International laboratory for Applied Network Research

MASNA: Applied Statistics with Network Analysis

Moscow 2023, June 29

Pajek program for analysis and visualization of complex networks

Workshop of the MASNA online programme

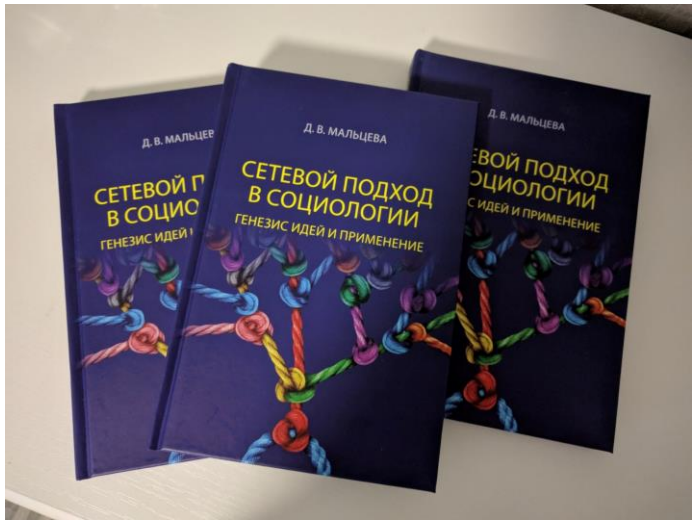
Daria Maltseva, Ph.D.

About me

- Head of the International Laboratory for Applied Network Research
- Leading Research Fellow, Visiting Lecturer
- Ph.D. | Candidate of Sciences in Sociology
- Degree in Sociology, Russian State University for the Humanities
- Young Faculty Support Program (Group of Young Academic Professionals), Category "New Researchers" (2018-2019)
- Grant of the Russian Scientific Foundation “Collaboration patterns in the Russian sociological community: the structure of scientific schools and their growth potential” 2021-2023
- Have a 2.5-year-old son

E-mail: dmaltseva@hse.ru

Expertise



Сетевой подход как феномен социологической теории

версия для печати

Мальцева Д. В.
Кандидат социологических наук, зам. заведующего, Международная лаборатория прикладного сетевого анализа НИУ «Высшая школа экономики», Москва, Россия d_malceva@mail.ru

DOI: 10.7868/S0132162518040013
ID статьи на сайте журнала: 7135

Рубрика: Методология и методы социологических исследований

Ссылка при цитировании:

Мальцева Д. В. **Сетевой подход как феномен социологической теории // Социологические исследования. 2018. № 4. С. 3-14.**
DOI: 10.7868/S0132162518040013

Текст статьи.

Аннотация

Представлен сравнительный анализ трех направлений сетевого подхода в социологии – анализа социальных сетей, реляционной социологии и акторно-сетевой теории. Основаниями для анализа выступают время и контекст появления, основные теоретические положения, методология и методы эмпирических исследований. Делаются выводы о сходстве (до определенной степени) направлений реляционной социологии и акторно-сетевой теории и их отличия от анализа социальных сетей. Каждое из направлений автономно и занимает определенное место в структуре социологического знания. Делается вывод о некорректности представления «сетевого подхода» как единого теоретического блока, объединения его направлений под унифицирующими названиями «сетевая теория», «сетевая парадигма» и др. Подчеркнуто, что это собирательное название для разных теорий и подходов, оперирующих понятием сети в разных смыслах.

Springer Link

Published: 19 April 2022

Collaboration between authors in the field of social network analysis

Daria Maltseva & Vladimir Batagelj

Scientometrics (2022) | Cite this article
267 Accesses | **2** Altmetric | Metrics

Abstract

This paper presents a study of authors writing articles in the field of SNA and groups the
means of bibliographic network analysis. The dataset consists of works from the Web of
Science database obtained by searching for "social network*", works highly cited in the f
works published in the flagship SNA journals, and written by the most prolific authors
(70,000+ publications and 93,000+ authors), up to and including 2018. Using a two-mo
network linking publications with authors, we constructed and analysed different types

Springer Link

Published: 30 August 2019

Social network analysis as a field of invasions: bibliographic approach to study SNA development

Daria Maltseva & Vladimir Batagelj

Scientometrics **121**, 1085–1128 (2019) | Cite this article
994 Accesses | **10** Citations | **1** Altmetric | Metrics

