

## Temporal blockmodeling

#### V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror news

Block mode

Conclusions

References

# Toward an indirect approach for blockmodeling temporal networks

Vladimir Batagelj

IMFM Ljubljana, IAM UP Koper and NRU HSE Moscow

XXXX Sunbelt Social Networks Conference on Zoom, July 13-17, 2020



#### Outline

### Temporal blockmodeling

#### V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror new

Block mode

. . .

Conclusions

- 1 Temporal blockmodeling
- 2 Clustering of TQs
- 3 Example: Terror news
- 4 Block model
- 5 Conclusions
- 6 References



Vladimir Batagelj: vladimir.batagelj@fmf.uni-lj.si

Current version of slides (July 11, 2020 at 22:08): slides PDF

https://github.com/bavla/SocNet/blob/master/Temp/SNAauthors.pdf



### Blockmodeling temporal networks

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

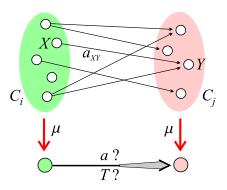
Clustering of TQs

Example: Terror new

Block mode

\_ . .

\_\_\_\_\_



To the traditional blockmodeling (BM) scheme we add the time dimension. We assume that the network is described using temporal quantities [6] for nodes/links activity/presence, nodes properties and links weights. Then also the BM partition  $\pi$  can be described for each node v with a temporal quantity  $\pi(v,t)$  with the meaning:

 $\pi(v,t)=i$  means that in time t node v belongs to cluster i. The clusters  $C_i(t)=\{v:\pi(v,t)=i\}$  structure and activity can change through time, but they preserve their identity.



### Blockmodeling temporal networks

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror news

Block mode

. . .

For the BM  $\mu$  the clusters are maped into BM nodes  $\mu: C_i \to [i]$ . 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 temporal quantities: its value  $a_{ij}(t)$  and, in the case of generalized BM, its type  $T_{ij}(t)$ . In general the value can be an object of different type than the values of the links in the original temporal network.

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



### Blockmodeling temporal networks

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror news

Block mode

Conclusions

References

This is a general framework. In the following years we intend to develop methods case by case.

- constant partition nodes stay in the same cluster all the time
  - indirect approach based on clustering of temporal quantities
  - temporal version of clustering with relation constraint
  - local optimization of the criterion function P over  $\Phi$
- dynamic partition nodes can move between clusters through time. The details are still to be elaborated.



# Clustering of temporal quantities

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

## Clustering of TQs

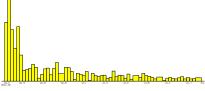
Terror new

Block mode

Conclusions

Referenc







A temporal quantity is described by a sequence

$$\mathbf{a} = [(s_r, f_r, v_r) : r = 1, 2, \dots, k]$$

where  $[s_r, f_r)$  determines a time interval and  $v_r$  is the value of variable on this interval.

We adapted traditional leaders and agglomerative hierarchical clustering methods for clustering units described by

variables that have for their values temporal quantities [2, 5].

ClusTQ at https://github.com/bavla/TQ.





## Clustering of temporal quantities

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

# Clustering of TQs

Example: Terror news

Block mode

Conclusions

Referenc

For a unit  $X_i$ , each variable  $V_j$  is described with a weight  $h_{ij}$  and a temporal quantity  $\mathbf{x}_{ij}$ 

$$X_{ij}=(h_{ij},\mathbf{x}_{ij})$$

In our algorithms we use *normalized* values of temporal variables  $Y' = (h, \mathbf{p})$  where

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

In the case, when  $h = tot(\mathbf{x}) = \sum v_r$ , the normalized TQ is essentially a probability distribution.

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



#### September 11th Reuters terror news

Temporal blockmodeling

V. Batagelj

Temporal blockmodelin

Clustering of TQs

Example: Terror news

Block model

Conclusion

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 of this network are important words (terms). There is an edge between two words iff they appear in the same utterance (for details see the paper [?]). The weight of an edge is its frequency. The network has n=13332 nodes (different words in the news) and m=243447 edges, 50859 with value larger than 1. There are no loops in the network.

