Mining Temporal Networks

Polina Rozenshtein
Nordea Data Science Lab
Finland
rozenshtein.polina@gmail.com

Aristides Gionis
Aalto University
Finland
aristides.gionis@aalto.fi

ABSTRACT

Networks (or graphs) are used to represent and analyze large datasets of objects and their relations. Naturally, real-world networks have a temporal component: for instance, interactions between objects have a timestamp and a duration. In this tutorial we present models and algorithms for mining temporal networks, i.e., network data with temporal information. We overview different models used to represent temporal networks. We highlight the main differences between static and temporal networks, and discuss the challenges arising from introducing the temporal dimension in the network representation. We present recent papers addressing the most well-studied problems in the setting of temporal networks, including computation of centrality measures, motif detection and counting, community detection and monitoring, event and anomaly detection, analysis of epidemic processes and influence spreading, network summarization, and structure prediction.

CCS CONCEPTS

• **Information systems → Data mining**; • **Networks → Network algorithms**.

KEYWORDS

data mining, graph mining, temporal networks

ACM Reference Format:
Polina Rozenshtein and Aristides Gionis. 2019. Mining Temporal Networks. In *The 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '19), August 4–8, 2019, Anchorage, AK, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3292500.3332295

TARGET AUDIENCE AND PREREQUISITES

The target audience for this tutorial is graduate students, researchers, and practitioners, who are interested in the analysis of networks with a temporal component. A basic familiarity with methods and concepts of network analysis is assumed. In the first part of the tutorial we broadly overview models and measures, thus, this part does not require any technical background. In the second part we cover problem formulations and algorithmic techniques. This part it targeted to researchers who (wish to) work in this research area, as well as developers and practitioners looking for solutions.

KDD '19, August 4–8, 2019, Anchorage, AK, USA

ACM ISBN 978-1-4503-6201-6/19/08.
https://doi.org/10.1145/3292500.3332295

TUTORS

Polina Rozenshtein received a MSc and an PhD degrees from Aalto University, Espoo, Finland, in 2014 and 2018. Her PhD thesis is on the topic of Temporal Networks Analysis [27]. She is currently a senior data scientist at Nordea Data Science Lab, prior to that she was a post-doctoral researcher in Data Mining group of Computer Science Department at Aalto University. Her research interests include data mining, combinatorial optimization, dynamic graph mining, and social-networks analysis. She has published several papers on the topic of graph mining and temporal-network analysis. Her MSc thesis received an award by the Finnish Society of Computer Science for best MSc thesis in Finland in 2015. Her paper "Event detection in activity networks" received the best student paper award in ECML PKDD 2014.

Aristides Gionis is a professor in the department of Computer Science in Aalto University. He is currently a fellow in the ISI foundation, Turin. Previously he has been a senior research scientist and group leader in Yahoo! Research, Barcelona. He obtained his PhD in 2003 from Stanford University, USA. He is currently serving as an action editor in the Data Management and Knowledge Discovery journal (DMKD), an associate editor in the ACM Transactions on Knowledge Discovery from Data (TKDD), and an associate editor in the ACM Transactions on the Web (TWEB). He has contributed in several areas of data science, such as algorithmic data analysis, web mining, social-media analysis, data clustering, and privacy-preserving data mining. His current research is funded by the Academy of Finland (projects Nestor, Agra, AIDA) and the European Commission (project SoBigData).

TUTORIAL OUTLINE

The total presentation time is 3 hours. The outline of the tutorial is as follows.

Part one:

(1) **Introduction [15 min]**
  (a) Motivation and application areas.
  (b) Main definitions and different types of temporal [6, 12, 13, 19, 23].
(2) **Models of temporal networks [30 min]**
  (a) Representation of temporality: static networks with aggregated temporal information [2]; dynamic and time-evolving networks with sequential updates [17]; sequences of network snapshots [1]; sequences of interactions with meta information [13].
  (b) Combinatorial models: multi-graphs, labeled graphs, sets of temporal edges [23]; statistical models: generative models, parameter fitting [10, 16]; dynamical processes [4, 11].
(3) **Characterization of temporal networks and efficient computation of network measures [30 min]**

(a) Centrality, connectivity, density measures, paths, trees, subgraphs, cycles, motifs, etc. [13, 19].
(b) Frequent patterns and episodes in temporal networks [14, 33].
(c) Measures of regime change in temporal networks, rules of evolution, etc. [5, 15].

