



# A Comparative Review on Different Social Network Analytical Tools

Jacob P. Cherian<sup>(✉)</sup>, Jubilant J. Kizhakkethottam, and Aneena Ann Alexander

Department of Computer Science and Engineering, Saintgits College of Engineering, Kottayam, Kerala, India

cherianjacobp@gmail.com, jubilantjob@gmail.com,  
aneenaalex@gmail.com

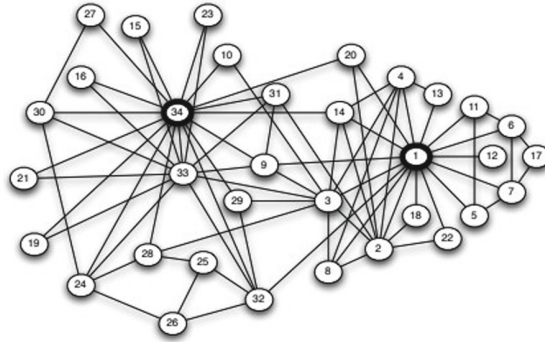
**Abstract.** With the evolution of Web 2.0, Social Networking sites like Facebook, Twitter, Instagram and similar platforms have gained wide popularity. They provide an in-depth of knowledge about users and the relationships between them. In order to analyse and extract meaningful information from these vast social network data, special graphical mining tools are needed that can efficiently model the characteristics of social networks. A variety of tools are available for Social Network Analysis, where raw network information can be formatted in an edge list, adjacency list, or adjacency matrix, often coupled with attribute information. We compare major tools and packages for analysing and visualising social networks that have a broad variety of applications covering genetics, economics, sociology, network theory, and several other domains. This work provides a comparison based on platform, license, file formats supported, layout and visualization and available metrics.

**Keywords:** Social networks · Tools · Metrics

## 1 Introduction

Social Network Analysis [1] is a methodology of quantitatively and qualitatively measuring and analyzing a social network. Social Network Analysis (SNA) is the method of analyzing social systems by use of networks and graph theory. It describes networked structures in terms of the nodes and the ties, edges, or links (interactions or relationship) that exist among them. SNA tools helps in identifying various relationships and changes in relationship between different entities. Entities or nodes may include computers, web pages, humans, group of authors collaborating on an article, organizations etc.

A famous example of connections among various nodes were described by Wayne W. Zachary who researched a social network of a karate club for a three-year period from 1970 to 1972. The network captures 34 karate club members, documenting connections between pairs of participants who interacted outside the club [2]. Our technological and economic systems have also become highly complex networks. This has made it harder to reason about their conduct and progressively dangerous to tinker with it. It made them vulnerable to disruptions spreading through the fundamental buildings of the network, sometimes turning localized breakdowns into cascading mistakes or financial crises (Fig. 1).



**Fig. 1.** The social network of friendships within a 34- person karate club (Zachary's Karate Club Dataset) [2]

Studying the various characteristics and perspective of a network has become a key research topic among social scientists. With the evolution of various tools for social network analysis, enormous amount of information has been gathered which influence the structure of the network thereby throwing light on marketing prospects, financial leaps, social exploration and so on. An interesting question is to why we should ever perform social media analytics? The answer lies in understanding the following aspects:

1. **Helps to understand the audience:** Knowing the social crowd can help maximize social media yields and help better conversions.
2. **Social data helps to create better content:** Tracking of a social network can help in identifying the content which drives traffic on a website.
3. **Help to understand competitors:** Your rivals also create content and execute policies on social media. This will result in their own unique information. You will be able to figure out what works and what isn't if you analyze this data.
4. **Social metrics can help you create a better strategy:** If you frequently study your social media analytics, you can find out what are the mistakes which resist your growth.
5. **Helps to identify the most influential node:** Many social network tools allow you to identify the key person in a network or the most influential node in a social graph.

All the above mentioned prospects of a social network can be well analyzed and studied by efficient tools. A wide variety of tools exist both to study the network structure of a social graph and understanding the structural and spatial characteristics of social data. In this paper, we focus on tools used to visualize and analyze the network structure of a social graph. Various tools like Gephi [3], Pajek [4], UCINET, Cytoscape [5], NetworkX [6] package, igraph package Graphviz [7], NodeXL [8] are to name a few.

