Public call for co-financing of research projects in 2017

20. BRIEF DESCRIPTION OF THE RESEARCH PROJECT

1. Scientific background

In last two decades, the increased interest in temporal networks bas been generated in many different sources. This triggered a need for developing methods for handling such networks. Particularly important for analyzing temporal networks were the study of travel-support services and the analysis of sequences of events including e-mail traffic, news events and phone calls. These approaches and results are surveyed by Holme and Saramäki (2012,2013) and Holme (2015). See also Nicosia (2013) and Kempe, Kleinberg, and Kumar (2000). Another overview can be found in Casteigts, Flocchini, Quattrociocchi, and Santoro (2012) and Casteigts and Flocchini (2013) based on their formalization of temporal networks using time-varying graphs (TVGs).

In this project we build upon our solutions and experiences with temporal network in program Pajek (Pajek 0.47, July 1999) by incorporating ideas from these literatures to create a broader framework for analyzing temporal networks.

The idea of adding the time dimension to networks is not new and has been formulated in many disciplines. Some of the earliest efforts are in the transport(ation) network analysis Bell and Iida (1997) and Correa and Stier-Moses (2011); project scheduling (Critical Path Method (CPM), Pert) in Operations Research Moder and Phillips (1970) and constraints networks in Artificial Intelligence Dechter (2003).

While there are also qualitative approaches to temporal networks, for example Allen (1983) and Vilain, Kautz, and van Beek (1990), in this project we build on the quantitative approach. For statistical approaches, see Kolaczyk (2009) and the Siena page of Snijders and Ripley, Snijders, Boda, Vörös, and Preciado (2013).

Based on our survey of these sources, there are two useful views for temporal networks:

- Networks can be seen as providing constraints on the activities of actors in them or be determined by constraints (for example, the networks linking airports as determined by airline time tables), and
- Networks can represent interactions of events among actors (for example, the KEDS networks and citation networks).

It is useful to think of processes on networks in terms of 'what moves' over the network ties. Examples of this include:

- Travel (where people and goods move between places) and transmission networks such as electrical grids and postal systems;
- Diffusion (of cultural ideas, styles, information and technologies), broadcasting (of news, entertainment and advertisements) and election results stemming for network processes (funding candidates, mobilizing grass-root actions, facilitating voting and preventing voting and
- Flows of, for example, people commuting, water delivered over pipes or oil and gas being delivered.

Thinking of what moves or flows over network ties is helped by using some additional concepts.

Public call for co-financing of research projects in 2017

A lot of research is still confused with the terminology and the terms used in communication network analysis, transport networks, computer networks, etc. that are similar or even the same, define the same phenomenon with different notation and different words. For example temporal distance Xuan, Ferreira, and Jarry (2003), reachability time Holme (2005), latency of the information and other terms name the same thing in different areas. The same thing happens with journeys Xuan et al. (2003) that are named temporal paths, time respecting paths or paths with schedules by other authors. There is no established formal description of temporal networks.

The common point of all current research is the time component and that the changes of the network are one of the key information about the network. The beginnings of temporal network analysis are based on time slices of the network Santoro, Quattrociocchi, Flocchini, and Casteigts (2011). The temporal network is represented as a sequence of static networks, representing the state of the temporal network at a chosen time point (interval). Two different approaches aim to unify temporal networks theory in a way that could be used for all the different uses. One is the time-aggregated graph from George and Kim (2013). The other is the time-varying graph from Casteigts et al. (2012). We feel that both descriptions lack the possibility of adding arbitrary information to the network vertices or links. They are both describing the presence with explicit functions which also seems too complex. In Batagelj and Praprotnik (2016) we proposed a new way for the temporal network description which remedies both of these shortcomings.

2. Problem identification

In a temporal network, the presence and activity of nodes and links can change through time. In Batagelj and Praprotnik (2016) we introduced the notion of temporal quantities to describe temporal networks. We define the addition and multiplication of temporal quantities in a way that can be used for the definition of addition and multiplication of temporal networks. The corresponding algebraic structures are semirings. The usual approach to (data) analysis of temporal networks is to transform the network into a sequence of time slices -- static networks corresponding to selected time intervals and analyze each of them using standard methods to produce a sequence of results. The approach proposed in this paper enables us to compute these results directly. We developed fast algorithms for the proposed operations. They are available as an open source Python library TQ (Temporal Quantities). The proposed approach enables us to treat as temporal quantities also other network characteristics such as degrees, connectivity components, centrality measures, Pathfinder skeleton, etc. The main goal of the paper was to provide a theoretical foundations and to show “It can be done”. Many questions are still to be answered and the library TQ should be made more efficient for large networks. These are the tasks of the proposed project.