Abstract

In this paper, the results of a study on the development of social network analysis (SNA) and its evolution over time, using the analysis of bibliographic networks are presented. The dataset consists of articles from the Web of Science Clarivate Analytics database obtained by

Springer Link

Published: 25 January 2020

Towards a systematic description of the field using keywords analysis: main topics in social networks

Daria Maltseva & Vladimir Batagelj

Scientometrics **123**, 357–382 (2020) | Cite this article
806 Accesses | **8** Citations | **1** Altmetric | Metrics

Abstract

This paper presents the results of the analysis of keywords used in Social Network Analysis (SNA) articles included in the WoS database and main SNA journals, from 1970 to 2018.

Springer Link

Published: 25 February 2021

Journals publishing social network analysis

Daria Maltseva & Vladimir Batagelj

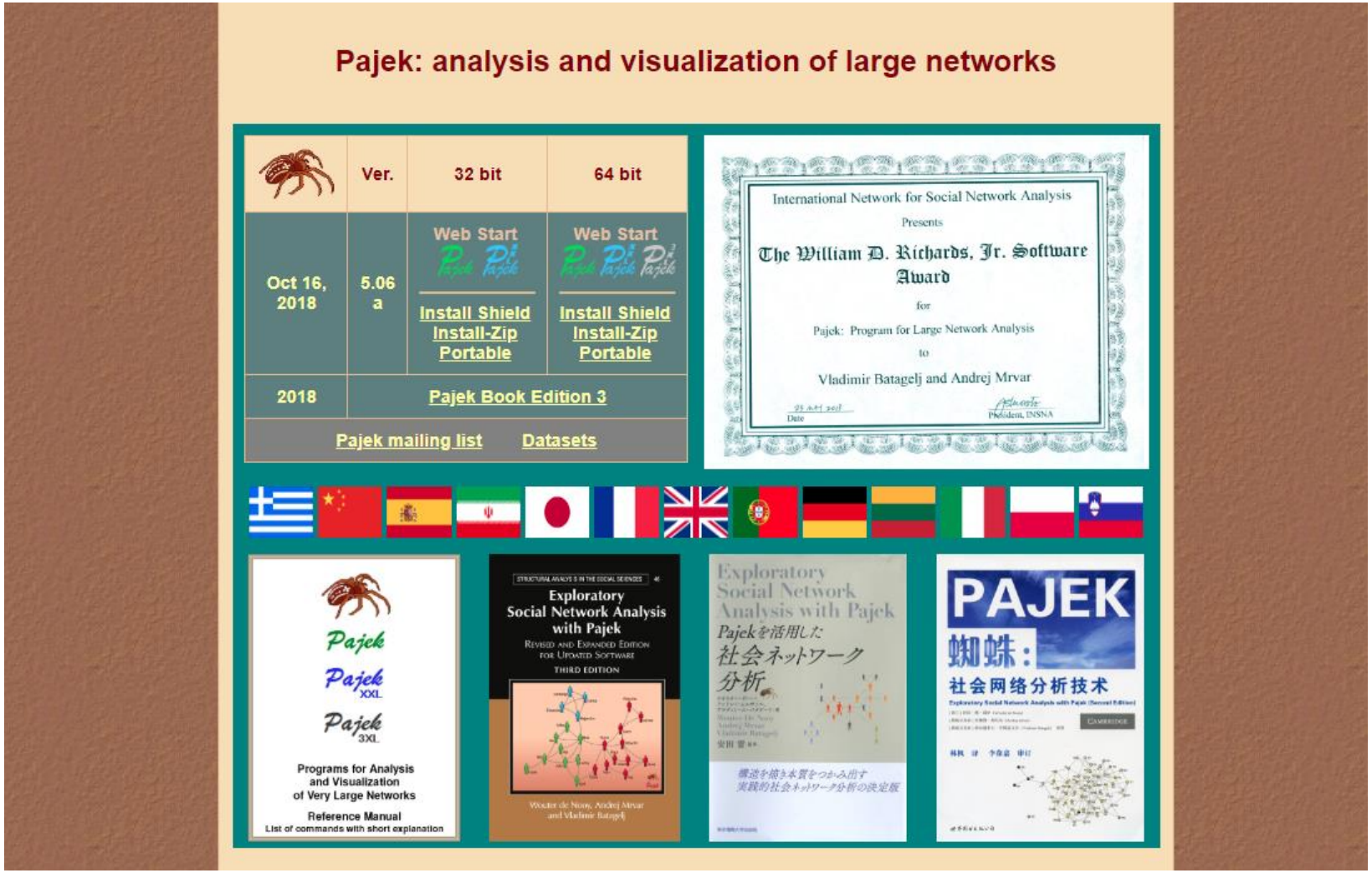
Scientometrics **126**, 3593–3620 (2021) | Cite this article
585 Accesses | **3** Citations | **1** Altmetric | Metrics