## 2 Tools for Social Network Analysis

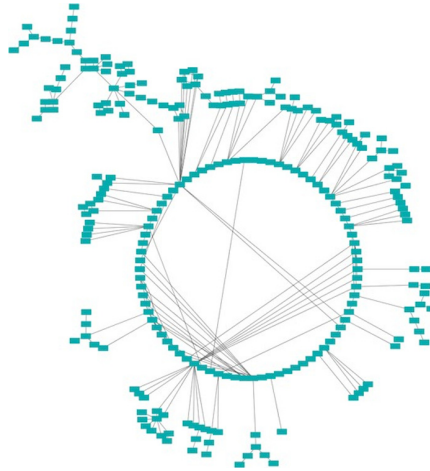
The social network analysis tool enables quantitative or qualitative study of social networks by defining the attributes of the social network, either through graphical or numerical representation [9]. A network analysis tool explores interactions and links within a dataset. Through network analysis, the networks are generated not by particular content types, but by relations among different content components. Typically, a network is made up of nodes and ties among these nodes. A variety of tools help in visualization and analysis of social networks thereby providing help to researchers in understanding how nodes are connected together in a network. Such analysis tools also aid in the creation of research questions and ultimately allow the researcher to reach conclusions. We have classified network analysis tools into two categories: Focused desktop tools and Developer tools. Focused desktop tools are standalone software tools which are primarily focused on network analysis and visualization. Developer tools are libraries or packages for network analysis which can be integrated along with another program.

### 2.1 Focused Desktop Tools

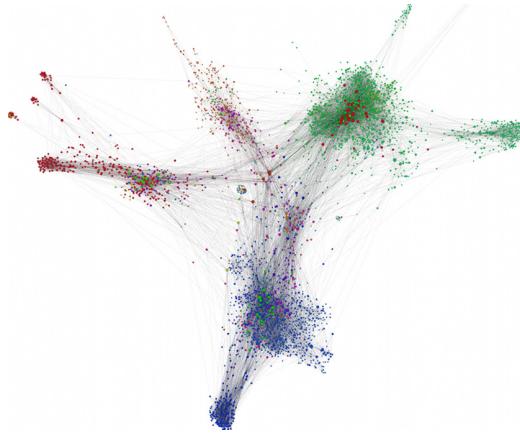
**Cytoscape.** Cytoscape is an open source software platform for the visualization and integration of molecular interaction networks and biological pathways with gene expression profiles, annotations and other state information. Although initially intended for biological studies, Cytoscape is now a general platform for complex analysis and visualization of networks [10]. Cytoscape was originally built by the Institute of Systems Biology in Seattle in 2002. Now, it is being developed by an internationally acclaimed consortium of free and open source software developers. The networks can be exported as ready to publish images with high quality. File formats like PDF, PS, SVG, PNG, JPEG, and BMP files are supported.

**Gephi.** Gephi is a free and open source tool primarily developed for Windows, Mac OS and Linux platforms. Gephi offers real time visualization with features for cartography, dynamic filtering and facilitation of logical reasoning. Gephi was initially developed by a group of students of the University of Technology, Compaigne in France. Gephi has been widely used in a number of research projects, journalism, academia and for examining traffic in Twitter networks. Gephi is widely used in Digital Humanities and various branches of social and political science. The purpose is to help data scientists make assumptions, intuitively discover patterns, isolate structure singularities or flaws during data sourcing. Gephi uses a 3D render engine to visualize large networks in real time and to accelerate the process of exploration [3].

Gephi is operated by its ad-hoc OpenGL engine, pushing the envelope on how customizable and effective network exploration can be. Networks that contain 100,000 nodes and 1,000,000 edges can be visualized using dynamic filtering Rich tools for constructive graph manipulation (Figs. 2 and 3).