3. Objective of the proposed research with particular emphasis on the originality of the proposed research and its potential impact for the development of new research directions

- Collect additional examples (data sets) of temporal networks. We intend to develop a new network data description format netJSON based on JSON that will be used to describe all temporal networks in our collection.
- Extend the approach based on temporal quantities to other concepts of network analysis such as centrality measures based on eigen-vectors, cores, islands etc.
- Extend the algebraic approach to problems with journeys (temporal walks with non-zero transition times) by constructing appropriate semirings.
- Develop efficient algorithms for analysis of large temporal networks (based on temporal quantities) taking into account that most of large networks are sparse.
- Evaluate the feasibility to use libraries for parallel processing, such as Solomonik and Hoefler

2015, to speed-up critical computations.

- Develop a methodology for analysis of temporal networks for selected fields of application such as scientometrics, genealogies, e-mail communications, phone-communications, etc.
- A serious problem when dealing with temporal networks is the large size of data and also of obtained results. The visualization of temporal quantities proved to be very helpful. We will try to develop alternative approaches for user-friendly browsing the data and results. We will develop a javascript library netD3 of network visualization procedures based on the library D3.js.

4. The originality of the expected results

We expect to produce several original theoretical results (construction of new semirings) and new and original efficient algorithms for analysis of large temporal networks. We shall offer to the network analysis community a format netJSON as a general “language” for describing networks.

5. Working methods

Although we already know the state of the art in the field of temporal networks, we will need to make a deeper study of some topics. We will organize a special seminar to exchange the among project members knowledge and trace the project activities. Some tasks are relatively routine (collecting and encoding the data), but require a lot of work. Other tasks require theoretical solutions first, followed by implementation of the corresponding program, testing it and its application on the real-life data. Obtained results will be presented on seminars, scientific meetings and published as papers and (free) software libraries.

6. Relevance and potential impact of the results

The project results will provide new tools and methodologies for analysis of temporal networks. The format netJSON will provide a general “language” for describing networks and could be used as an option for transferring networks between different network analysis programs. The library netD3 will be based on netJSON and will provide interactive, web-based visualization of networks. The authors of network analysis programs can visualize obtained results by simply exporting them in netJSON and submit the netJSON file to a selected netD3 visualization procedure. We shall provide our collection of temporal networks on the web.

7. Exceptional socio-economic or culturally relevant achievements of the project leader

The project leader started in 1996 with Andrej Mrvar to develop Pajek – a program for analysis and visualization of large networks. Pajek is still now the only general purpose program for analysis of networks with millions of nodes. In 2013 Pajek received the INSNA's William D. Richards Jr., Software Award.

He was involved for many years in introducing of IT in Slovenian education. Important is also his contribution in introducing several software tools in Slovenia: pascal, logo, TeX and LaTeX, Postscript, Python. In 2011 he received for his contributions The Slovenian Information Society Award.

In 1991 he started with Anuška Ferligoj and Pat Doreian a new approach to blockmodeling of social networks. The results of more than ten years of research were collected in a book “Generalized blockmodeling” (2004). For this book the authors were awarded in 2007 by INSNA Harrison White Outstanding Book Award.

For more than fifteen years he is giving workshops and short courses on network analysis at different universities and other institutions all around the world. He was teaching for 10 years a course on network analysis at the ECPR summer schools in Ljubljana. He is teaching a network analysis course also at the doctoral study of statistics at University of Ljubljana.

8. Organisational structure and feasibility of the project

For all of tasks from section 20.3 we already have ideas how to proceed. The most "foggy" task is extending the algebraic approach to problems with journeys, but also in this case we expect some interesting results.

As already mentioned, we shall organize the work on the project around a special seminar that will serve for exchange of knowledge and tracing the project activities. We shall establish a project wiki that will provide a single place for project documentation and collaboration. We shall also provide a web site with temporal network data and software.