Abstract

This paper presents the analysis of journals publishing articles on social network analysis (SNA). The dataset consists of articles from the Web of Science database obtained by searching for "social network*", works intensively cited, written by the most prominent

Expertise

Series of internships in Center for Methodology and Social Informatics (Faculty of Social Sciences, University of Ljubljana) – 2017, 2018(2), 2019



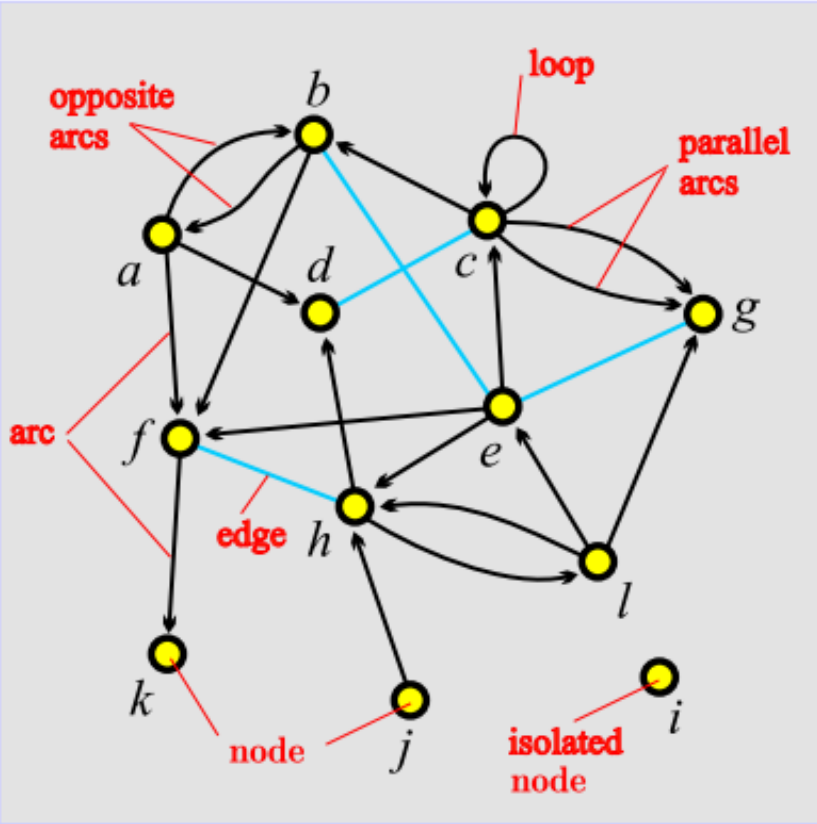
Aim of the seminar

To perform basic operations in Pajek in practice:

- generate a set of network data and modify the resulting networks,
- calculate the main metrics for individual nodes and the whole network,
- visualize the resulting networks.

To practice working in Pajek with the real empirical data collected by the ANR-Lab members.

Network Conceptualization



The **network** is based on two sets - a set of **nodes** (vertices) representing the selected units of analysis and a set of **lines** (links) representing the connections between the units of analysis, which together form a **graph**.