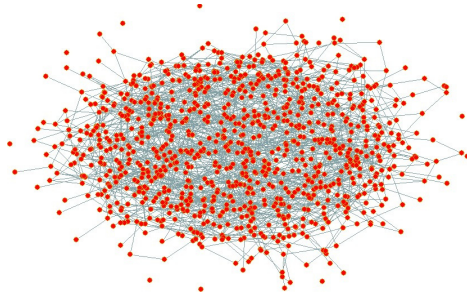


**Fig. 2.** Circular layout representation of a social network using Cytoscape



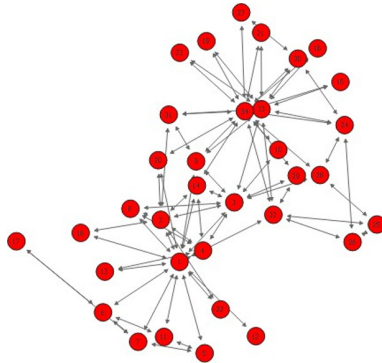
**Fig. 3.** Community detection using Gephi [11]

**Pajek.** Pajek is a program, for Windows, for visualization and analysis of large networks having thousands or millions of nodes. Pajek is useful in hierarchical data manipulation and provides powerful and accessible data manipulation functions. 3D visualization and its export in VRML are also available. 3 It can be used to analyze and visualize the large networks. It is developed by Vladimir Batagelj and Andrej Mrvar from the University of Ljubljana. Pajek is available freely, for noncommercial use. The key design aspects of Pajek was to facilitate the reduction of complex networks into smaller networks which can be further analyzed using sophisticated tools, to enable the user with powerful visualization tools, to implement a selection of efficient network algorithms [12] (Fig. 4).



**Fig. 4.** Representation of a random network in Pajek using Fruchterman-Reingold layout

**SocNetV.** Social Network Visualizer (SocNetV) is a cross-platform, customer-friendly free software application for analysis and visualization of social networks. It provides tools for drawing social networks with a few clicks on a virtual canvas, loading field data from a folder in a compatible format, or browsing the internet to create a social network of related web pages. It also permits editing of actors and ties through point-and-click, analyze graph and social network properties, generates beautiful HTML reports and embed visualization layouts to the network (Fig. 5).

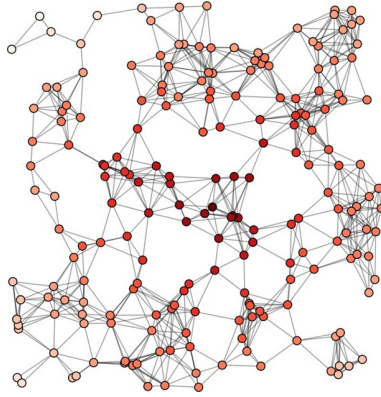


**Fig. 5.** Representation of Zachary's Karate Club using Kamada-Kawai layout model in SocNetV

## 2.2 Developer Tools

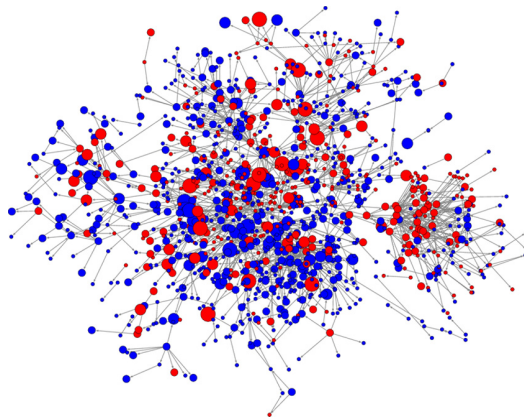
**NetworkX.** NetworkX is a Python language package for exploration and analysis of networks and network algorithms <sup>4</sup> NetworkX is appropriate for large-scale processing of real-world graphs, suitable for graphs reaching 10 million nodes and 100 million edges. NetworkX is a highly effective, exceptionally scalable, extremely mobile framework for social network analysis platform because it relies on a pure Python “dictionary” knowledge structure. With NetworkX, we can load and store networks in standard and

non-standard data formats, create many types of random and traditional networks, analyze network composition, create network models, develop new network algorithms, formulate networks, and much more (Fig. 6).



