $C66$. In the right diagram some values are outside the display area: $L96[15] = 0.3524$, $C66[4] = 0.1961$, $C66[5] = 0.2917$.

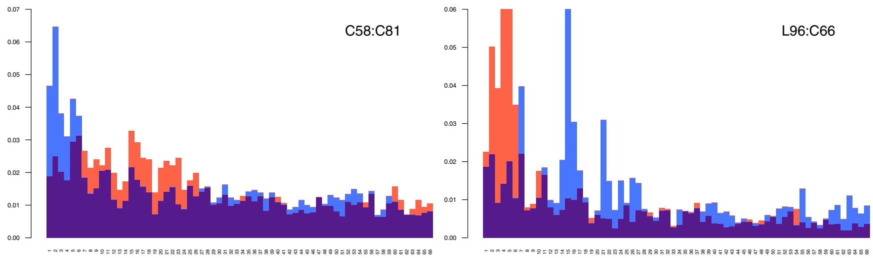
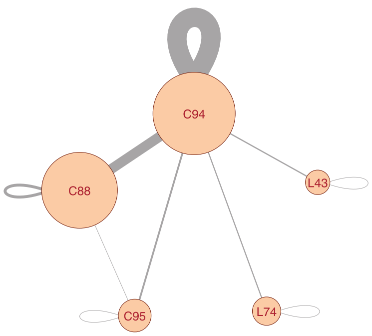


Fig. 4 Comparing activities of clusters (blue – first cluster, red – second cluster, violet – overlap).

We decided to consider in the BM the clustering of Terror News into 5 clusters $\mathbf{C} = \{C94, C88, C95, L43, L74\}$. The split of cluster C95 gives clusters of sizes 325 and 629 (for sizes, see the right side of Figure 5). Both clusters C94 and C88 have a chaining pattern at their top levels.

Because of large differences in the cluster sizes, it is difficult to interpret the total intensities image matrix. An overall insight into the BM structure we get from the geometric average intensities image matrix (right side) and the corresponding BM network (cut level 0.3), left side of Figure 5.



i	cluster	1	2	3	4	5
1	$C94$	23.85	12.23	2.26	1.57	1.42
2	$C88$		3.58	0.33	0.22	0.19
3	$C95$			0.56	0.07	0.07
4	$L43$				0.38	0.08
5	$L74$					0.39
	size	6018	5109	954	535	716

Fig. 5 Block model and image matrix.

A more detailed BM is presented by the activities (p-components) image matrix in Figure 6.

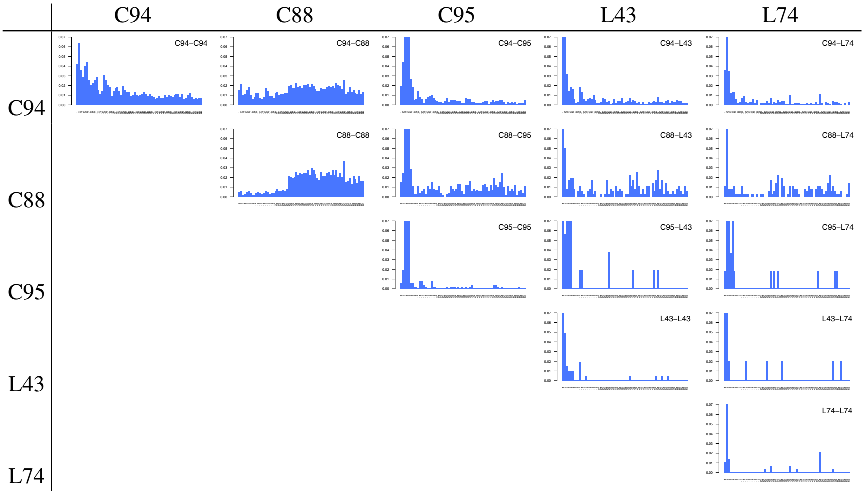


Fig. 6 BM represented as p-components of temporal activities of links between pairs of clusters.

A more compact representation of a temporal BM is a heatmap display of this matrix in Figure 7. Because of some relatively very large values, it turns out that the display of the matrix with logarithmic values provides much more information.