(4) **Group work [15 min]**

Part two:

(5) **Algorithmic approaches [25 min]**
(a) Streaming model, sliding-window model, sequential updates [18, 22]
(b) Theoretical foundations of network analysis [23]

(6) **Data Mining problems [40 min]**
(a) Community detection [26, 30].
(b) Event detection [3, 7, 9, 24, 28, 34].
(c) Epidemics analysis and influence spreading [20, 25, 29, 35].
(d) Network summarization [21, 31, 32].
(e) Structural prediction [8, 36].

(7) **Challenges, open problems, and trends [10 min]**

(8) **Group work [15 min]**

PREVIOUS EDITIONS AND SIMILAR TUTORIALS

This is a new tutorial and it has not been presented previously. To the best of our knowledge, a similar tutorial covering a broad range of topics in mining temporal networks has not been presented before. However, some parts of this tutorial overlap with previous tutorials that have been presented by other researchers. In particularly, Part 4.1 on algorithmic approaches and streaming models is partially related to the tutorial on "Sampling, Sketching, Streaming, Small-Space Optimization: Algorithmic Approaches for Analyzing Large Graphs,"[1] presented by Sudipto Guha and Andrew McGregor in KDD 2018. A number of surveys [6, 12, 13, 19, 23] on the topic of temporal networks is available.

TUTORIAL MATERIAL

The KDD 2019 edition of the tutorial is accompanied by a website,[2] which contains a full list of references with links to electronic editions, and the slides used in the tutorial presentation.

Acknowledgements. This work was supported by three Academy of Finland projects (286211, 313927, and 317085), and the EC H2020 RIA project "SoBigData" (654024).