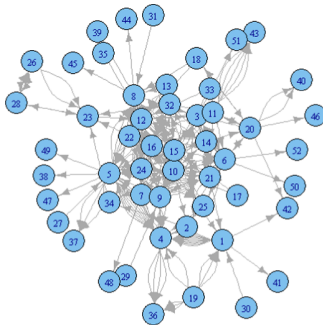
**Fig. 6.** Random geometric graph generated using networkx [13]

**SNAP.** Stanford Network Analysis Platform (SNAP) is a high performance, general purpose system for the manipulation and analysis of large networks. Graphs consists of nodes and directed, undirected or multiple edges between the nodes of the graph. Written in C++, the core SNAP library is designed for peak performance and concise graph representation. It is quickly scalable for large networks of hundreds of millions of nodes and billions of edges. Efficiently manipulates large graphs, measures structural properties, produces normal and random graphs, and supports node and edge attributes (Fig. 7).



**Fig. 7.** Representation of a full network in SNAP, where the red nodes denote news media sites and the blue nodes denote blogs. (Color figure online)

**iGraph.** Igraph is a suite of network analysis tools focused on performance, portability & user-friendliness. Igraph is a free and open-source package. Igraph can be programmed in R, Python, Mathematica and C/C++. igraph can be used to generate graphs, compute centrality measures and path length based properties as well as graph components and graph motifs. It also can be used for Degree-preserving randomization. Igraph can read and write Pajek and GraphML files, as well as simple edge lists. The library contains several layout tools as well. Igraph consists of a large collection of generators that can be separated into two groups: deterministic and stochastic graph generators. With the same variables, deterministic generators create the same graph, whereas stochastic generators create a different graph. The igraph library was developed due to the lack of network analysis software which can handle complex graphs effectively, which can be embedded into a high level programming platform and which can be used both interactively and non-interactively [14] (Fig. 8).



**Fig. 8.** Representation of a directed graph using igraph. [15]

**sigma.js.** Sigma is a JavaScript library devoted for drawing graphs. It allows easily publishable networks on Web pages, and helps developers to integrate network exploration in dynamic Web applications. Sigma offers a lot of different settings to make it simple to customize the way networks are drawn and connected. We can also directly add



**Fig. 9.** Representation of a social network using sigma.js [16]

functions to scripts for rendering nodes and edges the way we exactly want to. Custom rendering is provided in `sigma.js` with support for Canvas or WebGL (Fig. 9).

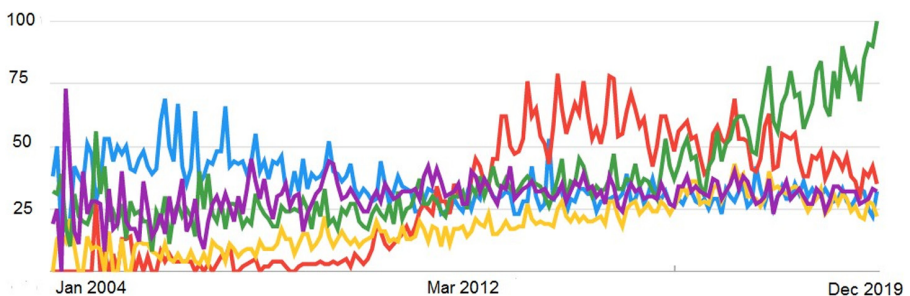
### 3 Evaluation

A wide variety of tools are available for social network analysis. In this paper, we have considered Pajek, SocNetV, Gephi, IGraph, NetworkX, Cytoscape and NodeXL for evaluation and comparison. The tools were selected purely based on the overall functionalities and ease of use provided by these tools. We have performed evaluation based on the following aspects:

- a) Popularity Trend of Search Term corresponding to the tools
- b) Metrics
- c) Layout & Visualization
- d) File Formats
- e) General Information (License, Platform, Type).