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.



#### **Details**

## Temporal blockmodeling

#### V. Batagelj

Temporal blockmodeling

Clustering of TQs

#### Example: Terror news

Block mode

Conclusion

References

We transformed the Pajek version of the network into NetsJSON format used in  $\mathsf{TQ}$  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.

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 the 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.



#### Leaders method

# Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror news

Block mode

Conclusions

Reference

To cluster all 13332 words (nodes) in Terror news we used an adapted leaders method searching for 100 clusters:

After 50 steps we stopped the search Pr. — 41012 04

After 50 steps we stopped the search.  $P_1 = 41012.94$ ,  $P_{50} = 23897.74$ 

```
{74: 716, 43: 535, 82: 378, 2: 372, 9: 338, 69: 325, 96: 307, 46: 307, 100: 291, 26: 275, 62: 257, 13: 241, 85: 238, 81: 237, 34: 233, 98: 229, 27: 228, 29: 222, 19: 203, 22: 199, 10: 196, 88: 195, 37: 192, 12: 191, 30: 186, 54: 183, 72: 180, 66: 177, 14: 175, 33: 172, 25: 162, 23: 162, 92: 156, 71: 155, 93: 153, 87: 152, 51: 151, 24: 150, 16: 148, 58: 146, 59: 146, 47: 145, 63: 143, 61: 141, 4: 139, 45: 132, 89: 130, 8: 128, 17: 126, 56: 114, 50: 105, 77: 102, 42: 101, 1: 96, 55: 96, 97: 95, 53: 91, 94: 89, 5: 87, 70: 83, 90: 81, 44: 76, 3: 75, 60: 74, 40: 73, 52: 64, 15: 63, 84: 56, 20: 51, 57: 45, 64: 44, 67: 44, 66: 43, 21: 39, 31: 37, 11: 37, 39: 36, 79: 35, 38: 35, 36: 34, 76: 32, 91: 32, 99: 32, 80: 31, 73: 31: 68: 31, 7: 29, 6: 29, 35: 29, 28: 28, 49: 27, 86: 26, 78: 23, 83: 18, 41: 18, 32: 18, 95: 15. 18: 14. 48: 13, 75: 129
```



## Hierarchical clustering

## Temporal blockmodeling

#### V. Batagelj

Temporal blockmodeling

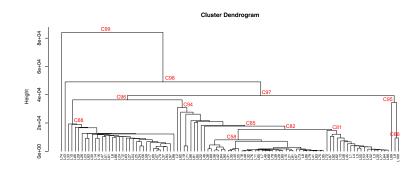
Clustering of TQs

#### Example: Terror news

Block mode

Conclusion

Reference





#### Word clouds for L74 and L43

 $|L74| = 716, \quad |L43| = 535$ 

## Temporal blockmodeling

#### V. Batagelj

blockmodeling

Clustering of TQs

#### Example: Terror news

Block mode

Conclusions

Reference

#### resumption manual 500 morgan\_stanlly\_dw good morning america crude jerusalem salt nous rkforce control surplus sadnes & Samuella sadnes & Samuella surplus sadnes & Samuella sadnes & Sa prime-time tuel-jaden/52-year swettending lagship gocasional swettending lagship gocasional gocasio t maine patriotism gasoline patriotism motorist petroleum akamai logan\_intl\_airport wholesale united\_states-bound liquidity controller

lady pearl\_barbor louisianapotomac suburban combination ted shut shard retail spectacular billood temperature plantus sorrow tremendous sorrow sorrow sorrow tremendous sorrow



# Word clouds for L46 and C46 |L46| = 307, |C46| = 358

## Temporal blockmodeling

V. Batagelj

blockmodeling

Clustering of TQs

Example: Terror news

Block mode

Conclusion

Reference

```
baggaley hekmatyar predecessor lavenda actorsweeney gerotesque e lavenda actorsweeney e lavenda actorsweeney e lavenda actorsweeney e lavenda actorsweeney e lavenda e la lavenda e lav
```

inc web metal manhattan firm sinceple analyst a gemergency sinceple analyst unprecedented morning tuesday