REFERENCES

[1] C. Aggarwal and K. Subbian. Evolutionary network analysis: A survey. *ACM Computing Surveys (CSUR)*, 47(1):10, 2014.
[2] C. C. Aggarwal and H. Wang. Graph data management and mining: A survey of algorithms and applications. In *Managing and mining graph data*, pages 13–68. Springer, 2010.
[3] L. Akoglu, H. Tong, and D. Koutra. Graph based anomaly detection and description: a survey. *Data mining and knowledge discovery*, 29(3):626–688, 2015.
[4] A.-L. Barabási et al. *Network science*. Cambridge university press, 2016.
[5] M. Berlingerio, F. Bonchi, B. Bringmann, and A. Gionis. Mining graph evolution rules. In *ECML PKDD*, pages 115–130. Springer, 2009.
[6] A. Casteigts, P. Flocchini, W. Quattrociocchi, and N. Santoro. Time-varying graphs and dynamic networks. *International Journal of Parallel, Emergent and Distributed Systems*, 27(5):387–408, 2012.
[7] M. Cordeiro and J. Gama. Online social networks event detection: a survey. In *Solving Large Scale Learning Tasks. Challenges and Algorithms*, pages 1–41. Springer, 2016.
[8] Y. Dhote, N. Mishra, and S. Sharma. Survey and analysis of temporal link prediction in online social networks. In *2013 International Conference on Advances in Computing, Communications and Informatics*, pages 1178–1183. IEEE, 2013.
[9] A. Goswami and A. Kumar. A survey of event detection techniques in online social networks. *Social Network Analysis and Mining*, 6(1):107, 2016.
[10] S. A. Hill and D. Braha. Dynamic model of time-dependent complex networks. *Physical Review E*, 82(4):046105, 2010.
[11] P. Holme. Epidemiologically optimal static networks from temporal network data. *PLoS computational biology*, 9(7):e1003142, 2013.
[12] P. Holme. Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88(9):234, 2015.
[13] P. Holme and J. Saramäki. Temporal networks. *Physics reports*, 519(3):97–125, 2012.
[14] H.-P. Hsieh and C.-T. Li. Mining temporal subgraph patterns in heterogeneous information networks. In *2010 IEEE Second International Conference on Social Computing*, pages 282–287. IEEE, 2010.
[15] A. Impedovo, C. Loglisci, and M. Ceci. Temporal pattern mining from evolving networks. *arXiv preprint arXiv:1709.06772*, 2017.
[16] H.-H. Jo, R. K. Pan, and K. Kaski. Emergence of bursts and communities in evolving weighted networks. *PloS one*, 6(8):e22687, 2011.
[17] R. Kumar, J. Novak, and A. Tomkins. Structure and evolution of online social networks. In *Link mining: models, algorithms, and applications*, pages 337–357. Springer, 2010.
[18] R. Kumar, T. Calders, A. Gionis, and N. Tatti. Maintaining sliding-window neighborhood profiles in interaction networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 719–735, 2015.
[19] M. Latapy, T. Viard, and C. Magnien. Stream graphs and link streams for the modeling of interactions over time. *Social Network Analysis and Mining*, 8(1):61, 2018.
[20] S. Lee, L. E. Rocha, F. Liljeros, and P. Holme. Exploiting temporal network structures of human interaction to effectively immunize populations. *PloS one*, 7 (5), 2012.
[21] Y. Liu, T. Safavi, A. Dighe, and D. Koutra. Graph summarization methods and applications: A survey. *ACM Computing Surveys (CSUR)*, 51(3):62, 2018.
[22] A. McGregor. Graph stream algorithms: a survey. *ACM SIGMOD Record*, 43(1): 9–20, 2014.
[23] O. Michail. An introduction to temporal graphs: An algorithmic perspective. *Internet Mathematics*, 12(4):239–280, 2016.
[24] A. Nurwidyantoro and E. Winarko. Event detection in social media: A survey. In *International Conference on ICT for Smart Society*, pages 1–5. IEEE, 2013.
[25] M. G. Rodriguez and B. Schölkopf. Influence maximization in continuous time diffusion networks. *arXiv preprint arXiv:1205.1682*, 2012.
[26] G. Rossetti and R. Cazabet. Community discovery in dynamic networks: a survey. *ACM Computing Surveys (CSUR)*, 51(2):35, 2018.
[27] P. Rozenshtein. *Methods for analyzing temporal networks*. PhD thesis, Aalto University, Helsinki, Finland, 2018.
[28] P. Rozenshtein, A. Anagnostopoulos, A. Gionis, and N. Tatti. Event detection in activity networks. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1176–1185. ACM, 2014.
[29] P. Rozenshtein, A. Gionis, B. A. Prakash, and J. Vreeken. Reconstructing an epidemic over time. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1835–1844. ACM, 2016.
[30] P. Rozenshtein, N. Tatti, and A. Gionis. Finding dynamic dense subgraphs. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 11(3):27, 2017.
[31] P. Rozenshtein, N. Tatti, and A. Gionis. The network-untangling problem: From interactions to activity timelines. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 701–716. Springer, 2017.
[32] P. Rozenshtein, F. Bonchi, A. Gionis, M. Sozio, and N. Tatti. Finding events in temporal networks: Segmentation meets densest-subgraph discovery. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 397–406. IEEE, 2018.
[33] B. Wackersreuther, P. Wackersreuther, A. Oswald, C. Böhm, and K. M. Borgwardt. Frequent subgraph discovery in dynamic networks. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, pages 155–162. ACM, 2010.
[34] H. Xiao, P. Rozenshtein, and A. Gionis. Discovering topically-and temporally-coherent events in interaction networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 690–705. Springer, 2016.
[35] H. Xiao, P. Rozenshtein, N. Tatti, and A. Gionis. Reconstructing a cascade from temporal observations. In *Proceedings of the 2018 SIAM International Conference on Data Mining*, pages 666–674. SIAM, 2018.
[36] L. Zhu, D. Guo, J. Yin, G. Ver Steeg, and A. Galstyan. Scalable temporal latent space inference for link prediction in dynamic social networks. *IEEE Transactions on Knowledge and Data Engineering*, 28(10):2765–2777, 2016.

[1] https://people.cs.umass.edu/~mcgregor/graphs/
[2] https://rozensp.github.io/KDD19-tutorial-temporal/