#### 3.1 Popularity Trend

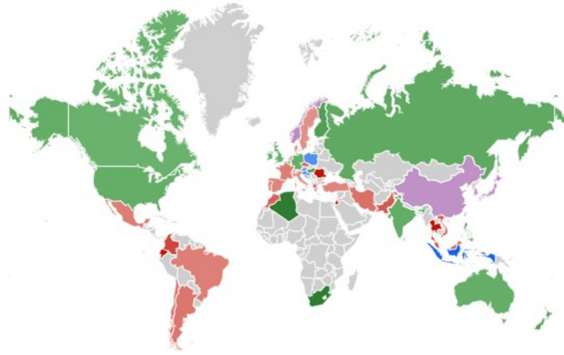
The popularity trend of search term corresponding to the tools mentioned above were performed on Google Trends. The search trend over a period spanning 15 years (2004 January to 2019 December) is as shown in Fig. 10. The search trend clearly indicates that NetworkX is the most popular network tool (since 2017) on search volume compared to the other tools, followed by Gephi. Pajek had been the favorite network analysis tool in the first decade of the 21st century. Worldwide search interest for the various tools are provided in Fig. 11. Geographic breakdown of search trend for each tool is provided form Fig. 12, 13, 14, 15 and 16. Table 1 provides information on country wise search percentages calculated out of searches for all five terms in various countries. Search trends for NodeXL and SocnetV were not considered due to their low search volume compared to the other tools.



**Fig. 10.** Search trends for various tools from January 2004 to December 2019, Google Trends



● pajek ● Gephi ● Igraph ● NetworkX ● Cytoscape



**Fig. 11.** Worldwide search interest for different social network analysis tools from January 2004 to December 2019. The color intensity represent percentage of search volume. (Google Trends)



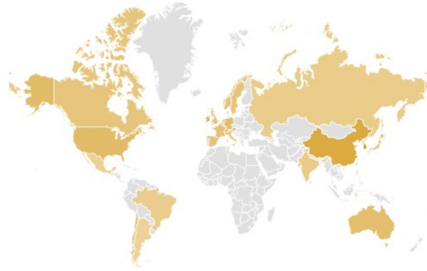
**Fig. 12.** Search interest for Pajek (Breakdown by Region), Google Trends



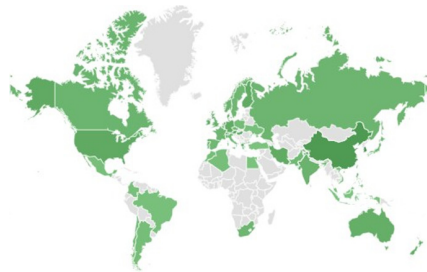
**Fig. 13.** Search interest for Gephi (Breakdown by Region), Google Trends

### 3.2 Metrics

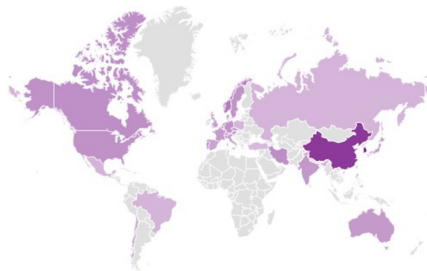
Metrics or indices play a crucial role in identifying the most important nodes within a social graph. Identifying the most important person in a social network, identifying key infrastructure nodes on the internet, spreaders of epidemics are vital in network theory. This paper provides a comparative study on the different types of metrics available on the



**Fig. 14.** Search interest for IGraph (Breakdown by Region), Google Trends



**Fig. 15.** Search interest for NetworkX (Breakdown by Region), Google Trends



**Fig. 16.** Search interest for Cytoscape (Breakdown by Region), Google Trends

various tools in the study. The paper focus on centrality measures [17] like betweenness centrality, closeness centrality, degree centrality, eigen vector centrality, power centrality, eccentricity centrality, stress centrality and other indices like triads, cliques and page rank. It is observed that SocnetV and igraph package provide support to all the metrics under consideration. Betweenness centrality and closeness centrality measures can be calculated with all the tools under the study.

### 3.3 Visualization and Layout

The biggest challenges of graph exploration are geared towards high-level visualization. There are numerous ways in which a graph can be drawn and represented visually.

**Table 1.** Search percentages calculated out of searches for all five terms in various countries.

Country	Pajek	Gephi	IGraph	NetworkX	Cytoscape
Australia	0%	30%	18%	34%	18%
Brazil	13%	38%	15%	20%	14%
Canada	7%	26%	14%	26%	27%
China	16%	19%	13%	25%	27%
Germany	8%	26%	16%	31%	19%
India	5%	23%	11%	34%	27%
Norway	0%	26%	14%	26%	34%
Russia	5%	28%	14%	41%	12%
Sweden	0%	30%	17%	27%	26%
United States of America	5%	24%	17%	34%	20%