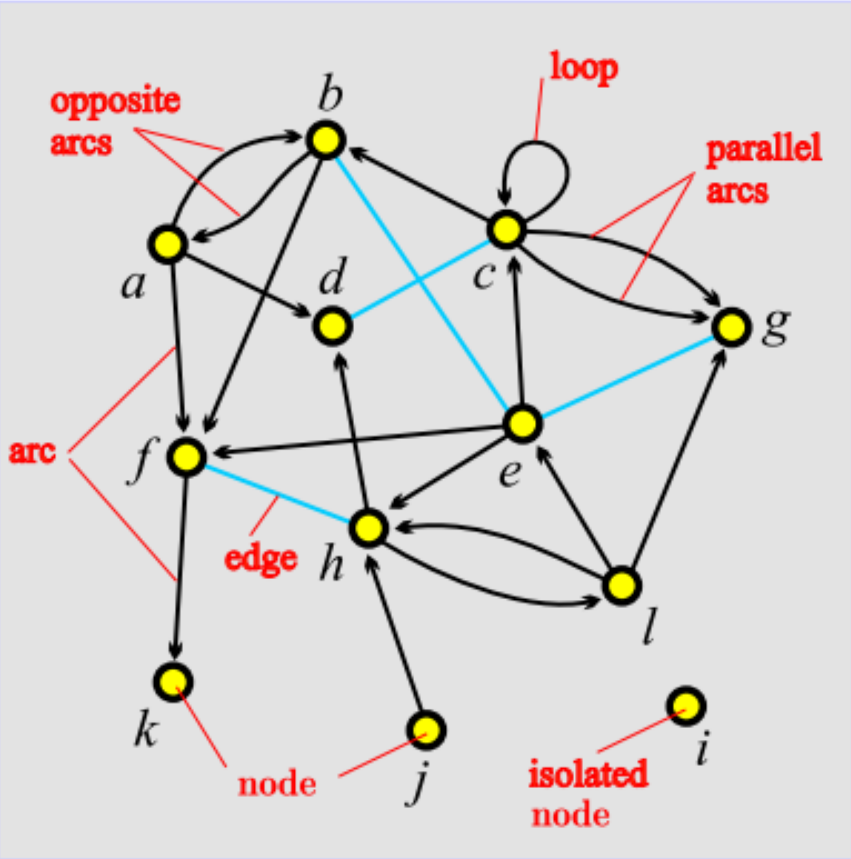
The line can be directed (**arc**) or undirected (**edge**).

Nodes and lines can contain additional data - characteristics / **attributes** (name, type, value) - that can be measured or calculated.

