

Pajek – Analysis and Visualization of Large Networks

Vladimir Batagelj¹ and Andrej Mrvar²

¹ Department of mathematics, FMF, University of Ljubljana

² Faculty of Social Sciences, University of Ljubljana

1 Short Description



Pajek (spider, in Slovene) is a program package, for Windows (32 bit), for analysis and visualization of *large networks* (having thousands of vertices). It is freely available, for noncommercial use, at its home page:
<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

We started the development of Pajek in November 1996. The main goals in the design of Pajek are [1]:

- to support abstraction by (recursive) factorization of a large network into several smaller networks that can be treated further using more sophisticated methods;
- to provide the user with some powerful visualization tools;
- to implement a selection of efficient (subquadratic) algorithms for analysis of large networks.

Besides ordinary networks Pajek supports also 2-mode networks and temporal networks.

Pajek is essentially a collection of procedures based on 6 data types: network, partition, cluster, vector, permutation, and hierarchy. These procedures are available through the main window menus. Frequently used sequences of operations can be defined as macros. This allows also the adaptations of Pajek to special groups of users.

We developed efficient algorithms for determining main parts in acyclic networks, cores, counting triads, and for pattern (subnetwork) searching [2,3,4]. Pajek contains also some data analysis procedures such as clustering and block-modeling.

Pajek is still under development. The latest version is available for download at its home page.

2 Areas of Application

Pajek was applied by researchers in different areas: social network analysis [2], chemistry (organic molecule), genealogies [5], Internet networks, citation networks, diffusion networks (AIDS, news), data-mining (2-mode networks), ...

It is also used at universities (Ljubljana, Rotterdam, Irvine, The Ohio State University, Penn State, Madrid. . .) as a support in courses on network analysis.

Together with Wouter de Nooy from University of Rotterdam we wrote a course book *Exploratory Social Network Analysis With Pajek*.

3 Layout Algorithms and Layout Features

Special emphasis is given to automatic generation of network layouts. Several standard algorithms for automatic graph drawing are implemented: spring embedders (Kamada-Kawai and Fruchterman-Reingold), layouts determined by eigenvectors (Lanczos algorithm), drawing in layers (genealogies and other acyclic structures), fish-eye views and block (matrix) representation.

These algorithms were modified and extended to enable additional options: drawing with constraints (optimization of the selected part of the network, fixing some vertices to predefined positions, using values of lines as similarities or dissimilarities), drawing in 3D space. Pajek also provides tools for manual editing of graph layout.

Properties of vertices/lines (given as data or computed) can be represented using colors, sizes and/or shapes of vertices/lines.

4 Architecture

Pajek is implemented in Delphi and runs on Windows 32 operating systems.

5 Interfaces

Pajek supports also some non-native input formats: UCINET DL files; chemical MDLMOL and BS; and genealogical GEDCOM.

The layouts can be exported in the following output graphic formats that can be examined by special 2D and 3D viewers: Encapsulated PostScript, Scalable Vector Graphics, VRML, MDLMOL, and Kinemages.

References

1. Batagelj, V., Mrvar, A. (1998): Pajek – A Program for Large Network Analysis. *Connections*, **21** (2), 47-57.
2. Batagelj, V., Mrvar, A. (2000): Some Analyses of Erdős Collaboration Graph. *Social Networks*, **22**, 173-186.
3. Batagelj, V., Mrvar, A. (2001): A Subquadratic Triad Census Algorithm for Large Sparse Networks with Small Maximum Degree. *Social Networks*, **23**, 237-243.
4. Batagelj, V., Mrvar, A., Zaveršnik M. (1999): Partitioning Approach to Visualization of Large Graphs. Kratochvíl, J. (Ed.) Graph Drawing. 7th International Symposium, GD'99, Štířín Castle, Czech Republic, Proceedings. *Lecture Notes in Computer Science*, 1731. Springer-Verlag, Berlin/Heidelberg, 90-97.
5. White, D.R., Batagelj, V., Mrvar, A. (1999): Analyzing Large Kinship and Marriage Networks with Pgraph and Pajek. *Social Science Computer Review*, **17** (3), 245-274.