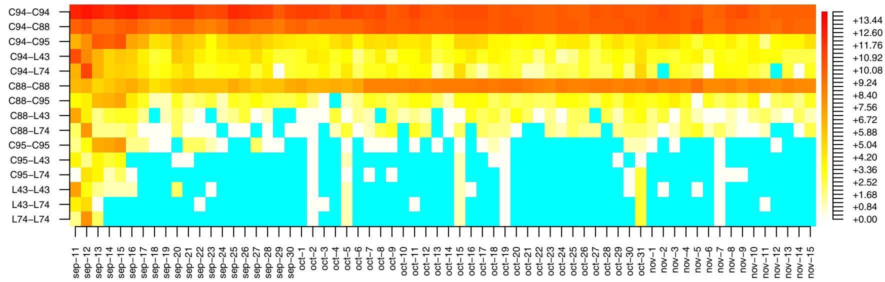


Fig. 7 BM heatmap with log2 values.

To the Terror News network, we applied also the clustering with relational constraints approach. Because of the limited space available for each paper, we can not present it here. A description of the analysis with the corresponding code is available at `https://github.com/bavla/TQ/wiki/BMRC` .

5 Conclusions

The presented research is a work in progress. It only deals with the two simplest cases of temporal blockmodeling. We provided some answers to the problem of normalization of model weights TQs when comparing them and some ways to present/display the temporal BMs.

We used different tools (R, Python, and Pajek) to obtain the results. We intend to provide the software support in a single tool – probably in Julia. We also intend to create a collection of interesting and well-documented temporal networks for testing and demonstrating the developed software.

Acknowledgements The paper contains an elaborated version of ideas presented in my talks at the XXXX Sunbelt Social Networks Conference (on Zoom), July 13-17, 2020 and at the EUSN 2021 – 5th European Conference on Social Networks, Naples (on Zoom), September 6-10, 2021.

This work is supported in part by the Slovenian Research Agency (research program P1-0294 and research projects J1-9187, J1-2481, and J5-2557), and prepared within the framework of the HSE University Basic Research Program.

References

1. Anderberg, M.R.: Cluster Analysis for Applications. Academic Press, New York (1973)
2. Batagelj, V., Praprotnik, S.: An algebraic approach to temporal network analysis based on temporal quantities. Soc. Netw. Anal. Min. **6**(1), 1–22 (2016)
3. Batagelj, V., Ferligoj, A.: Clustering relational data. In: Gaul, W., Opitz, O., Schader, M. (Eds.) Data Analysis / Scientific Modeling and Practical Application, pp. 3–15. Springer (2000)
4. Batagelj, V.: Generalized Ward and related clustering problems. In: Bock H.-H. (ed) Classification and Related Methods of Data Analysis, pp. 67–74. North-Holland, Amsterdam (1988)
5. Batagelj, V., Ferligoj, A., Doreian, P.: Indirect blockmodeling of 3-way networks. In: Brito, P., Bertrand, P., Cucumel, G., de Carvalho, F. (eds.) Selected Contributions in Data Analysis and Classification, pp. 151–159. Springer (2007)
6. Batagelj, V., Doreian, P., Ferligoj, A., Kejžar, N.: Understanding Large Temporal Networks and Spatial Networks: Exploration, Pattern Searching, Visualization and Network Evolution. Wiley (2014)
7. Batagelj, V., Kejžar, N.: Clamix – Clustering Symbolic Objects (2010) Program in R `https://r-forge.r-project.org/projects/clamix/`
8. Kejžar, N., Korenjak-Černe, S., Batagelj, V.: Clustering of modal-valued symbolic data. Adv. Data Anal. Classif. **15**, pp. 513—541 (2021)
9. Corman, S. R., Kuhn, T., McPhee, R. D., Dooley, K. J.: Studying complex discursive systems: Centering resonance analysis of communication. Hum. Commun. Res. **28**(2), 157–206 (2002)
10. Diday, E.: Optimisation en Classification Automatique. Tome 1.,2. INRIA, Rocquencourt (in French) (1979)
11. Doreian, P., Batagelj, V., Ferligoj, A.: Generalized Blockmodeling. Structural Analysis in the Social Sciences. Cambridge University Press (2005)
12. Doreian, P., Batagelj, V., Ferligoj, A. (Eds.) Advances in Network Clustering and Blockmodeling. Wiley (2020)
13. Hartigan, J. A.: Clustering Algorithms. Wiley-Interscience, New York (1975)
14. Ward, J. H.: Hierarchical grouping to optimize an objective function. J. Am. Stat. Asso.c **58**, 236–244 (1963)