Network = Graph + Data

Visualization by: Vladimir Batagelj

Network data format: set of nodes and links

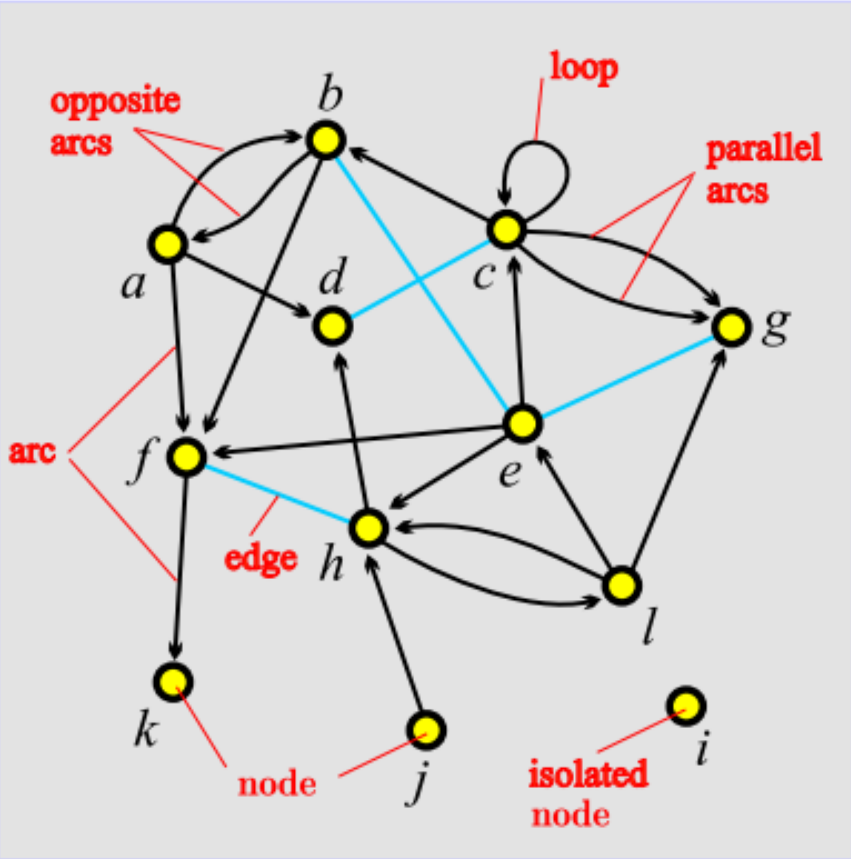


$$
\begin{aligned}
\mathcal{V} &= \{a, b, c, d, e, f, g, h, i, j, k, l\} \\
\mathcal{A} &= \{(a, b), (a, d), (a, f), (b, a), \\
&\quad (b, f), (c, b), (c, c), (c, g)_1, \\
&\quad (c, g)_2, (e, c), (e, f), (e, h), \\
&\quad (f, k), (h, d), (h, l), (j, h), \\
&\quad (l, e), (l, g), (l, h)\} \\
\mathcal{E} &= \{(b\colon e), (c\colon d), (e\colon g), (f\colon h)\} \\
\mathcal{G} &= (\mathcal{V}, \mathcal{A}, \mathcal{E}) \\
\mathcal{L} &= \mathcal{A} \cup \mathcal{E}
\end{aligned}
$$

```
*Vertices  12
  1 "a"     0.1020  0.3226
  2 "b"     0.2860  0.0876
  3 "c"     0.5322  0.2304
  4 "d"     0.3259  0.3917
  5 "e"     0.5543  0.4770
  6 "f"     0.1552  0.6406
  7 "g"     0.8293  0.3249
  8 "h"     0.4479  0.6866
  9 "i"     0.8204  0.8203
 10 "j"     0.4789  0.9055
 11 "k"     0.1175  0.9032
 12 "l"     0.7095  0.6475
*Arcs
  1  2
  2  1
  1  4
  1  6
  2  6
  3  2
  3  3
  3  7
  3  7
  5  3
  5  6
  5  8
  6 11
  8  4
 10  8
 12  5
 12  7
  8 12
 12  8
*Edges
  2  5
  3  4
  5  7
  6  8
```

Link weight can be added (third column)

Network data format: links of nodes with “neighbors”

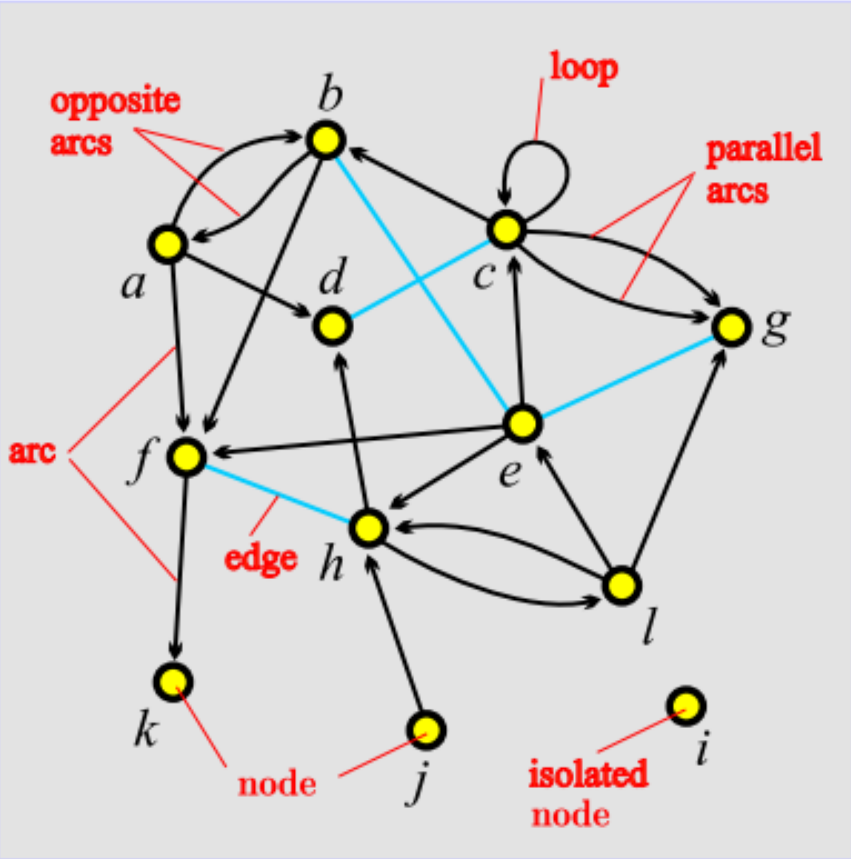


$$N_A(a) = \{b, d, f\}$$
$$N_A(b) = \{a, f\}$$
$$N_A(c) = \{b, c, g, g\}$$
$$N_A(e) = \{c, f, h\}$$
$$N_A(f) = \{k\}$$
$$N_A(h) = \{d, l\}$$
$$N_A(j) = \{h\}$$
$$N_A(l) = \{e, g, h\}$$