## Word clouds for C88, C95 $|C88| = 5109, \quad |C95| = 954$

Temporal blockmodeling

V. Batageli

Clustering of

Example: Terror news

al\_quaeda troop emocrat at a fight of the state suspicious

postal abdullah wave senate fighter special anatoge night special and special warfare diseaseweapon\_c antibiotic process mail warnnorthern

process mail warmorthern weapont should be acterium new form southern test campaign from campaign from targetun abdul future rumsfeld

afghanistan talihan

# resolution

al-shehhi recorder engineer bookum oklahoma signal electronic john\_f\_kennedyqaeda communist marwan aznar

wireless desk grasso pull rain habib laguardia barrier do delaware vigilauthor delaware vigilauthorize restrict oracletrooper of the market of the amidtown bangzao

latin biden applaud On glederal reservebank applaud Convention iik a

schumer cripple reservist advertise commandeer forthamburg remembrance

construction kennedy

4 D > 4 A > 4 B > 4 B >



#### Word clouds for C58 and C81

 $|C58| = 1396, \quad |C81| = 2226$ 

Temporal blockmodeling

V. Batagelj

blockmodeling

Clustering of TQs

Example: Terror news

Block mode

Conclusions

Referenc

thursday maional timepercent newspaper world trade\_ctr major airlinealwindry white\_nouse and was a saturday of the major airlinealwindry white\_nouse and tell aircraft way office to tell aircraft way





#### Word cloud for C94

|C94| = 6018,  $C58, C81 \subset C94$ 

### Temporal blockmodeling

#### V. Batagelj

Temporal blockmodeling

Clustering of TQs

#### Example: Terror news

Block mode

Conclusions

References

```
united states
                                                                                          percent possible agency nation government new_york
                                                                                                                                                                                                                                                                                                flight 능
                            operation washington military time
                                           passerger & 5 fbi COUNTRY airlearter auroriter auroriter
                                                                                                                                                                                                                                                                                                                                 owhite_house
                       foreign way threat pakistanstrike
                                  network number behand meter a tuesday by the federal authority by facility chief capital day will be a symbol and the federal authority by facility chief capital day will be a symbol and the federal authority by facility chief pentagon leader include Security
ores_bush
                                                                          company attack thursday employee president sounday american terrorismplane member spokesman wednesday news official aircraft year saturday
                               office company
                                                                                                                           bin läden
```



# Comparisons of leaders and cluster representatives L74:C98, L43:C97, C96:C95, C88:C94

#### Temporal blockmodeling

#### V. Batagelj

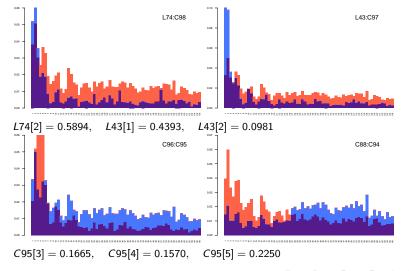
nporal kmodelinį

Clustering of TQs

#### Example: Terror news

Block mode

Conclusions





# Comparisons of leaders and cluster representatives C58:C81, L96:C66, C46:C955, C46

Temporal blockmodeling

V. Batagelj

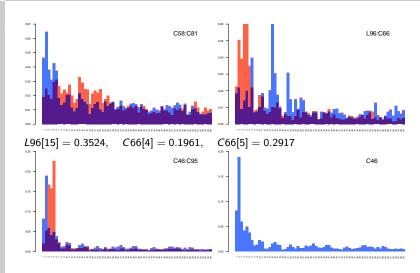
Temporal blockmodeling

Clustering of TQs

Example: Terror news

Block mode

Conclusions





#### Block model

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror new

Block model

Conclusions

Referenc

To produce a block model we have to specify the values on its links. Because the original network is undirected, so is the BM. There are different options.

#### Temporal quantities:

$$\mathbf{a}(([i]:[j])) = \mathsf{activity}(C_i,C_j) = \sum_{u \in C_i, v \in C_j} \mathbf{a}(u:v), \quad \mathsf{for} \ i \neq j$$

and 
$$\mathbf{a}(([i]:[i])) = \frac{1}{2} \operatorname{activity}(C_i, C_i)$$
.