In initial phase (first year) of the project we shall develop the netJSON format and convert our collection of temporal networks in it. At the same time we can start a theoretical work on new semirings and algorithms.

When a new algorithm is developed it will be implemented and tested (month 7 – end of the project). We will start intensive development of netD3 in the second year.

We expect that after 1 year and half the version of TQ library for large networks will be functional. We will start the development of methodologies for analysis of temporal networks from selected fields.

Project results will be disseminated on seminars, workshops and conferences, on the project web site and published in scientific journals.

If available, we will engage in project activities also some doctoral students.

9. References

Allen JF 1983 Maintaining knowledge about temporal intervals. Communications of the ACM 26(11), 832–843.

Batagelj V and Zaveršnik M 2011 Fast algorithms for determining (generalized) core groups in social networks. Advances in Data Analysis and Classification 5(2), 129–145.

Batagelj V and Praprotnik S 2016 An algebraic approach to temporal network analysis based on temporal quantities. https://arxiv.org/abs/1505.01569 . In print in Social Network Analysis and Mining.

Bell MGH and Iida Y 1997 Transportation Network Analysis. Wiley, Chichester.

Casteigts A and Flocchini P 2013 Deterministic Algorithms in Dynamic Networks: Formal Models and Metrics. Technical Report DRDC Ottawa CR 2013-020, Defence R&D Canada.

Casteigts A, Flocchini P, Quattrociocchi W and Santoro N 2012 Time-varying graphs and dynamic networks. International Journal of Parallel, Emergent and Distributed Systems 27(5), 387–408.

Correa JR and Stier-Moses NE 2011 Wardrop equilibria In Wiley Encyclopedia of Operations Research and Management Science (ed. Cochran JJ, Cox LAJ, Keskinocak P, Kharoufeh JP and Smith JC) John Wiley & Sons, Inc.

D3.js - Data-Driven Documents. https://d3js.org/

Dechter R 2003 Constraint Processing. Morgan Kaufmann, San Francisco.

George B, Kim S 2013 Spatio-temporal Networks; Modeling and Algorithms. Springer Briefs in Computer Science. Springer.

Holme, P. 2015 Modern temporal network theory: a colloquium. Eur. Phys. J. B 88, 234.

Holme P and Saramäki J 2012 Temporal networks. Physics Reports 519(3), 97–125.

Holme P and Saramäki J 2013 Temporal networks. Understanding Complex Systems.Springer.

Holme P, Kim BJ, Yoon CN and Han SK 2002 Attack vulnerability of complex networks. Physical Review E 65(5), 056109.

JSON. http://www.json.org/

Kempe D, Kleinberg J and Kumar A 2000 Connectivity and inference problems for temporal networks Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing, pp. 504–513. ACM, New York.

Kolaczyk ED 2009 Statistical Analysis of Network Data: Methods and Models. Springer, New York.

Moder JJ and Phillips CR 1970 Project Management with CPM and PERT second edn. Van Nostrand Reinhold Company, New York.

Nicosia V, Tang J, Mascolo C, Musolesi M, Russo G and Latora V 2013 Graph metrics for temporal networks In Temporal Networks (ed. Holme P and Saramäki J) Understanding Complex Systems Springer Heidelberg, New York, Dordrecht, London pp. 15–40.

Ripley RM, Snijders TAB, Boda Z, Vörös A and Preciado P 2013 Manual for SIENA version 4.0 (version November 1, 2013) University of Oxford, Department of Statistics; Nuffield College Oxford.

Santoro N, Quattrociocchi W, Flocchini P, Casteigts A 2011 Time-varying graphs and social network analysis: Temporal indicators and metrics. 3rd AISB Social Networks and Multiagent Systems Symposium (SNAMAS, pages 32–38.

Solomonik E, Hoefler T 2015 Sparse Tensor Algebra as a Parallel Programming Model. http://arxiv.org/abs/1512.00066

Vilain M, Kautz H and van Beek P 1990 Constraint propagation algorithms for temporal reasoning: A revised report In Readings in Qualitative Reasoning About Physical Systems (ed.Weld DS and de Kleer J) Morgan Kaufmann Publishers Inc. San Francisco, CA, USA pp. 373–381.

Xuan BB, Ferreira A, Jarry A 2003 Computing shortest, fastest, and foremost journeys in dynamic networks. International Journal of Foundations of Computer Science, 14(2), 267–285.