**Table 2.** Comparison of the metrics available on various tools

Indices	SocnetV	networkx	igraph	Cytoscape	Gephi	Pajek	NodeXL
Degree centrality	C	C	C	C		C	C
Betweenness centrality	C	C	C	C	C	C	C
Closeness centrality	C	C	C	C	C	C	C
Stress centrality	C		C	C			
Power centrality	C		C				
Eccentricity centrality	C	C	C		C		
Eigen vector centrality	C	C	C	C			C
Triad census	C	C	C			C	
Page rank	C	C	C	C	C		C
Cliques	C	C	C			C	

Each visualization layout has its own merits. In Table 3, a comparison on the different layout options available on the tools are provided. It is observed that not all the layout models considered are currently available in the tools under discussion. Each tool

provides certain layout features as mentioned in the table. While most of the tools support Fruchterman Reingold layout and the circular layout, Sugiyama layout model is supported by igraph and nodexl.

**Table 3.** Comparison of the types of layout available on various tools.

Layout	SocnetV	networkx	igraph	Cytoscape	Gephi	Pajek	NodeXL
Fruchterman Reingold	C	C	C		C	C	C
Kamada Kawai	C	C	C			C	
Sugiyama			C				C
Spring layout	C	C		C			
Circular	C	C	C	C	C	C	C
Polar							C
Grid		C	C	C			C
Force atlas					C		
Random	C	C	C				C

### 3.4 File Formats

A social network graph can be represented in a variety of file formats. For example, a social graph can be represented in the form of an adjacency matrix, edge list, GML [18], GraphML [19] and various other formats. Table 4 identifies the different file formats that are supported by the major tools. The GraphML, GML and Pajek (.net) format is supported by most of the tools. However, adjacency matrix, edge list formats are supported only by SocnetV, networkx and igraph.

### 3.5 General Information

As a variety of tools are available for social network, it is always desirable to understand the general information pertaining to each tool such as the license, platform support and type of tool. This information is provided in Table 3. Most of the tools under discussion are updated and stable versions are released on a periodic basis. New features are being incorporated in every new release. Packages like networkx and igraph are available as part of R or Python Language whereas, socnetv, cytoscape and gephi are standalone softwares which can be downloaded and installed on any platform. Pajek is a standalone software designed especially for the Windows platform, while NodeXL is a windows based application integrated with Microsoft Excel.

**Table 4.** File formats supported by the various social networking tool

File format	SocnetV	networkx	igraph	Cytoscape	Gephi	Pajek	NodeXL
Graph ML	C	C	C	C	C	C	C
GML	C	C	C	C	C	C	
Pajek (.net)	C	C	C		C	C	C
UCINET (DL)	C						C
GEXF		C			C	C	
CSV					C	C	
DOT	C					C	
GXL						C	
LEDA		C					
JSON		C		C			
Adjacency matrix (.am)	C	C	C				
Edge list (.lst)	C	C	C				

## 4 Conclusion

Social Network Analytics is placed among various domains like sociology, mathematics, computer science, biological networks etc. Various types of visualization and analytical tools are available for different applications. How to choose a software varies from user to user. For example, Cytoscape is very efficient in visualization and integration of molecular interaction networks with genomic profiles and other state data. Pajek can be useful for analyzing complex networks with very large dataset but the user interface of Pajek is quite sophisticated. The igraph package and the networkx package are very useful from a command based GUI point of view. The igraph and networkx packages are updated periodically and built in algorithms are available for most of the graph operations. They are also very powerful for statistical computations.

SocnetV is an excellent tool for network analysis, given its clean and modern UI. It includes advanced features (Structural Equivalence, Hierarchical Clustering, and most Centrality metrics, FDP layouts), and primarily focus to help users understand what Social Network Analysis is. Tools like NodeXL (Microsoft) are very user-friendly and it only requires MS-EXCEL skills to handle medium-sized network datasets. NodeXL has the advantage of a simple UI design integrated with Microsoft Excel and provides facility to obtain the twitter datasets or facebook dataset of a user (Tables 2 and 5).

Thus, the choice of a social networking tool is up to the user, considering the type of application, complexity of the social network, ease of use or other parameters or features as already discussed in this paper.