Total intensities: 
$$a(([i]:[j])) = tot(a(([i]:[j])))$$
.

Geometric average intensities: 
$$g(([i]:[j])) = \frac{a(([i]:[j]))}{\sqrt{|C_i| \cdot |C_j|}}$$
.



#### Block model

# Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror nev

Block model

Conclusions

Referenc

#### **Total intensities:**

i	cluster	size	1	2	3	4	5
1	C94	6018	143549	67801	5422	2816	2939
2	C88	5109	0	18288	739	357	357
3	C95	954	0	0	535	53	54
4	L43	535	0	0	0	205	51
5	L74	716	0	0	0	0	281

#### Geometric average intensities:

i	cluster	1	2	3	4	5
1	C94	23.85	12.23	2.26	1.57	1.42
2	C88	0	3.58	0.33	0.22	0.19
3	C95	0	0	0.56	0.07	0.07
4	L43	0	0	0	0.38	0.08
5	L74	0	0	0	0	0.39



# Block model Geometric average intensities, cut at level 0.3

Temporal blockmodeling

V. Batagelj

Temporal

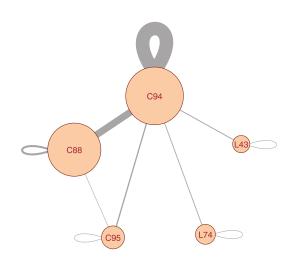
Clustering of TQs

Example: Terror news

Block model

Conclusions

References





# Hierarchical clustering of Franzosi's violence network

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

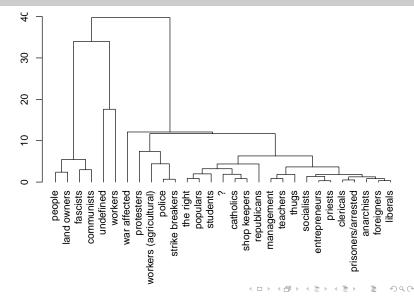
Clustering of TQs

Example: Terror news

Block model

Conclusions

. . . . . . . .





## Conclusions

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror news

Block mode

Conclusions

References



### Acknowledgments

Temporal blockmodeling

V. Batagelj

Temporal blockmodelin

Clustering of TQs

Example: Terror new

Block mode

Conclusions

References

This work is supported in part by the Slovenian Research Agency (research program P1-0294 and research projects J1-9187 and J7-8279), project COSTNET (COST Action CA15109), and by Russian Academic Excellence Project '5-100'.



#### References I

Temporal blockmodeling

V. Batagelj

Temporal blockmodeling

Clustering of TQs

Example: Terror new

Rlock mode

. . .

References

Anderberg, M.R. (1973) Cluster analysis for applications. Academic Press, New York

Batagelj, V., Kejžar, N., Korenjak-Černe, S. (2015) Clustering of Modal Valued Symbolic Data. arXiv:1507.06683

Batagelj, V., Kejžar, N. (2010) Clamix – Clustering Symbolic Objects. Program in R

https://r-forge.r-project.org/projects/clamix/

Batagelj, V. (1988) Generalized Ward and related clustering problems. In: Bock HH (ed) Classification and related methods of data analysis, North-Holland, Amsterdam, pp 67-74

Batagelj, V. (2020) TQ / clusTQ – clustering temporal quantities. https://github.com/bavla/TQ

Batagelj, V., Praprotnik, S. (2016) An algebraic approach to temporal network analysis based on temporal quantities. Social Network Analysis and Mining, 6(1), 1-22.

Diday, E. (1979) Optimisation en classification automatique. Tome 1.,2. INRIA, Rocquencourt (in French)



#### References II

Temporal blockmodeling

V. Batagelj

Temporal blockmodelin

Clustering of TQs

Example: Terror new

Block model

Conclusions

References



Hartigan, J.A. (1975) Clustering algorithms. Wiley-Interscience, New York



Ward, J.H. (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236-244