$$N_E(e) = \{b, g\}$$
$$N_E(c) = \{d\}$$
$$N_E(f) = \{h\}$$

```
*Vertices  12
  1 "a"    0.1020   0.3226
  2 "b"    0.2860   0.0876
  3 "c"    0.5322   0.2304
  4 "d"    0.3259   0.3917
  5 "e"    0.5543   0.4770
  6 "f"    0.1552   0.6406
  7 "g"    0.8293   0.3249
  8 "h"    0.4479   0.6866
  9 "i"    0.8204   0.8203
 10 "j"    0.4789   0.9055
 11 "k"    0.1175   0.9032
 12 "l"    0.7095   0.6475
*Arcslist
  1  2  4  6
  2  1  6
  3  2  3  7  7
  5  3  6  8
  6 11
  8  4 12
 10  8
 12  5  7  8
*Edgeslist
  2  5
  3  4
  5  7
  6  8
```

Network data format: adjacency matrix



	a	b	c	d	e	f	g	h	i	j	k	l
a	0	1	0	1	0	1	0	0	0	0	0	0
b	1	0	0	0	1	1	0	0	0	0	0	0
c	0	1	1	1	0	0	2	0	0	0	0	0
d	0	0	1	0	0	0	0	0	0	0	0	0
e	0	1	1	0	0	1	1	1	0	0	0	0
f	0	0	0	0	0	0	0	1	0	0	1	0
g	0	0	0	0	1	0	0	0	0	0	0	0
h	0	0	0	1	0	1	0	0	0	0	0	1
i	0	0	0	0	0	0	0	0	0	0	0	0
j	0	0	0	0	0	0	0	1	0	0	0	0
k	0	0	0	0	0	0	0	0	0	0	0	0
l	0	0	0	0	1	0	1	1	0	0	0	0

Difficulty with distinguishing between displaying two reciprocal links (a, b) and non-directional links (b, e)

```
*Vertices  12
  1 "a"     0.1020   0.3226
  2 "b"     0.2860   0.0876
  3 "c"     0.5322   0.2304
  4 "d"     0.3259   0.3917
  5 "e"     0.5543   0.4770
  6 "f"     0.1552   0.6406
  7 "g"     0.8293   0.3249
  8 "h"     0.4479   0.6866
  9 "i"     0.8204   0.8203
 10 "j"     0.4789   0.9055
 11 "k"     0.1175   0.9032
 12 "l"     0.7095   0.6475
*Matrix
 0 1 0 1 0 1 0 0 0 0 0 0
 1 0 0 0 1 1 0 0 0 0 0 0
 0 1 1 1 0 0 2 0 0 0 0 0
 0 0 1 0 0 0 0 0 0 0 0 0
 0 1 1 0 0 1 1 1 0 0 0 0
 0 0 0 0 0 0 0 1 0 0 1 0
 0 0 0 0 1 0 0 0 0 0 0 0
 0 0 0 1 0 1 0 0 0 0 0 1
 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 1 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0
 0 0 0 0 1 0 1 1 0 0 0 0
```

Additional information

- **clu** – clustering: partition of nodes – nominal or ordinal data about nodes (node belongs to the cluster/group)
- **vec** – vector: numeric data about nodes (the property has value on node);
- **per** – permutation: ordering of nodes (node is at the certain position)

When collecting the network data consider to provide as much properties as possible.

Wolfe Monkey Data

inter.net

```
*Vertices  20
  1 "m01"
  2 "m02"
  3 "m03"
  4 "m04"
  5 "m05"
  6 "f06"
  7 "f07"
  8 "f08"
  9 "f09"
 10 "f10"
 11 "f11"
 12 "f12"
 13 "f13"
 14 "f14"
 15 "f15"
 16 "f16"
 17 "f17"
 18 "f18"
 19 "f19"
 20 "f20"
*Edges
  1   2   2
  1   3  10
  1   4   4
  -   -   -
```

inter.net

```
  1   6   5
  1   7   9
  1   8   7
  1   9   4
  1  10   3
  1  11   3
  1  12   7
  1  13   3
  1  14   2
  1  15   5
  1  16   1
  1  17   4
  1  18   1
  2   3   5
  2   4   1
  2   5   3
  2   6   1
  2   7   4
  2   8   2
  2   9   6
  2  10   2
  2  11   5
  2  12   4
  2  13   3
  2  14   2

  ...
```

sex.clu

```
*vertices 20
 1
 1
 1
 1
 1
 2
 2
 2
 2
 2
 2
 2
 2
 2
 2
 2
 2
 2
 2
```

age.vec

```
*vertices 20
 15
 10
 10
  8
  7
 15
  5
 11
  8
  9
 16
 10
 14
  5
  7
 11
  7
  5
 15
  4
```