**Table 5.** Information on license, type and platform

Tool	License	Type	Platform
SocnetV	GNU GPL3	Stand-alone	Multiple
networkx	BSD	Python package	Multiple
igraph	GNU GPL2	R package	Multiple
Cytoscape	GNU LGPL	Stand-alone	Multiple
Gephi	GNU GPL3	Stand-alone	Multiple
Pajek	Free for non commercial use	Stand-alone	Windows
NodeXL	Commercial user license	MS Excel integrated	Windows

## References

1. Wikipedia: Social network analysis. [https://en.wikipedia.org/wiki/Social\\_network\\_analysis](https://en.wikipedia.org/wiki/Social_network_analysis)
2. Easley, D., Kleinberg, J.: Cornell university: networks, crowds, and markets. In: Networks, Crowds, and Markets - Reasoning about a Highly Connected World. Cambridge University Press, July 2010
3. Bastian, M., Heymann, S., Jacomy, M.: Gephi: an open source software for exploring and manipulating networks. In: International AAAI Conference on Weblogs and Social Media, pp. 361–362 (2009)
4. Batagelj, V.: Pajek - program for large network analysis. *Connections* **21**(2), 47–57 (1988)
5. The Cytoscape Consortium: Cytoscape user manual 3.7.2. [http://manual.cytoscape.org/en/stable/Navigation\\_and\\_Layout.html](http://manual.cytoscape.org/en/stable/Navigation_and_Layout.html)
6. Hagberg, A.A., Schult, D.A., Swart, P.J.: Exploring network structure, dynamics, and function using NetworkX. In: 7th Python in Science Conference (SciPy 2008) (SciPy), pp. 11–15 (2008)
7. Ellson, J., Gansner, E., Koutsofios, L., North, S., Woodhull, G.: Graphviz – open source graph.pdf, pp. 3–4 (2000)
8. Hansen, D., Shneiderman, B., Smith, M.A.: Analyzing social media networks with NodeXL. In: Insights from a Connected World. Elsevier, August 2010
9. Wikipedia: Social network analysis software. [https://en.wikipedia.org/wiki/Social\\_network\\_analysis\\_software](https://en.wikipedia.org/wiki/Social_network_analysis_software)
10. Cytoscape Consortium: Cytoscape. [https://cytoscape.org/what\\_is\\_cytoscape.html](https://cytoscape.org/what_is_cytoscape.html)
11. Gephi: The open graph viz platform. <https://gephi.org/features>
12. Huisman, M., van Duijn, M.A.: Software for social network analysis. In: Models and Methods in Social Network Analysis, pp. 270–316 (2005)
13. NetworkX Developers: NetworkX: software for complex networks. [https://networkx.github.io/documentation/stable/auto\\_examples/drawing/plot\\_random\\_geometric\\_graph.html](https://networkx.github.io/documentation/stable/auto_examples/drawing/plot_random_geometric_graph.html)
14. Csardi, G., Nepusz, T.: The igraph software package for complex network research. *Inter-J. Complex Syst.* **1695**(5), 1–9 (2006)
15. Toshkov, D.: Network visualization in R with the igraph package. <http://re-design.dimiter.eu/?p=547>
16. Jacomy, A.: SCIENCESPO - MÉDIALAB: Sigma JS. <http://sigmajs.org/>
17. Marsden, P.V., Elsevier: measures of network centrality. In: International Encyclopedia of the Social & Behavioral Sciences, pp. 532–539. Science Direct (2015)

18. Himsolt, M.: GraphEd: a graphical platform for the implementation of graph algorithms (extended abstract and demo). In: Tamassia, R., Tollis, I.G. (eds.) Graph Drawing. GD 1994. Lecture Notes in Computer Science, vol. 1190, pp. 233–240. Springer, Heidelberg (1996). [https://doi.org/10.1007/3-540-58950-3\\_370](https://doi.org/10.1007/3-540-58950-3_370)
19. Brandes, U., Pich, C.: GraphML transformation. In: Pach, J. (ed.) GD 2004. LNCS, vol. 3383, pp. 89–99. Springer, Heidelberg (2005). [https://doi.org/10.1007/978-3-540-31843-9\\_11](https://doi.org/10.1007/978-3-540-31843-9_11)