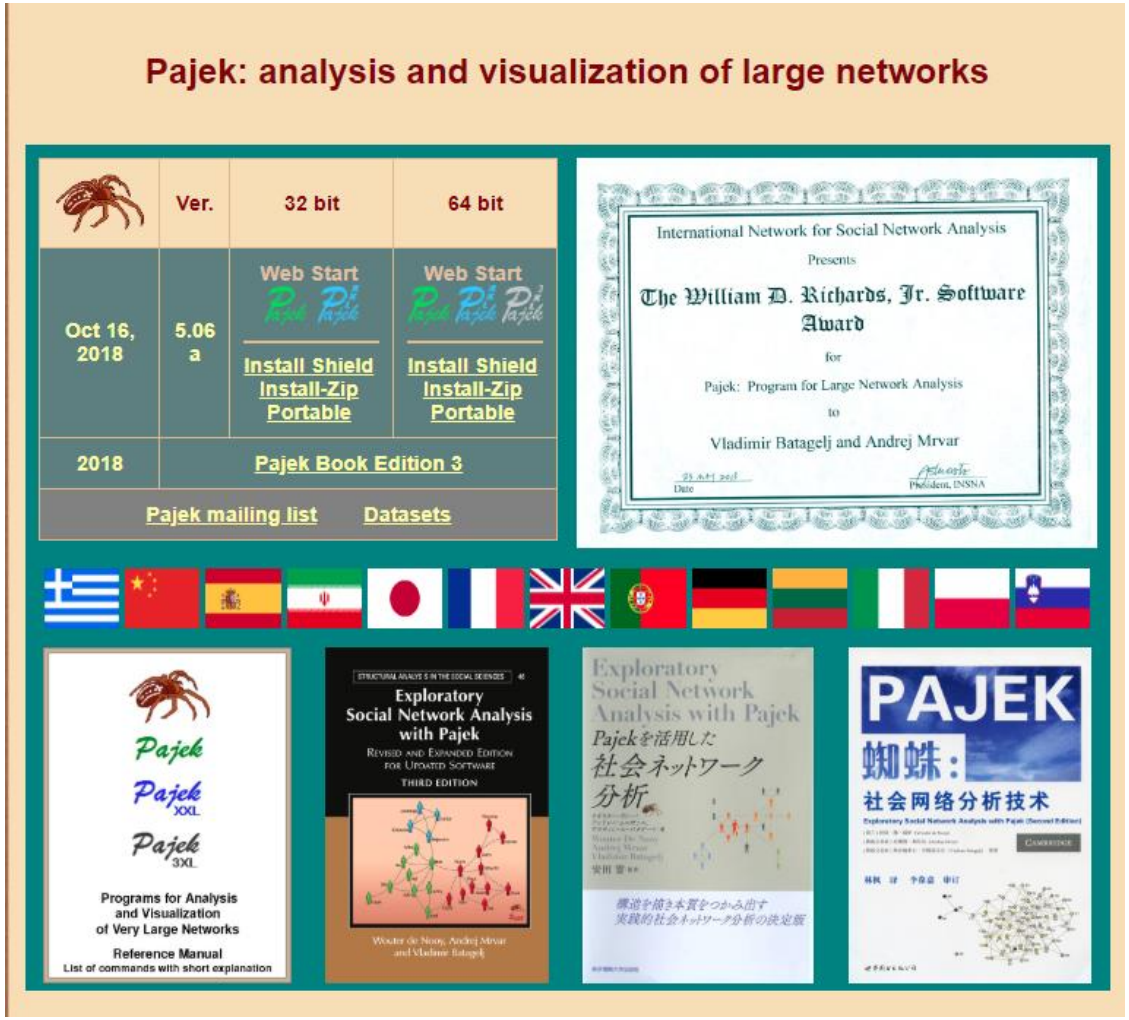
rank.per

```
*vertices 20
  1
  2
  3
  4
  5
 10
 11
  6
 12
  9
  7
  8
 18
 19
 20
 13
 14
 15
 16
 17
```

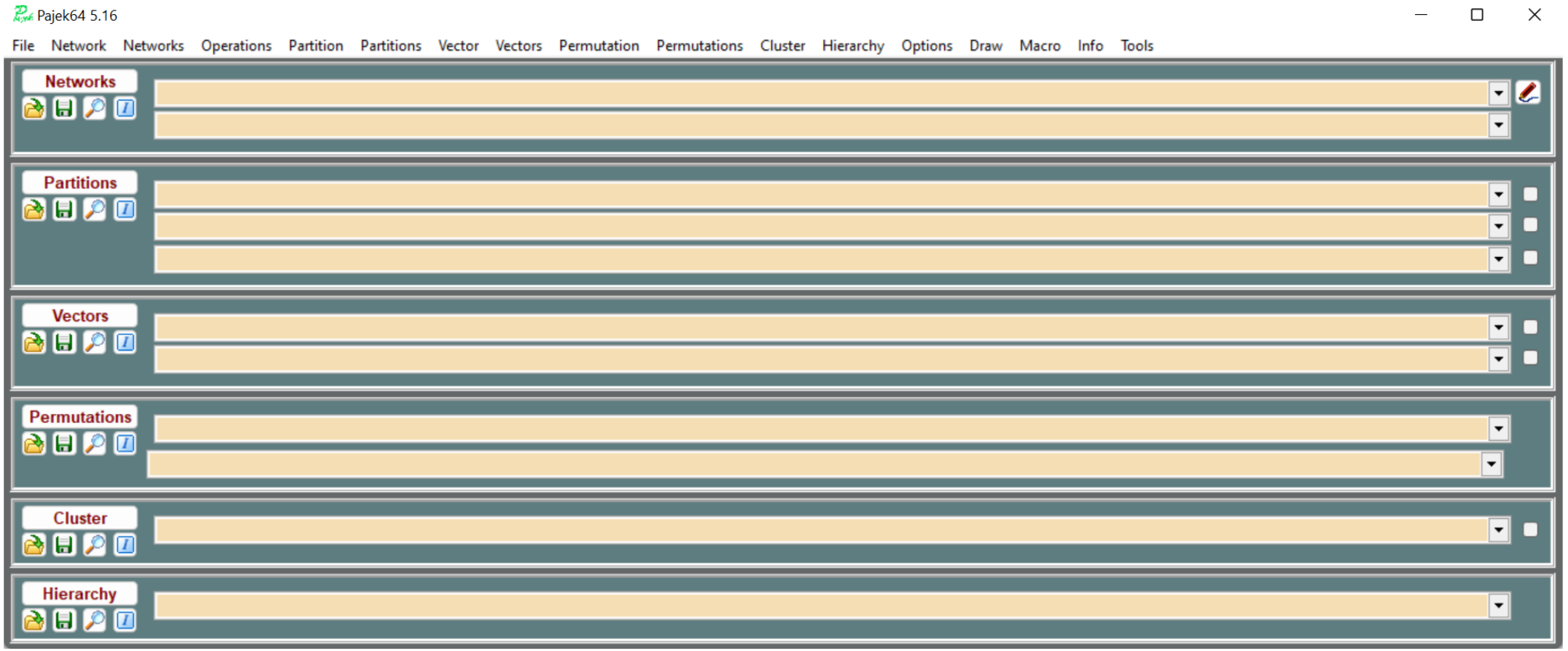
0 is not allowed as node number

Pajek

- Pajek is a program, for Windows, for analysis and visualization of large networks having some thousands or even millions of vertices. The latest version of Pajek is freely available, for noncommercial use.
- Link: http://mrvar.fdv.uni-lj.si/pajek/
- Creators: Andrej Mrvar and Vladimir Batagelj



Pajek



Link to the workshop materials:
https://github.com/Daria-Maltseva/pajek/wiki/workshop_masna2023

announcement

14th Summer School
'Methods and tools of social network analysis'

Date

July 4-8

Save the date: July 4, 5.45 – 7.15 PM (Moscow)
Daria Maltseva, Analysis and visualization of complex networks in Pajek